



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

**Transcriptional response to prolonged acid exposure and
genetic divergence in *Helicobacter pylori***

A thesis submitted for the award of Doctor of Philosophy

By

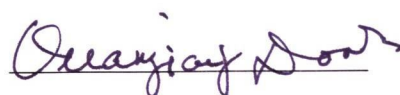
Quanjiang Dong

Department of Clinical Medicine
at Adelaide and Meath Hospital
Faculty of Health Science
University of Dublin
Trinity College

September, 2002

I hereby declare that this thesis has not been previously submitted for a Degree at this or any other University and that it represents my own unaided work except where duly acknowledged in the text.

I agree that this thesis may be lent or copied at the discretion of The Librarian, University of Dublin, Trinity College.

A handwritten signature in dark ink, appearing to read 'Quanjiang Dong', written over a horizontal line.

Quanjiang Dong

Acknowledgement

I wish to express my gratitude to my supervisor, Professor Colm O'Morain for giving me the opportunity to undertake the project and for his support and guidance throughout the three years of the study. I am extremely grateful to Dr. Martin Buckley, my supervisor, for his guidance, support and generosity. Without his encouragement and support, the thesis would not have been completed.

I would like to thank Professor Conor Keane from Department of Clinical Microbiology, for the opportunity to work in his laboratory. I would also like to thank Dr. Celine Herra at Sir Patrick Duns's Research Laboratory, St James's Hospital for her support and helpful discussion. Thanks also go to my colleagues in the SPD research laboratory for sharing their expertise and assistance.

I would like to thank Dr. Maria O'Sullivan and Dr. Wendy Hall, from Department of Gastroenterology, Adelaide and Meath Hospital, for editorial advice, discussion and help. I am grateful to Gwen Murphy and staff of Department of Gastroenterology.

I would like to thank Dr. Tony Moran and Professor Cyril Smyth for introducing me to the world of molecular biology.

I wish to express my sincere thanks to my parents and brothers for encouragement and support for my life. Finally, I would like to thank my wife who has supported, understood and encouraged me throughout.

Table of Contents

	Page
List of Figures -----	VII
List of Tables -----	IX
Summary -----	X
Abbreviations -----	XII
Publications -----	XV
 Chapter 1: General Introduction -----	 1
1.1 History of Discovery -----	2
1.2. Other <i>Helicobacter</i> species isolated from humans -----	3
1.3. Bacteriological Characteristics -----	3
1.4. Epidemiology -----	4
1.5. Consequences of <i>H. pylori</i> infection -----	7
1.6. Diagnosis -----	9
1.16.1. Invasive tests -----	9
1.16.2. Noninvasive tests -----	12
1.7. Treatment -----	13
1.8. General features of the genome -----	14
1.9. Virulence associated factors -----	15
1.9.1. Type IV secretion system and the effector molecule CagA -----	15
1.9.2. Vacuolating cytotoxin A (VacA) -----	16

1.9.3. Lipopolysaccharides (LPS) -----	17
1.9.4. Neutrophil-activating protein (HP-NAP) -----	17
1.9.5. Blood-group antigen-binding adhesin A (BabA) -----	18
1.9.6. Other virulence associated factors -----	18
1.10. Genetic diversity -----	19
1.11. Techniques for analysis of genomic diversity -----	21
1.12. Genotypes and clinical outcomes -----	22
1.12.1. Status of <i>cagA</i> -----	22
1.12.2. The <i>vacA</i> genotypes -----	23
1.12.3. The <i>iceA</i> genotypes -----	24
1.12.4. The combined genotypes -----	24
1.13. Transcriptional regulation -----	25
1.13.1. Transcription apparatus -----	26
1.13.2. Two-component system -----	29
1.13.3. Transcriptional regulation of the <i>cag</i> genes -----	31
1.13.4. Transcriptional regulation of <i>vacA</i> -----	32
1.13.5. Flagellar regulons -----	33
1.14. Adaptation and the consequences of <i>H. pylori</i> infection -----	35
1.14.1. Acid adaptation -----	36
1.14.2. Genetic divergence and adaptation -----	37
1.14.3. Genetic diversity and pathogenesis -----	38
1.15. Aims -----	39
Chapter 2: General Materials and Methods -----	40
2.1 General Materials -----	41

2.1.1. Molecular reagents -----	41
2.1.2. <i>H. pylori</i> strains -----	42
2.2. DNA Manipulation -----	43
2.2.1. Extraction of Genomic DNA -----	43
2.2.2. Purification and Concentration of DNA -----	45
2.2.3. Agarose Gel Electrophoresis -----	45
2.2.4. Polyacrylamide Gel electrophoresis -----	46
2.2.5. Labelling of DNA probes with Digoxigenin (DIG) -----	47
2.2.6. Southern Hybridisation -----	47
2.2.7. Polymerase Chain Reaction (PCR) -----	48
2.2.8. Cloning and sequencing -----	49
2.3. RNA manipulation -----	49
2.3.1. Northern Hybridisation -----	49
 Chapter 3: Identification of <i>Helicobacter pylori</i> genes	
regulated by prolonged acid exposure -----	51
3.1 Introduction -----	52
3.2. Materials and Methods -----	55
3.2.1. Strain and culture conditions -----	55
3.2.2 Differential display PCR -----	55
3.2.3 Re-amplification and cloning of putative candidates -----	56
3.2.4. DNA sequencing and analysis -----	57
3.2.5. Northern Blot -----	57
3.3. Results -----	58
3.3.1 Variation in morphology of colonies -----	58

3.3.2. Identification of differentially expressed genes	58
3.3.3. Nucleotide sequence analysis	59
3.4. Discussion	60
 Chapter 4: Divergence in the genomic contents among <i>Helicobacter pylori</i> strains	66
4.1. Introduction	67
4.2. Materials and Methods	69
4.2.1 Bacterial culture	69
4.2.2 Rapid subtraction hybridisation	69
4.2.3. Southern blot and PCR amplification	71
4.2.4. Statistical analysis	72
4.3. Results	73
4.3.1. Genotyping the tester and the driver strains	73
4.3.2. Identification of sequences specific for the tester strain C1	73
4.3.3. Analysis of the identified sequences specific to strain C1	74
4.3.4. Distribution of the identified strain-specific sequences among <i>H. pylori</i> strains	75
4.4. Discussion	77
4.4.1. Strain-specific genes	78
4.4.2. Chromosomal rearrangement	81
4.4.3. Allelic polymorphism	82
4.4.4. Extensive genomic divergence	82
 Chapter 5: Heterogeneity in the <i>cag</i> pathogenicity island of <i>Helicobacter pylori</i>	84

5.1. Introduction ----- 85

5.2. Materials and Methods ----- 88

5.2.1 *H. pylori* strains ----- 88

5.2.2 Polymerase chain reaction (PCR) ----- 88

5.2.3. Southern Hybridisation ----- 89

5.2.4. Cloning and sequencing ----- 89

5.2.4. Sequencing analysis and Statistical analysis ----- 89

5.3. Results ----- 90

5.3.1. Prevalence of the *cagA*⁺ strains ----- 90

5.3.2. Composition of the *cag* PAI ----- 90

5.3.3. Variation at the *cagA* 3' region ----- 91

5.3.4. Variation at the *cag7* 5' region ----- 93

5.3.5. The combined genotypes ----- 94

5.3.6. The Mixed populations ----- 95

5.4. Discussions ----- 96

5.4.1. Heterogeneity in the compositions of the *cag* PAI ----- 96

5.4.2. Variation at the *cagA* 3' region ----- 97

5.4.3. Variation at the *cag7* 5' region ----- 98

5.4.4. The combined genotypes ----- 99

5.4.5. The Mixed populations ----- 99

5.4.6. Conclusions ----- 100

Chapter 6: General Discussion ----- 101

6.1. Transcriptional response of *H. pylori* to acid exposure ----- 102

6.1.1 Groups of acid regulated genes ----- 102

List of Figures

Inserted after page

Figure 1.1. Alignment of the -10, extended -10 and -35 promoter sequences in 21 <i>H. pylori</i> genes -----	28
Figure 3.1. DD-PCR for <i>H. pylori</i> cells grown on pH 5.5 (a) and pH 7.4 (n) media for 5 successive passages -----	58
Figure 3.2. Northern blot analysis of total RNA from <i>H. pylori</i> cells grown on pH 5.5 (a) and pH 7.4 (n) media for 5 successive passages -----	58
Figure 3.3. Sequence of <i>pap57</i> and alignment with the consensus sequence of the EAL domain -----	59
Figure 4.1. Schematic illustration of the procedures of RaSH -----	68
Figure 4.2. Representative dot-blot analysis of products of subtractive hybridisation -----	73
Figure 4.3. Nucleotide sequence and the deduced amino acid sequence of Clone F5 (A), Clone P32 (B) and alignment of the amino acid sequence of Clone P32 with <i>Dichelobacter nodosus</i> GepA (C) -----	74
Figure 4.4. Schematic illustration of the locations of clones F10, F3 and F17 in the sequenced strains 26695 and j99 -----	74
Figure 4.5. Prevalence of clone P32 and F5 in strains from different diseases -----	75

Figure 5.1. Schematic illustration of the orientation and organisation of the <i>cag</i> PAI -----	86
Figure 5.2. Distribution of <i>cagA</i> ⁺ strains among different diseases -----	90
Figure 5.3. Representative agarose gel analysis of products of the <i>cagA</i> 3' region -----	91
Figure 5.4. Primary structure of the <i>cagA</i> 3' region -----	91
Figure 5.5. Representative agarose gel analysis of products of the <i>cag7</i> 5' region -----	93
Figure 5.6. Distribution of the <i>cag7</i> genotypes in different disease groups -----	94

List of Tables

	Inserted after page
Table 1.1. Non- <i>H. pylori</i> <i>Helicobacters</i> isolated from humans -----	3
Table 1.2. General comparative features of the <i>H. pylori</i> genomes -----	14
Table 1.3. Genes and crossed nomenclature of the <i>cag</i> PAI and effects of inactivation on IL-8 secretion, NF- κ B activation and CagA translocation -----	16
Table 1.4. Functions of the two-component systems in <i>H. pylori</i> -----	30
Table 3.1. Sequences of the 11-mer and 10-mer primers for DD-PCR -----	55
Table 4.1. The sequences of oligonucleotides used for RaSH and PCR -----	69
Table 4.2. Characteristics of the sequences specific to <i>H. pylori</i> strain C1 -----	74
Table 4.3. Distribution of the strain-specific sequences and the <i>vacA</i> genotypes among strains of <i>H. pylori</i> -----	75
Table 5.1. Status of <i>cagA</i> , <i>cag7</i> and <i>cag2</i> in strains from different disease groups -----	90
Table 5.2. Prevalence of the <i>cagA</i> 3' region genotypes in different patient groups -----	93
Table 5.3. Distribution of the combined types of <i>cagA</i> and <i>cag7</i> in different groups -----	94

Summary

Helicobacter pylori infection is strongly associated with gastritis, peptic ulcer and gastric adenocarcinoma. Molecular mechanisms involved in *H. pylori* adaptation and clinical consequences of infection are unclear. This study aimed to investigate transcriptional response to acid exposure and inter-strain genetic diversity.

The mRNA expression profiles of *H. pylori* cells grown on pH 5.5 and 7.4 media were compared using differential display PCR. Expressions of seven genes, including *ureB*, *flaA*, *topA*, *tufB*, *topA* and *pap57* (a novel sequence containing an EAL domain) and two sequences which had no significant homologues, were up-regulated in cells grown on acidic media. The mRNA level of *atoE*, however, was decreased. These differentially expressed genes may be involved in maintaining internal pH and counteracting acid damage on cellular biomolecules. The possibility that *topA* encoding topoisomerase A could function as a master-like regulator is also discussed.

The inter-strain diversities in genomic contents and the *cag* pathogenicity island were investigated using a number of molecular techniques. Several types of genetic diversity were identified. (1) Strain-specific genes that were possibly acquired by horizontal gene transfer or loss by genomic reduction. These strain-specific genes included two novel sequences (clone F5 and P32) and homologues of *resI* encoding a restriction enzyme ResI and JHP 704. (2) Chromosomal rearrangements that were possibly generated by accessory genetic elements. Three rearrangement loci including a locus of the “plasticity zone” randomly occurred in the clinical isolates. (3) Variation in the sizes of *cagA* and *cag7* that were generated by long DNA repetitive elements. A new segment (β segment) involved in *cagA* variation was identified. (4) Allelic polymorphisms generated by mutation. Non-synonymous mutation occurred in *cagA* resulting in loss of tyrosine phosphorylation sites. (5) Instability of the *cag*

pathogenicity island. Deletion of *cagA* and/or *cag7* was observed, which would conceivably abolish the function of the type IV secretion machinery in *H. pylori*. The status or the genotypes of nine genes/sequences, which represented these identified types of genetic diversity, were detected in 19 strains related to duodenal ulcer, intestinal metaplasia or gastric cancer. Each strain exhibited a unique distribution profile of these genes/sequences, indicating that extensive diversities exist within the *H. pylori* species.

Clone F5, a homologue of *Dichelobacter nodosus* GepA, was absent from the examined strains from patients with intestinal metaplasia, but present in a significantly higher fraction of strains from duodenal ulcer patients. The prevalence of the *cag7* B-*cagA* γ_1 type strains was higher in patients with duodenal ulcer than that in patients with gastritis. The results suggested that these genetic traits (status of Clone F5 and the *cag7* B-*cagA* γ_1 genotype) may be used as markers for predicting clinical outcomes of *H. pylori* infection.

This study is the first large scale investigation on the transcriptional response of *H. pylori* to prolonged acid exposure. Tremendous genetic divergences were observed in *H. pylori*. These results are useful in the understanding of bacterial adaptation and pathogenesis.

Abbreviations

aa	Amino acid
BHI	Brain heart infusion
bp	Base pair
<i>cag</i>	Cytotoxin associated gene
CSPD	Disodium 3-{4-methoxyspiro [1,2-dioxetane-3,2'-(5'-chloro)tricyclo(3.3.1.1 ^{3,7})decaan]-4-yl}phenyl phosphate
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dNTP	Deoxyadenosine triphosphate
DU	Duodenal ulcer
DNase	Deoxyribonuclease
EAL	Glutamic acid (E)-alanine (A)-leucine (L)
EBI	European Bioinformatics Institute
EDTA	Ethylenediamine-tetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
g	Gram
GC	Gastric cancer
G+C	guanine and cytosine
h	Hour
IS	Insertion sequence
IL	Interleukin
IM	Intestinal metaplasia
IS	Insertion sequence
kb	Kilobase pairs

kDa	Kilodalton
L	Litre
LiPA	Line probe assay
LPS	Lipopolysaccharide
M	Molar
Mb	Megabase
μ	Micro
μ g	Microgram
μ l	Microlitre
μ mol	Micromole
mg	Milligram
M-MuLV	Molonety murine leukemia virus
min	Minute
ml	Millilitre
NCTC	National Collection of Type Cultures
NUD	Non-ulcer dyspepsia
ORF	Open reading frame
PAI	Pathogenicity island
PAGE	Polyacrylamide gel electrophoresis
PFGE	Pulsed field gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
RAPD	Random amplified polymorphic DNA
RaSH	Rapid subtraction hybridisation

RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
RT	Room temperature
RUT	Rapid urease test
SDS	Sodium dodecyl sulphate
sec	Second
SHP	Src Homologous protein
SSC	Saline-sodium citrate buffer
TAE	Tris acetate buffer
TBE	Tris borate buffer
TE	Tris EDTA buffer
TPS	Tyrosine phosphorylation site
vol.	Volume
v/v	Volume per volume
w/v	Weight per volume
Y	Tyrosine

CHAPTER 1
GENERAL INTRODUCTION

1.1. History of Discovery

Helicobacter pylori (*H. pylori*), a spiral-shaped bacterium, was first isolated from the human stomach by Warren and Marshall in 1983 (Warren 1983; Marshall, 1983). However, this was not the first observations that spiral organisms were present in the stomach (Buckley and O'Morain, 1998). In 1893, Bizzozzero described the presence of gastric spiral bacteria in dogs. The bacteria were also found in the stomach of cats and mice in 1896. In 1906, Krienitz provided the first evidence of spiral bacteria in humans when examining the gastric specimens from a patient with gastric cancer. In 1938, Doenges reported spirochaetes in 40% of human stomachs *post mortem*. In 1954, Palmer reported the absence of bacteria in biopsies from 1,180 human stomachs, and thereafter, a general view was established that the bacteria seen in the stomachs were contaminants. In the 1970s several papers continued to report the presence of bacteria in the stomachs of humans and associated the presence with gastric ulcers and chronic gastritis (Fung *et al.*, 1979). Finally, Marshall and Warren succeeded in culturing *H. pylori* from gastric biopsy specimens in 1983 and correlated the presence of the bacterium with gastric diseases. This was an important event in the acceptance of this spiral bacterium's role in gastric pathophysiology.

Originally, the spiral bacterium was named as *Campylobacter pyloridis* since it resembles other *Campylobacters* and is usually found in the pylorus of the stomach (Marshall and Warren, 1984). In order to conform to Latin grammar, the bacterium was renamed as *Campylobacter pylori* (Marshall and Goodwin, 1987). The bacterium was finally given the name *H. pylori* in 1989 and was recognized as a new bacterial species (Goodwin *et al.*, 1989).

1.2. Other *Helicobacter* species isolated from humans

The species of *Helicobacter pylori* is a member of the *Helicobacter* genus. This genus belongs to rRNA homology group VI that includes *Campylobacter*, *Arcobacter*, *Wolinella* and *Flexispira* (Vandamme *et al.*, 1991). The number of species in the genus *Helicobacter* has rapidly expanded since the identification of *H. pylori*. Currently, at least 24 formally named *Helicobacters* have been identified and an additional 35 or more novel *Helicobacters* await formal naming (Fox, 2002). Ten of these *Helicobacter* species have been isolated from humans (listed in Table 1.1).

1.3. Bacteriological Characteristics

H. pylori is a Gram-negative, slow-growing bacterium that colonizes the human stomach and is often found beneath the mucus layer. Isolation of *H. pylori* is achieved under microaerophilic conditions in an atmosphere of 5-10% CO₂. However, some strains may grow aerobically (Xia *et al.*, 1994). The bacterium grows over a temperature range of 33 to 40 °C but optimally at 37 °C (Goodwin and Armstrong, 1990). *H. pylori* is neutrophilic and grows within pH range 5.5-8.5 *in vitro*, but grows optimally between pH 6.9 and 8.0 (Owen, 1998). Typically, the colonies of *H. pylori* are grey, convex, circular (1-2 mm in diameter) and translucent. When cultured in liquid medium, the cells may enter stationary phase after 24 h or 48 h culture (Joyce *et al.*, 2000; Worku *et al.*, 1999).

When observed *in vivo* the bacterium has a spiral morphology (0.5-1.0 µm wide and 2-4 µm long) but frequently has a curved rod-like appearance *in vitro*. The bacteria in older cultures or those growing in nutrient deficient media can transform from normal bacillary morphology to a coccoid morphology (Young *et al.*, 2001; Worku *et al.*,

Table 1.1. Non-*Helicobacter pylori* Helicobacters isolated from humans^a

Species	Other hosts	Primary site	Other sites
<i>H. rappini</i> ^b	Sheep, dog, mice	Intestine	Blood (humans); liver (sheep); stomach (dogs)
<i>H. canis</i>	Dog, cat	Intestine	Blood (humans); liver (dogs)
<i>H. cinaedi</i> ^b	Hamster, rhesus monkey, dog	Intestine	Blood, soft tissue, joints (humans); liver (monkey)
<i>H. fennelliae</i>	Dog, macaque	Intestine	Blood
<i>H. pullorum</i> ^b	Chicken	Intestine	Liver (chicken)
<i>H. canadensis</i>	NR	Intestine	NR
<i>H. westmeadii</i>	NR	NR	Blood
<i>H. winghamensis</i>	NR	Intestine	NR
<i>H. heilmannii</i> ^b	Dogs, cats, monkeys, cheetahs, wild rats, swine	Stomach	NR
<i>H. felis</i>	Dogs, cats, cheetahs	Stomach	NR

^a Adapted from Fox, 2002; ^b some data suggest zoonotic potential; NR, not recorded

1999). It is unclear whether the coccoid form is viable or not. Kusters *et al* (1997) reported that this form was a morphological manifestation of bacterial cell death (Kusters *et al.*, 1997). Others provided evidence that coccoid *H. pylori* cells are actually viable and can revert to the spiral morphology (Andersen *et al.*, 1997; Wang *et al.*, 1997).

In artificially contaminated milk, *H. pylori* could survive for up to 6 days at 4°C and 3-4 days at room temperature (RT). The bacterium does not survive any length of time in yoghurt (Vaira *et al*, 2001). When suspended in saline or chicken stock, the bacterium survives 2 days at RT and 3 days if refrigerated, but at -20°C it survives at least 4 weeks. Surface survival on wood, plastic and ceramic has also been investigated and viability is maintained from 30 to 90 min only, but if cooled survival was 3 h (Vaira *et al*, 2001).

H. pylori produces substantial amounts of urease which is used as a biochemical marker for identifying the bacterium (Megraud *et al.*, 1985). Other enzymes including catalase, oxidase, acid phosphatase and alkaline phosphatase are also produced by the bacterium (Megraud *et al.*, 1985).

1.4. Epidemiology

H. pylori is one of the most common gastrointestinal pathogens, infecting almost 50% of the human population (Taylor and Blaser, 1991). In developing countries, the bacterium infects 70 to 90% of the population, while in the developed world the prevalence of infection is much lower at approximately 25% (Dunn *et al.*, 1997).

Although several workers have isolated *H. pylori* from domestic cats (Fox *et al.*, 1995) and non-human primates (Dubois *et al.*, 1994) and have detected it in water sources (Hulten *et al.*, 1996), the only significant reservoir of *H. pylori* appears to be humans (Megraud and Broutet, 2000). The human gastric epithelium is the preferred environment for *H. pylori* but the bacterium has been shown to infect other epithelia such as the duodenum, oesophagus and the rectum (Sahay and Axon, 1996). However, in each of these cases heterotopis (presence of tissue in an abnormal location) and gastric metaplasia (transformation of gastric tissue) were prerequisites. *H. pylori* has been found in the lower oesophagus as a result of gastrooesophageal reflux disease and has also been cultured from gastric juices (Sahay and Axon, 1996; Newton *et al.*, 1997; Weston *et al.*, 2000). However, more significantly, the bacterium has also been isolated and detected with varying success from both human faeces and saliva (Sahay *et al.*, 1995; Thomas *et al.*, 1992).

Peak acquisition of *H. pylori* occurs in childhood (Amarri and Celinska-Cedro, 2001). However, the prevalence of infection in the developed world increases with age, gradually rising from infancy to an age of about 60 years (Calam, 1996). This suggests that in the developed world more elderly people are infected than children. The explanation for this paradox lies in the finding that the incidence of *H. pylori* infection has declined in the developed world due to improvement in living conditions and sanitation (Banatvala *et al.*, 1993; Taylor and Parsonnet, 1995).

There is a significant correlation between seropositivity and the ethnic origin of the population (Dwyer, *et al.*, 1988). This may be related to, among other factors, cultural practices or the relative lack of a public health infrastructure and social deprivation

(Dwyer *et al.*, 1988). Increased prevalence of *H. pylori* colonisation was found in individuals with lower family incomes (Fiedorek *et al.*, 1991). Also, demographic factors (urban more than rural), lack of a fixed hot-water supply and domestic overcrowding were demonstrated to be correlated to colonization by *H. pylori* (Elitsur *et al.*, 1998; Mendall *et al.*, 1992; Mitchell *et al.*, 1992).

Close personal contact is important for transmission of *H. pylori* (Malaty *et al.*, 1991; Miyaji *et al.*, 2000). A study of 1684 individuals indicated that seropositivity in children was more common in those whose mothers or older siblings were seropositive or in those who attended a nursery school (Miyaji *et al.*, 2000). Another study also revealed that 40% of children of *H. pylori*-positive parents were colonized by *H. pylori*, compared with 3% of children of *H. pylori*-negative parents (Malaty *et al.*, 1991). Transmission between couples in close contact, however, was unlikely (Perez-Perez *et al.*, 1991; Polish *et al.*, 1991).

H. pylori is possibly transmitted via the gastro-oral route and the faecal-oral route (Vaira *et al.*, 2001). The oral route of infection is consistent with the epidemiological transmission during childhood, and is supported by data from animals (Lee *et al.*, 1991). Gastric secretions or saliva are probably the vehicle of transmission (Axon, 1995; Polonczyk *et al.*, 1996). Faecal-oral spread can occur either directly from an infected patient as in occupational exposure or from close contact (Rudi *et al.*, 1997; Thomas *et al.*, 1992; Kelly *et al.*, 1994), or indirectly from the environment through the vehicle of sewage contaminating food or water (Klein *et al.*, 1991; Lindkvist *et al.*, 1995).

1.5. Consequences of *H. pylori* infection

It is well established that *H. pylori* infection causes chronic active gastritis and is strongly associated with the development of duodenal ulcer and gastric ulcer diseases and gastric mucosa-associated lymphoid tissue lymphoma (MALToma) (Cover and Blaser, 1992; Moran, 1997; Malfertheiner *et al.*, 2002). It is a group I carcinogen for distal gastric adenocarcinoma (Parsonnet *et al.*, 1991).

Patients infected with *H. pylori* develop two different phenotypes of chronic gastritis with differing disease associations: antral predominant gastritis and pangastritis (Sepulveda *et al.*, 2002). Antral predominant gastritis rarely progresses over time into the body of the stomach. It is associated with the development of duodenal ulcer in a subgroup of patients. Pangastritis is characterised by mucosal inflammation in both antrum and corpus. Patients with pangastritis are prone to development of gastric ulcers and progression to atrophy (multifocal atrophic gastritis). Intestinal type gastric adenocarcinoma may ultimately develop in a small fraction of patients with pangastritis progressing through intestinal metaplasia and dysplasia. Patients with antral predominant gastritis and duodenal ulcer are relatively protected from gastric adenocarcinoma whereas those with pangastritis and gastric ulcer have a predisposition to the development of gastric adenocarcinoma (Molloy and Sonnenberg, 1997; Asaka *et al.*, 2001). It is hypothesized that the level of acid secretion is important in determining whether antral-predominant gastritis or pangastritis will develop (Lee *et al.*, 1995). Individuals with high acid output will have sparing of the gastric body and such patients are at risk for duodenal ulceration. In patients with lower levels of acid secretion, *H. pylori* may be more likely to cause inflammation in the gastric body and hence, pangastritis.

Recent reports provide further evidence confirming the association between *H. pylori* infection and the development of gastric adenocarcinoma. Gastric adenocarcinoma was induced in 37% of specific pathogen-free (SPF) Mongolian gerbils inoculated with *H. pylori* (Watanabe *et al.*, 1998). Histological examination of the stomach of SPF gerbils inoculated with *H. pylori* at different intervals revealed the sequential process of active gastritis-intestinal metaplasia-gastric adenocarcinoma. A prospective study from Japan (Uemura *et al.*, 2001) observed 1,166 patients, who underwent endoscopy and follow-up. The study revealed that *H. pylori* infection was a prerequisite for the development of gastric adenocarcinoma and *H. pylori*-positive patients with duodenal ulcer appeared to be at very low risk for the development of gastric adenocarcinoma. Regression of intestinal metaplasia was found in some patients after eradication of *H. pylori* (Asaka *et al.*, 2001; Correa *et al.*, 2000; Ohkusa *et al.*, 2001; Uemura *et al.*, 1997).

Intricate relationships between *H. pylori* infection and oesophageal diseases including gastro-oesophageal reflux disease (GORD) and oesophageal adenocarcinoma have been suggested (McNamara and O'Morain, 1999). *H. pylori* may offer protection against GORD and oesophageal adenocarcinoma. The suggestion comes from the facts including (1) a pronounced inverse relationship between infection rate of *H. pylori* and prevalence of these diseases (Chow *et al.*, 1998; Varanasi *et al.*, 1998); (2) Higher incidence of GORD after eradication of *H. pylori* (Hirschl *et al.*, 1994; Schwizer *et al.*, 2001). However, observations contrary to those aforementioned have been reported (Malfertheiner *et al.*, 2002; McNamara and O'Morain, 1999). Prospective, multicentre studies and studies in animal models are required to clarify the relationship.

H. pylori or other helicobacter species may be related to extra-gastrointestinal diseases. The association of *H. pylori* with any of these diseases, however, has not been clearly demonstrated (Vaira *et al.*, 2002). These diseases include hepato-biliary diseases, inflammatory bowel disease, coronary heart disease, respiratory, neurological, skin, autoimmune, growth and miscellaneous disorders (Tsang *et al.*, 1999; Vaira *et al.*, 2002).

1.6. Diagnosis

Numerous diagnostic tests are available for detecting *H. pylori* infection. These tests are based on either an invasive endoscopic procedure or non-invasive tests on blood samples, saliva and stool samples (Leodolter and Megraud, 1999). No single test is ideal for the detection of *H. pylori*, and hence, a combination of the results derived from two tests with acceptable sensitivity and specificity is recommended for accurate diagnosis (Vaira *et al.*, 2002).

1.6.1. Invasive tests:

Culture: Detection of *H. pylori* by culture is considered as a “gold standard” for diagnosis. This method has a specificity of 100%. The sensitivity of this test ranges from 77-95% and was suggested to depend on the adequate storage and transport of samples, antibiotic administration and personnel expertise (Dunn *et al.*, 1997). The biopsies are usually carried in special transport media such as nutrient broth, Brucella broth, Stuart’s medium, 20% glucose, chocolate agar slants, normal saline or commercially available media (Han *et al.*, 1995; Kjoller *et al.*, 1991; Siu *et al.*, 1998; van-der-Hulst *et al.*, 1996; Veenendaal *et al.*, 1993). To achieve a high rate of recovery, the specimens are maintained under 4 °C (Han *et al.*, 1995). The specimens, transported

properly from the endoscopy units to the laboratory, are smeared on the agar surface with the lumen side face to the agar. Otherwise, the specimens are homogenized and inoculated onto the agar plates. Media for isolation of *H. pylori* from gastric biopsies include Mueller-Hinton agar, Columbia agar, Wilkins Chalgren agar and Brain Heart Infusion agar that are supplemented with horse, sheep or human blood at concentrations ranging from 5 to 10% (V/V) (Ansorg *et al.*, 1991; Fresnadillo-Martinez *et al.*, 1997; Tee *et al.*, 1991; van der Hulst *et al.*, 1996; Xia *et al.*, 1993). To increase the sensitivity or minimize the risk of possible contamination, selective medium is a better choice in which vancomycin, nalidixic acid, trimethoprim, or amphotericin B are selectively supplemented in the media (Fresnadillo-Martinez *et al.*, 1997). A combination of selective and non-selective media would increase sensitivity (Piccolomini *et al.*, 1997; Tee *et al.*, 1991). Growth of the bacteria is at 37 °C under microaerobic environments in 98% humidity for five to ten days. Extended incubation may be required if the biopsies are taken from patients who have been administrated anti-*H. pylori* treatment. An advantage of culture is that the isolates recovered from the specimens can be used for antimicrobial susceptibility testing and detection of virulence factors, thus, benefiting the clinical management of *H. pylori* infection (Vaira *et al.*, 2002).

Rapid urease test (RUT): *H. pylori* produces large amounts of urease which catalyzes urea to ammonia and carbonic acid, and thus, increases pH (Moran *et al.*, 2002). RUT is based on the detection of an increase in pH in the gastric specimens. In untreated patients, the specificity of RUT ranges from 80 to 95% while the sensitivity range is 95 to 100% (Midolo and Marshall, 2000). For patients that have been administrated with anti-*H. pylori* treatment, however, RUT is not recommended due to decreased sensitivity and specificity (Vaira *et al.*, 2002).

Histology: Histological identification of *H. pylori* is widely used as it can simultaneously examine the pathology of gastric mucosa and the bacterium. Several staining methods have been used for the detection of *H. pylori* including haematoxylin and eosin staining, modified Giemsa, Warthin-Starry, modified McMullen's staining and immunohistochemical examination with anti-*H. pylori* antibodies (Fallone *et al.*, 1997; Madan *et al.*, 1988; Jonkers *et al.*, 1997; Rotimi *et al.*, 2000; Shimizu *et al.*, 1996; Simor *et al.*, 1990; Tazawa *et al.*, 1998). The sensitivity of histological detection of *H. pylori* is generally 93-98%, while the specificity is 95-98% (Dunn *et al.*, 1997).

PCR related techniques: Detection of *H. pylori* by PCR has a sensitivity of 85-96% and a specificity of 90-100% (Dunn *et al.*, 1997). PCR for this purpose is usually carried out by amplifying housekeeping genes, such as *ureC*, using specific primers (Lu *et al.*, 1999; Kawamata *et al.*, 1996). PCR restriction fragment length polymorphism (PCR-RFLP), PCR single strand conformational polymorphism (PCR-SSCP) and line probe assays (LiPA) have a good resolution of fingerprinting of the infected strains (Matsuoka *et al.*, 1999). Point mutation at 23S rRNA gene (usually A-G) resulting in clarithromycin resistance can be detected by PCR, PCR-RFLP (Sevin *et al.*, 1998) and LiPA (Ryan *et al.*, 2001). The results from PCR amplification for detection of *H. pylori*, however, should be interpreted cautiously considering the extensive genetic diversity of *H. pylori*. The target genes and the target regions on these genes to be amplified should be carefully selected to increase the specificity and also simultaneous amplification of multiple loci will certainly reduce the false positive rate.

Fluorescent in-situ hybridization: This technique was recently developed for detection of *H. pylori* (Russmann *et al.*, 2001; Trebesius *et al.*, 2000). The results can be obtained in 3 hours with excellent correlation with other diagnostic techniques. The technique uses a set of fluorescent labeled probes specifically annealed to the 16S ribosomal RNA gene (for detection of *H. pylori*) or 23S ribosomal RNA gene (for detection of clarithromycin resistance).

1.6.2. Noninvasive tests

Serological tests: The tests were the first noninvasive techniques based on the detection of a specific anti-*H. pylori* IgG antibody in patients' sera. Others include antibodies to CagA, VacA and Hsp60. The accuracy of serological tests may vary according to the antibodies to be detected, methods used and whether or not the patients had received anti-*H. pylori* treatment (Loy *et al.*, 1996). The sensitivity and specificity of the tests ranges from 88-95% and 86-95% respectively (Dunn *et al.*, 1997). The sensitivity and specificity of the ELISA test performed in an Irish population was 82.4% and 85%, respectively (Breslin *et al.*, 2000). Serological test are convenient and low-cost and can play an important role in clinical practice, but do need to be validated in local laboratories (Leodolter and Megraud, 1999).

Urea breath test (UBT): The principle of the UBT is based on the CO₂ produced from urea which is catalyzed by *H. pylori* urease. Following ingestion of ¹³C or ¹⁴C labeled urea, the CO₂ produced by *H. pylori* urease passes through mucosa into the blood and can be detected in the breath. The sensitivity and specificity of UBT are 97% and 95%, respectively (Vaira *et al.*, 2002). The accuracy of the test is altered in patients with

gastric surgery, patients treated with H₂ antagonists or proton pump inhibitors (Chey *et al.*, 1997) and patients with bleeding peptic ulcer disease.

Stool antigen test: An enzymatic immunoassay was developed recently which detects the presence of the *H. pylori* antigens in stool samples (Vaira *et al.*, 1999). The sensitivity before treatment is 93.1% and the specificity is 92.8% (Vaira *et al.*, 2001). This test is considered as an alternative for UBT, although its diagnosis of *H. pylori* is not convincing in post-treated patients (Malfertheiner *et al.*, 2002).

1.7. Treatment

H. pylori infection is associated with a wide range of gastroduodenal diseases (Malfertheiner *et al.*, 2002). Eradication of the bacterium, however, is only recommended in a subgroup of patients infected with *H. pylori*. Currently it is universally accepted that anti-*H. pylori* treatment should be used in gastric and duodenal ulcer diseases, gastric MALToma, atrophic gastritis, post-gastric cancer resection, and in first-degree relatives of gastric cancer (Qasim and O'Morain, 2002).

The most recent eradication guidelines proposed at the Maastricht 2-2000 recommend the use of proton pump inhibitors along with clarithromycin and amoxycillin or metronidazole for a minimum of 7 days (Malfertheiner *et al.*, 2002). The combination of amoxycillin and clarithromycin is preferred because it may favour best results with a second-line proton pump inhibitor quadruple therapy. Eradication rates of up to 88-95% have been achieved using current first-line regimens (Miwa *et al.*, 2000; Veldhuyzen *et al.*, 2000). The recommended second-line therapy includes a combination of a proton pump inhibitor with bismuth salt (subsalicylate/subcitrate),

metronidazole, and tetracycline for a minimum of 7 days. Extended proton pump inhibitor-based triple therapy can be used if bismuth is not available.

The most common reason for primary treatment failure is poor compliance with the eradication guidelines (Lee *et al.*, 2000). Antibiotic resistance may be a significant factor for treatment failure in certain geographical areas. The genetic background of patients may be an additional factor contributing to treatment failure (Furuta *et al.* 2001). The proton pump inhibitors are mainly metabolised by cytochrome P450 2C19 in the liver. The genotypes of this enzyme affected the cure rate of the treatment regime with proton pump inhibitors (Furuta *et al.* 2001).

1.8. General features of the genome

Two *H. pylori* strains, 26695 and J99, have been sequenced (Alm *et al.*, 1999; Tomb *et al.*, 1997). Strain 26695 was originally isolated from a patient with gastritis. This strain is able to colonise piglets, elicit an immune and inflammatory response and is also known to be toxigenic, plasmid-free and naturally competent (Akopyants *et al.*, 1995; Eaton *et al.*, 1989 & 1992; Tomb, *et al.*, 1997). Strain J99 was originally isolated from a patient with duodenal ulcer disease (Alm, *et al.*, 1999). The general features of the two genomes are shown in Table 1.2. Each strain contains a circular chromosome with a size of approximately 1.67 Mega-bp in 26695 and 1.64 Mega-bp in J99. The average G + C content of the chromosome is 39%. Ninety-one percent of the genome are protein coding regions, whereas 6.0% are intergenic regions. In strain 26695, 1,590 open reading frames (ORF) have been predicted, whereas 1,495 ORFs have been predicted in strain J99. The true number of the open reading frames in *H. pylori* genome, however, might be less than that of the original annotations based on the

Table 1.2. General comparative features of the *H. pylori* genomes

Genome features	<i>H. pylori</i> 26695	<i>H. pylori</i> J99
Size (base pairs)	1,667,867	1,643,831
(G+C) content (%)	39	39
Regions of different G/C content	8	9
AGAGATT repeats at bp=1	26	2
<i>vacA</i> genotypes	<i>slb/ml</i>	<i>sla/ml</i>

Open reading frames		
Percent of coding region	91.0	90.8
Predicted number	1,590	1,495
Functionally classified	875	895
Conserved with unknown function	275	290
<i>H. pylori</i> specific	345	367
Number with signal sequence	517	502
Average length (bp)	954	998
Percent of AUG initiation codons	81.8	82.7
Percent of GUG initiation codons	9.7	6.7
Percent of UUG initiation codons	8.1	10.4
Percent of other initiation codons	0.4	0.2

Insertion elements		
Complete <i>IS605</i> copies	5	0
Partial <i>IS605</i> copies	8	5
Complete <i>IS606</i> copies	2	1
Partial <i>IS606</i> copies	2	2

RNA elements		
Percent of genome (stable RNA)	0.75	0.75
23S-5S rRNA	2(3)	2
16s rRNA	2	2
tRNAs	36	36

Adapted from Alm *et al.*, 1999

analysis of the length distribution of the conserved protein-coding genes (Skovgaard *et al.*, 2001). Two-thirds of the predicted ORFs are of known functions revealed largely by analysis of their similarity to counterparts in other species.

1.9. Virulence genes

The bacterium produces several virulence genes that facilitate colonisation, induce inflammatory cell infiltrates and intoxicate epithelial cells. These virulence genes are discussed below.

1.9.1. Type IV secretion system and the effector molecule cagA

Bacteria possess type I, II, III and IV secretion systems to secrete toxins or translocate protein/protein complexes into host cells. The structure of the type IV secretion system resembles the conjugative pilus (Baron *et al.*, 2002). This secretion system is able to translocate DNA or proteins into host cells (Censini *et al.*, 2001). At least nine bacterial species such as *Agrobacterium tumefaciens* and *H. pylori* possess a type IV secretion system. The genes coding for the type IV secretion system in *H. pylori* are located in the *cag* pathogenicity island (PAI) (Censini *et al.*, 1996). Products of the *cag* PAI are primarily involved in assembling this secretion machinery. The functions of this system include: (1) induction of pro-inflammatory cytokine interleukin-8 (IL-8) release from gastric epithelial cells (Segal *et al.*, 1997). This induction is probably mediated by translocation of an unidentified protein into host cells or by contact of the machinery to the membrane of host cells (Fischer *et al.*, 2001). The cytokine IL-8 is a potent chemotactic and activating factor for neutrophils; (2) translocation of the cytotoxin associated gene A protein (CagA), the effector molecule, into the epithelial cells (Backert *et al.*, 2000; Odenbreit *et al.*, 2000; Stein *et al.*, 2000). CagA is a 128 kDa

protein encoded by *cagA* located at the right end of the *cag* pathogenicity island (Censini *et al.*, 1996; Covacci *et al.*, 1993). The translocated *cagA* is activated by phosphorylation at tyrosine residues (Odenbreit *et al.*, 2000). The host c-Src kinases are responsible for this process (Selbach *et al.*, 2002; Stein *et al.*, 2002). Subsequently the activated CagA binds to SHP-2 and induces growth factor-like cell changes characterised by actin polymerisation, elongation and spreading of cells (Higashi *et al.*, 2002; Stein *et al.*, 2000). The functions of *cag* genes are listed in Table 1.3. The majority of *cag* genes are essential in induction of IL-8 secretion and translocation of *cagA*. However, several genes (HP0520, HP 0521, HP0535, HP0536) are not required in these functions. HP0526, HP0538, HP0545 genes are partially required for both functions (Fischer *et al.*, 2001).

1.9.2. Vacuolating cytotoxin A (VacA)

H. pylori secretes a unique vacuolating cytotoxin (VacA, Atherton *et al.*, 1995). VacA is produced as a 140 kDa protoxin, which subsequently undergoes both N-terminal and C-terminal processing and yields an ~90 kDa mature VacA toxin (Nguyen *et al.*, 2001). It is secreted into the extracellular space and is also partially retained on the bacterial cell surface. VacA purified from broth culture appears as a large oligomeric complex with a mass >900 kDa (Cover and Blaser, 1992). Individual VacA oligomer has a flower or snowflake-like appearance (Iwamoto *et al.*, 1999). The secreted VacA is able to associate itself with VacA receptors on the eukaryotic cell surface and thus binds to the cell membrane (Seto *et al.*, 1998; Yahiro *et al.*, 1997). VacA induces formation of vacuoles containing both late endosomal and lysosomal markers (Molinari *et al.*, 1997; Papini *et al.*, 1998). VacA also induces pore formation in the plasma membrane and apoptosis in various epithelial cell lines (Peek *et al.*, 1999).

Table 1.3. Genes and crossed nomenclature of the *cag* PAI and effects of inactivation on IL-8 secretion, NF-κB activation and CagA translocation.

Nomenclature of the <i>cag</i> genes			Properties of individual genes inactivation		
Tomb	Akopyants	Censini (homology)	IL-8 secretion	NF-κB activation	CagA translocation
HP0520 <i>cag1</i>	ORF6	<i>cag</i> ζ	++	ND	++
HP0521 <i>cag2</i>	ORF7	<i>cag</i> ε	++	ND	++
HP0522 <i>cag3</i>	ORF8	<i>cag</i> δ	-	ND	-
HP0523 <i>cag4</i>	ORF9	<i>cag</i> β (<i>virD4</i>)	-	ND	-
HP0524 <i>cag5</i>	ORF10	<i>cag</i> α (<i>virB11</i>)	++	ND	-
HP0525	ORF11	<i>cag</i> Z	-	ND	-
HP0526 <i>cag6</i>	ORF12	ND	+	ND	+
HP0527 <i>cag7</i>	ORF13-14	<i>cag</i> Y (<i>virB10</i>)	-	ND	-
HP0528 <i>cag8</i>	ORF15	<i>cag</i> X (<i>virB9</i>)	-	ND	-
HP0529 <i>cag9</i>	ORF16	<i>cag</i> W (<i>virB8</i>)	-	ND	-
HP0530 <i>cag10</i>	ORF17	<i>cag</i> V	-	ND	-
HP0531 <i>cag11</i>	ORF18	<i>cag</i> U	-	ND	-
HP0532 <i>cag12</i>	ORF19	<i>cag</i> T (<i>virB7</i>)	-	ND	-
HP0534 <i>cag13</i>	ORF20	<i>cag</i> S	++	ND	+
HP0535 <i>cag14</i>	ORF21	<i>cag</i> Q	++	ND	++
HP0536 <i>cag15</i>	ORF22	<i>cag</i> P	++	ND	++
HP0537 <i>cag16</i>	ND	<i>cag</i> M	-	-	-
HP0538 <i>cag17</i>	ND	<i>cag</i> N	+	+	+
HP0539 <i>cag18</i>	ND	<i>cag</i> L	-	-	-
HP0540 <i>cag19</i>	ND	<i>cag</i> I	++	-	-
HP0541 <i>cag20</i>	ND	<i>cag</i> H	-	-	-
HP0542 <i>cag21</i>	ND	<i>cag</i> G	+	-	-
HP0543 <i>cag22</i>	ND	<i>cag</i> F	+	-	++
HP0544 <i>cag23</i>	ND	<i>cag</i> (<i>virB4</i>)	-	-	-
HP0545 <i>cag24</i>	ND	<i>cag</i> D	+	-	+
HP0546 <i>cag25</i>	ND	<i>cag</i> C	-	-	-
HP0547 <i>cag26</i>	ND	<i>cag</i> A	+/-	-	-

The list is based on the references from Covacci and Rappuoli, 2000 and Fischer *et al*, 2001. Twenty-seven *cag* genes are listed in the table. Other genes are excluded, as they are unlikely to encode proteins or to be involved in pathogenesis. These genes include HP0533, *tnpA*, *tnpB*, *cagR*, *cagO*, *cagB*, HP0548, HP0549 (*glr*). ++ efficient translocation or Il-8 secretion; + partial; - loss of functions; +/- uncertain; ND, not determined.

1.9.3. Lipopolysaccharides (LPS)

H. pylori LPS is composed of lipid A, an oligosaccharide core and the antigenic O-polysaccharide chain. Compared with LPS of other bacteria, *H. pylori* LPS and lipid A induce low immunological activities in a range of test systems (Moran, 1999). Nevertheless, these reduced levels of LPS-induced cytokines and toxic oxygen may facilitate the persistence of the bacterium at the gastric mucosal surface (Moran, 1999). Fucosylated oligosaccharides in *H. pylori* LPS mimic Lewis x and Lewis y blood group antigens present in humans (Appelmelk *et al.*, 1996; Monteiro *et al.*, 1998; Moran *et al.*, 1996; Moran, 1999; Moran and Prendergast, 2001). Approximately 85% of *H. pylori* isolates express at least one of the Lewis antigens (Moran and Prendergast, 2001; Wirth *et al.*, 1996). The molecular mimicry of *H. pylori* Lewis antigens might allow the bacterium to effectively evade the host immune response (Wang *et al.*, 2000) or induce autoantibodies directed against gastric parietal cell canaliculi (Negrini *et al.*, 1991). However, it has been suggested that Lewis antigens in *H. pylori* may mediate adhesion to gastric epithelial cells and facilitate colonisation, rather than mediating immune evasion or induction of autoantibodies (Appelmelk *et al.*, 2000; Moran *et al.*, 2001).

1.9.4. *H. pylori* neutrophil-activating protein (HP-NAP)

H. pylori usually does not enter the gastric epithelial cells or penetrate through the gastric epithelium. Therefore, dense infiltration of the gastric lamina propria by neutrophils and monocytes is probably recruited by soluble bacterial products. HP-NAP is a dodecamer composed of identical 17-kDa subunits (Evans *et al.*, 1995, Tonello *et al.*, 1999). It is a chemoattractant for neutrophils and monocytes and triggers the formation and release of toxic oxygen radicals (Satin *et al.*, 2000). HP-NAP is also antigenic since the majority of *H. pylori* infected individuals produce antibodies

specific for it. Therefore, it plays an important role in *H. pylori*-induced inflammation (Dundon *et al.*, 2001).

1.9.5. Blood-group antigen-binding adhesin A (BabA)

BabA is a 78-kDa protein produced by *H. pylori* *babA* gene. Two alleles, *babA1* and *babA2*, have been identified. The sequences of the two alleles are almost identical except for a 10 bp insert in *babA2*. The insert creates a frame shift and enables expression of *babA2*. BabA mediates adherence of *H. pylori* to human Lewis^b blood-group antigens on gastric epithelial cells (Boren *et al.*, 1993). *In vitro* analysis revealed that the binding is lineage-dependent (Falk *et al.*, 1993). In transgenic mice expressing the human Lewis^b epitope in gastric epithelial cells, Lewis^b mediates the binding of *H. pylori* attachment to epithelial cells and consequently colonisation of *H. pylori* induces gastritis and gastric atrophy (Falk *et al.*, 1995; Guruge *et al.*, 1998).

1.9.6. Other virulence associated factors

Urease is a hexameric molecule composed of UreA (30-kDa) and UreB (62-kDa) subunits (Austin *et al.*, 1991; Ha *et al.*, 2001). The enzyme plays a crucial role in acid survival of the bacterium by buffering the intracellular pH and is essential for colonisation of *H. pylori* in animal models (Eaton *et al.*, 1991; Karita *et al.*, 1995; Tsuda *et al.*, 1994). Two flagellin proteins, FlaA and FlaB, are essential for full motility of *H. pylori* and essential for colonisation (Eaton *et al.*, 1996; Suerbaum *et al.*, 1993). The protease of *H. pylori* may be involved in degradation of the gastric mucus and may be an important factor that contributes to duodenal ulcer disease (Moran, 1996).

1.10. Genetic diversity

Variability of the organisation and gene order in *H. pylori* has been revealed by pulse-field gel electrophoresis (PFGE) (Jiang, *et al.*, 1996; Saunders *et al.*, 1997) and restricted fragment length polymorphism (PCR-RFLP) analyses (Stone *et al.*, 1997). The variable PFGE or RFLP patterns between isolates are most likely resultant from the high rate of silent mutations and gain or loss of genes (Alm *et al.*, 1999).

Comparison of the two sequenced genomes also revealed that 6%-7% of the genes were only found in one of the two genomes. These genes are referred to as strain-specific genes (Alm *et al.*, 1999) and are usually not essential for survival of the organism. Bacteria gain such genes possibly for better interaction with, or adaptation to the host (Jain *et al.*, 2002; Ochman *et al.*, 2000). Loss of genes results in genomic reduction and probably occurs to remove potentially deleterious genes. From the bacterial perspective, gain or loss of genes is a major mechanism to achieve better adaptation. It is observed as variable genomic sizes and varied fingerprinting patterns among strains. From the host's perspective, however, this process often leads to modification of bacterial virulence and varied pathogenicity among bacteria. In *H. pylori*, three categories of strain-specific genes are frequently found: (1) *cag* pathogenicity island that is a major virulence determinant as described above (Censini *et al.*, 1996); (2) Type II restriction-modification systems which consist of a restriction endonuclease and a methyltransferase cleaving at the 4-5 bp sequences or preventing cleavage by methylation (Aras *et al.*, 2001; Lin *et al.*, 2001; Nobusato *et al.*, 2000; Takata *et al.*, 2002); (3) Genes encoding *H. pylori*-specific or conserved hypothetical proteins (Alm *et al.*, 1999).

Divergent sequences caused by mutation occur frequently in *H. pylori*. Comparison of two sequenced genomes revealed occurrence of substantial silent mutation in the genetic loci (Alm *et al.*, 1999). Examples of non-synonymous point mutation, however, have also been identified in *H. pylori*. Mutations in the 23S RNA gene (A-G transition at nucleotide position 2143 or 2144) confer resistance of the bacteria to clarithromycin (Debets *et al.*, 1998; Versalovic *et al.*, 1997).

Homologous recombination within a locus could lead to insertions and deletions. Varied numbers of a 108 bp repeat sequence are present in the 3' end of the *cagA* alleles (Yamaoka *et al.*, 1998). Strand slippage mispairing is another mechanism responsible for genetic diversity. A number of homopolymeric tracts and dinucleotide repeat regions have been identified from analyses of the two genomes (Alm *et al.*, 1999; Saunders, *et al.*, 1998). Slipped strand mispairing of these repeats within coding sequences may result in the loss or gain of a unit, such that the ORF may be in or out of frame with an upstream translation initiation codon, or may result in mutations preventing formation of a functional protein. A short homopolymeric sequence repeat has been identified in *fliP* (Josenhans *et al.*, 2000). Strains harbouring *fliP* with a C₈ repeat have full motility and normal flagella. Strains harbouring *fliP* with a C₉ repeat, however, have no flagella and have totally lost motility. In addition, sequence variations resulting in the varied phenotypes have also been identified. VacA alleles exhibit a mosaic structure in their signal region (*s* region) and mid-region (*m* region) (Atherton *et al.*, 1995). Three types, namely *sl**a*, *sl**b* and *s*2 were found in the *s* region and three types *m*1, *m*1*a* and *m*2 (Strobel *et al.*, 1998) occur in the *m* region. Strains harbouring a *sl*/*m*1 type *vacA* produce higher amounts of VacA than *vacA* *s*2/*m*2

strains. Extensive diversity has also been found in virulence-associated loci such as *flaA*, *flab* and *ureB* (Forbes *et al.*, 1995; Han *et al.*, 2000).

Genetic diversity may also occur at the promoter sequences. Sequence variability in the *vacA* promoter regions leads to altered promoter strength and therefore the varied transcriptional level among strains (Forsyth *et al.*, 1998). Analysis of the genome also revealed that 27 genes had homopolymeric tracts and dinucleotide repeat regions in their promoter regions (Saunders, *et al.*, 1998). Slipped strand mispairing is likely to alter the nucleotide length of tracts and hence switch on or off the expression of the genes.

1.11. Techniques for analysis of genomic diversity

Comparison of the complete genome sequences of different strains is no doubt the most powerful approach to investigate the genetic diversity. However, the substantial cost and labour involved with this approach limit its application. Two approaches are frequently used in the analysis of the genomic differences between organisms.

Microarray: This technique is used to compare the genomic compositions between a reference organism (driver) and an organism to be probed (tester). The analysis involves construction of DNA arrays, probing the arrays and data collection and analysis. Compared to genome sequencing, the technique is powerful and cost-effective. The application of microarray analysis however, is limited since the construction of gene arrays is dependent on knowledge of genomic sequences which makes identification of novel genes unlikely (Boucher *et al.*, 2001). In addition, the requirement for complex equipment also limits its use.

Subtractive hybridisation: An alternative approach for genomic composition comparison is subtractive hybridisation. This technique can also be used to identify the differentially expressed genes between two conditions (Jiang *et al.*, 2000). The principles of the technique are to remove common sequences between two closely related nucleic acid populations, thereby allowing the enrichment of differential sequences between the two populations. Several subtractive hybridisation approaches have been developed (Akopyants *et al.*, 1998; Jiang *et al.*, 2000). These approaches differ in the strategies for hybridisation to remove the common sequences to the tester and the driver nucleic acid and in the strategies for enriching the differential sequences. A PCR-based subtractive hybridisation has been used to compare the genomic contents between strains of *H. pylori* (Akopyants *et al.*, 1998). A simplified subtractive hybridisation approach, rapid subtractive hybridisation, was developed to identify the differentially expressed genes in mammalian cells (Jiang *et al.*, 2000, refer to Chapter 4.1). Cost-effectiveness and the capacity to identify novel sequences are advantages of subtractive hybridisation.

1.12. Genotypes and clinical outcomes

The outcomes of *H. pylori* infection are diverse. The correlation between genotypes and clinical diseases has been studied, with many of the works focusing on the *cagA*, *vacA* and *iceA* genotypes.

1.12.1. Status of *cagA*

Based on the status of *cagA* gene in the isolate, *H. pylori* is grouped into *cagA*⁺ or *cagA*⁻ strains (Covacci *et al.*, 1993). In most Western European and American countries

cagA⁺ strains are more prevalent in peptic ulcer diseases and distal gastric cancer (Blaser 1995; Covacci *et al.*, 1993; Ryan *et al.*, 2000). Furthermore, *cagA*⁺ strains are significantly associated with the presence of epithelial damage, higher degree of lymphocyte and neutrophilic cell infiltrates, atrophy and the presence of intestinal metaplasia (Nogueira *et al.*, 2001). The *cagA*⁺ strains can activate mitogen-activated protein kinase cascades and induce expression of proto-oncogenes *c-fos* and *c-jun* (Hirata *et al.*, 2001; Meyer *et al.*, 2000). Patients infected with *cagA*⁺ strains may have increased risk for the development of severe diseases (Blaser and Berg, 2001). However, conflicting results were also observed. In East Asia, nearly all *H. pylori* strains possess *cagA*, irrespective of the clinical outcomes (Yamaoka *et al.*, 1998; Kersulyte *et al.*, 2000). In addition, *cagA*⁺ strains are not associated with the clinical outcomes in certain America strains.

1.12.2. The *vacA* genotypes

The nucleotide sequence of *vacA* is significantly divergent among *H. pylori* isolates. The *vacA* alleles can be divided into conserved and highly divergent regions (Atherton *et al.*, 1995). Three signal sequence types have been identified, namely *sla*, *s1b* and *s2* (van Doorn *et al.*, 1998b). The mid-region is basically typed into *m1* and *m2*. Nevertheless, additional mid-region variants have been identified in different geographical regions, for example, *m1a* was found in German isolates (Strobel *et al.*, 1998) and *m1b* and *m1b-m2* in Chinese isolates (Pan *et al.*, 1998).

Strains with *sla/m1* mosaic genotypes produce higher levels of VacA *in vitro* and induce vacuolation of epithelial cells (Atherton *et al.*, 1995). Severe histological damage such as epithelial damage, atrophy and intestinal metaplasia has been observed

in patients infected with such strains. High prevalence of *s1* type strains was found in individuals with peptic ulcer diseases and gastric cancer (Miehlke *et al.*, 2000; Rudi *et al.*, 1998). However, most strains from East Asia have *s1a* types (van Doorn *et al.*, 1998b). The *m1* type is predominant in Japanese isolates, while *m2* is predominant in Chinese isolates (Kersulyte *et al.*, 2000).

1.12.3. The *iceA* genotypes

Adherence of *H. pylori* to gastric epithelial cells induces expression of a gene, designated as *iceA* (for induced by contact with epithelium) (Peek *et al.*, 1995). The *iceA* alleles can be divided into *iceA1* and *iceA2*. The association between the *iceA* genotypes and clinical outcome is uncertain. Strains of *iceA1* type were initially found to be more prevalent in peptic ulcer diseases and therefore were associated with the development of peptic ulcer diseases (van-Doorn *et al.*, 1998a). Subsequent studies however did not confirm this association (Wong *et al.*, 2001; Yamaoka *et al.*, 1999). The nucleotide sequences of 19 strains of *iceA1* type and 19 of *iceA2* have been determined. Sequence analysis indicates that there is a strong nucleotide sequence homology between *iceA* and *nlaIII*R encoding a restriction endonuclease. Frameshift mutations, insertions and deletions are frequent in *iceA* of either type (Figueiredo *et al.*, 2000). It is, therefore, unlikely that *iceA* would have significant impact on virulence considering its substantial diversity at amino acid level.

1.12.4. The combined genotypes

Simultaneous analysis of multiple genotypes has been suggested to be a better predictor of the clinical outcomes of infection than genotyping a single gene. Gerhard *et al* (1999) reported that strains of *cag*+/*vacA**s1/iceA1* were associated with severe gastric

diseases. Conflicting results, however, have been reported (Yamaoka *et al.*, 1999). In addition, genotyping of *H. pylori* strains may be complicated as an individual may be infected with multiple strains. An isolate from an individual, for example, may simultaneously possess *cagA*⁺/*cagA*⁻ subpopulations, *vacA* *s1/s2*, *vacA* *m1/m2* or *iceA1/iceA2* subpopulations (Arents *et al.*, 2001; van Doorn *et al.*, 1998). The results, therefore, should be interpreted with caution. Molecular typing of other virulence genes or marker genes is required to predict the outcomes of infection.

1.13. Transcriptional regulation

Bacteria usually survive and grow in rapidly changing and unfavourable environments. To cope with the harsh environments, bacteria evolve complex mechanisms to control transcription (Ishihama, 2000). Transcription is completed by a DNA-dependent RNA polymerase (RNAP) in three sequential steps: initiation, elongation and termination. Bacteria, such as *E. coli*, contain the simplest forms of cellular RNAP consisting of two large subunits (β and β') that associate with two copies of a smaller subunit, α , to form the core RNAP (Wosten, 1998). A sigma (σ) subunit binds to the core enzyme in a reversible manner to form the holoenzyme which is capable of initiating transcription at a specific start site on the DNA template (Burgess and Antony, 2001). The binding of RNAP holoenzyme to the promotor results in a so called “closed complex” which is converted into an “opened complex” by melting of a short DNA region within the sequence bound by the enzyme (Wosten, 1998). In the resulting ternary complex, the holoenzyme starts to produce 2 to 12 bp nascent mRNA molecules. Following disassociation of σ subunit, the RNAP core enzyme elongates through kilobases of DNA and finally stops at specific sequences on the terminator (Burgess and Antony, 2001). Transcription in each of these processes is elaborately modulated and hence

ensures the expression of genes at a required level (Wosten, 1998). The nascent mRNA, however, may undergo degradation that also contributes to regulation of gene expression in a cell (Akeda *et al.*, 2000).

H. pylori uniquely inhabits the acidic gastric environment. The bacterium is exposed to acidic gastric contents when passing through the gastric lumen and periodically encounters acid attack even after it successfully colonizes the gastric mucosa (McGowan *et al.*, 1996). In addition, *H. pylori* may have to cope with other stresses including the oxidative stress resulting from inflammation, temporal exposure to varying temperature, antibiotics and mutagens in the diet. The gene transcription in *H. pylori* therefore, must be regulated accordingly. Currently, the emerging data have begun to uncover the mechanisms of transcriptional regulation in *H. pylori*.

1.13.1. Transcription apparatus

RNA polymerase: The three polypeptides in the RNAP core enzyme are usually encoded by three genes (*rpoA*, *rpoB* and *rpoC*). In most Gram-negative bacteria, *rpoB* and *rpoC* are separated by a short, untranslated linker of 50-100bp. These two genes, however, are fused in *H. pylori* (Zakharova, *et al.*, 1998). The resultant β - β' fused protein is approximately 300 KD when separated on the SDS-PAGE gel. A fused *rpoBC* may be a common feature of the *Helicobacter* genus as it is also present in 10 species of *Helicobacter* (Zakharova, *et al.*, 1999). The fused β - β' protein does not provide any advantages over the growth or survival of *H. pylori*. The natural fused protein is possibly an evolutionary accident (Raudonikiene, *et al.*, 1999).

Sigma factors: Based on sequence similarity, the bacterial sigma factors are grouped into two families, σ^{70} -like factors and σ^{54} factors (Wosten, 1998). The σ^{70} family consists of three groups, the primary σ^{70} factor (group 1), non-essential σ^{70} like factors (group 2) and alternative factors (group 3). *H. pylori* contains only three sigma factors, σ^{80} , σ^{28} and σ^{54} , in contrast to the seven sigma factors in *E. coli* (Alm *et al.*, 1997; Blattner *et al.*, 1997). The genome of *H. pylori* lacks group 2 sigma factors such as RpoS which is a major factor during entry into the stationary phase and a global regulator in acid stress (Tomb *et al.*, 1997).

The group 1 bacterial sigma 70 factors or primary sigma factors are responsible for the transcription of most genes expressed in exponential growth (Wosten, 1998). Four conserved amino acid sequence regions are identified in *E. coli* σ^{70} (Burgess, 2001; Wosten, 1998). These four regions are divided into many subregions according to their respective functions. The subregions 2.4 and 4.2 recognize the -10 and -35 promotor elements, respectively. In *E. coli*, this promotor sequence recognized by σ^{70} is characterized by two stretches of six nucleotides, TATAAT and TTGACA, centered around positions -10 and -35 respectively, from the transcription start point (+1). A dinucleotide TG at position -15 and -14 is also conserved in *E. coli* promoters and is known as the “extended -10” promotor element (TGnTATAAT). The strength of the promotor is dependent on the conservation of these promoter elements as well as the size of the spacer located between the -35 and -10 elements.

A group 1 or primary sigma factor, namely σ^{80} factor as its molecular mass is approximately 78.7 kDa, is present in *H. pylori* (Beier *et al.*, 1998). The σ^{80} factor is comprised of approximately 680 amino acid residues. Region 1.1, a non-conserved

region is missing and replaced with homology to a regulatory protein from *Bacillus subtilis* phage SPO1 (Vuthoori *et al.*, 2001). The conserved region from 2.1 to 4.2 has greatest identity to a group of gram-positive bacteria including *Staphylococcus aureus*, *Listeria monocytogenes*, and *Enterococcus faecalis* (Solnick *et al.*, 1997). The function of σ^{80} might be less constrained than its counterpart in *E. coli* because of its greater rate of amino acid replacement. This might enable σ^{80} to recognize relatively less conserved promoters resulting from significant plasticity of the *H. pylori* genome. The consensus promoter sequences recognized by σ^{80} have been suggested as TTAAGC-x(19-23)-TATAAT (Vanet, 2000). The -10 promoter sequences from 21 genes with experimental evidence are aligned here (Figure 1.1). The results suggest that the most frequent sequence is TAAAAT at the -10 region. However, if one nucleotide substitution in the -10 element is allowed, the most frequent sequence (16/21) is TATAAT as suggested although this sequence itself is only present in two genes. The extended -10 element is found in 8 genes, suggestive of the frequent use of this element in *H. pylori*. The majority of genes under control of σ^{70} in *E. coli* have both -10 and -35 promoter elements. In the aligned *H. pylori* genes, however, -35 element is only present in 6/21. This suggests *H. pylori* does not frequently use the -35 promoter element for transcription initiation.

H. pylori also possesses two additional sigma factors: σ^{28} and σ^{54} (Tomb, *et al.*, 1997). Both sigma factors are involved in transcriptional regulation of classes of flagellar genes (Colland *et al.*, 2001; Josenhans *et al.*, 2002; Spohn and Scarlato, 1999b). The σ^{28} factor belongs to group 3 sigma factors or alternative sigma factors and recognises a 4 nucleotide (nt) sequence TAAA at -35 region and an 8 nt sequence GCCGATAA at the -10 region relevant to the transcription start point (TSP). The

AAACTCATTTTATAATTAGAATTTAAGCTAATTTTGTAGTAAATAACAAAGG	<i>rpoD</i>
ATAAGATAAAATAATAAAACGCATCATTAACCATTGATTGAATACCACTAGATAAATACCCCAT	HP1408
CTAAACCCTTATAAGTTATGGATAAAAACTTAAAGAATAGGGTAAATAAACGCA	HP1243
CGATTACATTTTTACCCACAATTAACAAAAATTGCTTAAATGAAAGAAC	<i>cagB</i> P2
ATTGATAGATTATTAACCTTTTATGCTATAATGCGAGGG	<i>dnaK</i>
TTTTTTCAATATTAATAATGATTAATGAAAAAAAAAAAAATGCTTGATATTGTGTGTATAATGAGAATG	<i>cagA</i>
ATCTTGTTGAGTTTTATGTTTATTACCTTAATTGATAAAAGTTTAATA	<i>vacA</i>
AGAAAAGACCAATAATCTAAAGATGTTAGGGTAAATTAACATTTTGAC	<i>hpaA</i>
TTAAGAAAAGATTTACCAAAAAGTATTAAAAAATGATTACAATACGGCTA	<i>sodB</i>
TAATAAACGCTTCTGTAAATATTAGTAAATCAAACATGCTACAATTACATCCAA	<i>ureA</i>
TTATTTTTAAGATATTTTCAAGGTATCTTTAATAATAAAATGCTAGAATTAGTCCC	HP166
AATTGAATCTTGTTAGTCTTTGTATAACAAATTATGTAATAATCACCACA	<i>katA</i>
ATTATTGTCGGTCAAGGTAATTTTAATCAAACATGCTATTATTGGAAAT	<i>cheY</i>
TTTGTTTTTATGGCTTGACTTATCCCTAAAAATGCGCTATAGTTGTGTCTG	<i>hspA</i>
TGACCCTTACAATTTTCAAACCAATCGTTAATGCGATTAAATATGGCTATATATACTACAATA	HP119
AGCGCCGTATAGTTTTTA	<i>topA</i>
TTGACTTATCCCTAAAAATGCTATAGTTATGTCTG	<i>groES</i>
TTGATTGACTAACACTAAAGATTATGATAGTATTC	<i>cbpA</i>
AATGGTTAGTCAAGGTGAGCACATGAAGCGTTCTAAAATAGTTTGGTTT	<i>cagB</i> P3
AGCGTGTTTATGGGGTTGTATTTGATTTATGGCATCATTCGGTGGCTCTT	<i>copA</i>
CCTTAGCGTTTTTAATAACCTTATAAGTCCGCCAAGACTTCTTAAGGGTT	<i>repA</i>

Figure 1.1. Alignment of *H. pylori* -10, extended -10 and -35 promoter sequences in 21 genes. The most frequent -10 promoter sequence is TAAAT which is present in 4 genes. If one nucleotide substitution in the promoter sequence is allowed, however, the most frequent sequence is TATAAT which is also recognized by *E. coli* σ^{70} (16/21). If two nucleotide substitutions in the promoter sequence are allowed, TATAAT is present in 19 of 21 genes. The extended -10 promoter element and -35 element is found in 8 genes and 6 genes, respectively. The -10, extended -10 and -35 elements are in red, pink and blue respectively. The TATAAT and those with one nucleotide substitution in the -10 element are highlighted.

length of the spacer between the two regions is 15 nt. (Burgess, 2001; Wosten, 1998). The holoenzyme comprised of RNA polymerase and σ^{54} requires an enhancer protein to be fully functional. The consensus promotor sequence recognized by σ^{28} is TGGCAC at the -24 region TTGCW at the -12 region with a 5nt spacer (Fujinaga *et al.*, 2001).

1.13.2. Two-component system

The two-component system, usually consisting of a sensor histidine kinase and a cognate response regulator, is a common mechanism for signal transduction in bacteria (Hoch, 2000). The sensor kinase has an N-terminal input domain perceiving environmental stimuli and a C-terminal core domain, referring to the transmitter domain (Fabret, 1999). Signal binding causes activation of the sensor kinase resulting in ATP hydrolysis and auto-phosphorylation of a histidine residue on the transmitter domain (Stock *et al.*, 2000). The activated sensor kinase subsequently transfers a phosphoryl group to a receiver domain of a cognate response regulator (Stock *et al.*, 2000). The phosphoryl group resides as an acyl phosphate of an aspartic acid residue on the receiver domain (Hoch, 2000; Stock *et al.*, 2000). The receiver domains normally inhibit the output domain on the response regulator and phosphorylation relieves this inhibition freeing the output domain to carry out its function (Falke *et al.*, 1997; Hoch, 2000). The response regulators usually function in transcription activation, but may also interact directly with other proteins (Falke *et al.*, 1997).

Genomic analysis reveals that *H. pylori* contains 4 histidine kinases and 6 response regulators (Tomb *et al.*, 1997). The smaller number of these genes, compared to that in *B. subtilis* and *E. coli*, is consistent with the paucity of global regulatory networks in *H. pylori* and reflects the tight adaptation of *H. pylori* to its restricted niche (Doig *et al.*,

1999). Each of the kinases couples with a cognate response regulator, while the remaining two orphan response regulators have no identified corresponding histidine kinases (Table 1.4). The functions of HP165-HP166, HP244-HP703 (*flgR*) and HP1364-HP1365 pairs are likely to be involved in transcription regulation (Beier *et al.*, 2000; Dietz *et al.*, 2002). The CheA-CheY two component system, however, is necessary for chemotaxis and colonization which are possibly carried out by direct interaction with flagellar motor proteins (Foynes *et al.*, 2000).

In strain 26695, HP165 was annotated as two open reading frames (Tomb *et al.*, 1997). This may be a sequencing error as it is a single open reading frame in strain J99 (Alm *et al.*, 1999). The predicted product of HP165, a 414 amino acid kinase is possibly attached to the cytoplasmic membrane. *In vitro* phosphorylation analysis demonstrated that it exclusively activates HP166 regulator (Beier *et al.*, 2000). HP166 is essential for survival as inactivation of this gene results in a lethal phenotype (Dietz *et al.*, 2002). The activated HP166 regulator at least positively regulates transcription of two subsets of genes which encode *H. pylori*-specific hypothetical proteins (Table 1.4). It binds to a 26 bp motif located upstream of the target genes. Mutation analysis, however, demonstrated that these target genes were not essential for survival. It is possible that the unphosphorylated HP166 regulator may be functional and regulate a set of essential genes. Transcription of HP166 is autoregulated, which is carried out by the binding of the phosphorylated HP166 regulator to the promotor sequence (Dietz *et al.*, 2002). HP703 regulator, a NtrC homologue, is exclusively activated by HP244 and regulates transcription of a set of flagellar genes (Dietz *et al.*, 2002).

Table 1.4. Functions of the two-component systems in *Helicobacter pylori*

Histidine kinases	Response regulators	Target genes	Adaptive roles
HP0165	HP0166 (OmpR)	HP1408-1412 regulons; HP118, HP119	Unknown, HP 166 is essential for survival <i>in vitro</i>
HP0244	HP0703 (NtrC), <i>flgR</i>	<i>flaA</i> , <i>flaB</i> , <i>flgE</i>	Flagellar biosynthesis
HP0392 (<i>cheA</i>)	HP1067 (CheY), <i>cheY</i>	Unknown	Motility
HP1364	HP1365 (OmpR)	Unknown	Unknown
N/A	HP1043 (OmpR)	Unknown	Essential
N/A	HP1021 (NH)	Unknown	Essential

N/A, no homologues of the corresponding kinases. NH, no homologue in the public database.

1.13.3. Transcriptional regulation of the *cag* genes

Transcription of *cagA* is under the control of the σ^{80} promotor as shown in Figure 1.2. An identifiable -35 and the extended -10 promoter sequences are not observed (Spohn *et al.*, 1997). Expression of *cagA* varies in response to environmental stimuli. Transcription levels of *cagA* increase gradually as the growth time extends and reaches the maximum level at 48 h when bacteria are grown in Brain Heart Infusion (BHI) broth (Karita *et al.*, 1996). Increased expression of *cagA* has also been observed when cells grown for 48 h in BHI broth were briefly exposed to acidic McIlvain's buffer. Duration of exposure (5-50 minutes) to acid does not appear to have a significant impact on *cagA* transcription. Free iron in the media contributes to transcription regulation of *vacA* but does not affect transcription of *cagA* (Szczebara *et al.*, 1999). The varied expression of *cagA* indicates that a regulatory factor is likely to be involved in the transcription regulation of *cagA* or that a posttranscriptional regulation mechanism is involved. Quorum sensing and two component systems are major global regulatory networks which are responsible for virulence gene expression in other bacteria (Latifi *et al.*, 1995; Sperandio *et al.*, 1999). In *H. pylori*, the quorum sensing system, however, does not affect transcription of *cagA* (Joyce, *et al.*, 2000).

Transcription of nine *cag* genes (*cagI*, *cagI0-11*, *cagI3-16*, *cag21* and *cag25*) has been examined using gene fusion with *ureB* (Joyce *et al.*, 2001). Each of these genes possesses an upstream region with promotor activity. Co-culture with an epithelial cell monolayer was shown to result in 14-fold increase in expression of *cagI5* and *cag21* (Joyce *et al.*, 2001). However, both genes are not involved in the translocation of CagA (Joyce *et al.*, 2001).

1.13.4. Transcriptional regulation of *vacA*

Transcription of *vacA* is initiated at nucleotide position 119 upstream of the AUG start codon (Forsyth *et al.*, 1998). Transcription of *vacA* is started by a single promoter in the majority of strains. In rare cases, transcription of *vacA* may start at other positions (Forsyth *et al.*, 1998). The sequences of promoter -10 and -35 elements have been identified as TAAAAA and TTTATG respectively (Forsyth and Cover, 1999). Like *cagA*, an extended -10 element (TG) is necessary for transcription of *vacA*. Expression of *vacA* is growth-dependent. Peak levels of *vacA* transcription are reached at the onset of stationary phase and subsequently decline over time (Forsyth *et al.*, 1998). Differential expression of *vacA* was seen between *vacA* s1/m1 and *vacA* s2/m2 strains (Forsyth and Cover, 1999). The higher *vacA* transcriptional level in *vacA* s1/m1 strains resulted in a higher amount of VacA in the supernatant. Heterogeneity in *vacA* transcription level among strains, therefore, is an important determinant for the different vacuolation phenotypes induced by VacA. Regulatory elements involved in growth phase-dependent transcriptional regulation of *vacA* are not fully understood (Forsyth and Cover, 1999). Mutation of the upstream region of -35 element has been shown not to affect the expression level of *vacA* (Forsyth *et al.*, 1998). The bacterial quorum sensing system is involved in diverse physiological processes. In *Agrobacteria* species, virulence gene expression is under control of a quorum sensing system (Luo and Farrand, 2001). Transcription of *vacA* in *H. pylori*, however, is not regulated by *luxS* (Forsyth *et al.*, 2000; Joyce *et al.*, 2000).

basal body and hook proteins, as well as the regulators FlgM and σ^{28} for transcription of the class 3 flagellar genes (Colland *et al.*, 2001; Josenhans *et al.*, 2002). FlgR is a response regulator of a two-component system with HP0244 as the histidine sensory kinase (Spohn and Scarlato, 1999b). HP0244 specifically transfers a phosphate to a conserved aspartate residue in FlgR *in vitro* and subsequently activates FlgR (Spohn and Scarlato, 1999a). The class 3 genes include the major filament subunit FlaA, hook proteins and four hypothetical proteins which are possibly involved in flagellar assembly. Transcription of the genes in this class is under control of a σ^{28} factor (FliA) and an anti- σ^{28} factor (FlgM, HP1122) (Colland *et al.*, 2001). No *flaA* transcript can be detected in the *fliA* defective mutants of *H. pylori*, indicating that *flaA* expression is positively regulated by *fliA*. The *fliA* defective mutants have truncated flagella, morphologically similar to those in *flaA* deficient mutants (Josenhans *et al.*, 2002). In *flgM* defective mutants, the level of *flaA* transcript is approximately 2.5 to 4 fold higher than that in the wild-type strain, whereas the morphology of the flagella in *flgM* defective mutants is similar to the wild-type strain (Colland *et al.*, 2001). FlgM interacts with the sigma factor FliA and subsequently prevents the interaction between FliA and the β -subunit of RNA polymerase antagonising the transcription mediated by σ^{28} -holoenzyme (Colland *et al.*, 2001). The interacting domain is a 29-residue region located at the C-terminal of FlgM (Colland *et al.*, 2001). In anti- σ^{28} factor containing bacterial species, transcription activation of late genes is usually achieved by secreting the anti- σ^{28} factor from the cell on sensing the completion of basal-body-hook assembly and enabling RNA polymerase to associate with σ^{28} transcribing late genes. In *H. pylori*, it is unclear if transcription of class 3 genes is dependent on secretion of FlgM. *H. pylori* FlgM cannot be exported efficiently in *S. typhimurium* although it can complement the *S. typhimurium* mutants deficient in FlgM (Josenhans *et al.*, 2002).

H. pylori cells can adapt the mechanical properties of the flagellar filament by varying the relative amounts of FlaA and FlaB in response to environmental signals (Suerbaum, 1995). In addition to the regulators described above, DNA topology affected by gyrase and topoisomerase I is likely to be associated with regulation of *flaA* and *flaB* expression (Spohn and Scarlato, 1999a). Evidence for the involvement of *gyrA* and *topA* comes from the observations including (1) *topA* and *flaB* are oppositely oriented and their promoters are overlapped (Suerbaum *et al.*, 1998a), (2) *gyrA* is co-transcribed with *flgR* which regulates *flaB* transcription (Spohn and Scarlato, 1999a); (3) expression of *flaB* increases when the gyrase activity is inhibited (Spohn and Scarlato, 1999a); (4) the short spacer region makes *flaA* expression very sensitive to DNA topology (Suerbaum *et al.*, 1998a);

1.14. Adaptation and the consequences of *H. pylori* infection

The basic requirements for a bacterium to establish colonisation involve overcoming the numerous physical, cellular and molecular barriers presented by the host (Ochman and Moran, 2001; Weiss, 2002). Typically, this entails contacting and invading the host body, exploiting the nutrient sources from the host, avoiding host defences, persisting in the particular niches, and finally infecting a new host. *H. pylori* uniquely colonises the human stomach (Blaser and Berg, 2001). To adapt to this niche, the organism faces a variety of challenges from the host. Acid stress is the most common challenge encountered by *H. pylori* when invading and persisting in the human stomach (Moran *et al.*, 2001). In addition, the organism has to adapt to diverse individual hosts that differ significantly in their genetic background, to cope with the dynamic environments resulting from physiological (e.g. age-related) and pathological changes (for example,

from superficial inflammation to atrophy) in a particular host, and to deal with exposure to a number of potential harmful factors (e.g. antibiotics) during ingestion.

1.14.1. Acid adaptation

The habitat of *H. pylori* is the human stomach or in rare cases, the primate stomach (Handt *et al.*, 1995). The organism usually colonises beneath the mucous layer where pH approaches neutrality. However, *H. pylori* experiences acid stress when passing the gastric lumen or even when persisting in the mucous layer (McGowan, *et al.*, 1996). The bacterium is able to survive at external pH 2 for 30 min *in vitro* (Bauerfeind *et al.*, 1997). Pre-adaptation to weak acid (such as pH 6) increases the survival rate of the bacterium (Karita *et al.*, 1998). The phenomenon is known as adaptive acid tolerance response (ATR). The bacterium is able to grow at pH 5.5 *in vitro*, although grows optimally between pH 6.9 and 8.0 (Owen, 1998).

Survival in the acidic conditions requires the production of urease. The activity of urea to produce ammonia by hydrolysis of urea is necessary for the survival of *H. pylori* in acidic environment *in vitro* (Mobley *et al.*, 1988). Surface-bound urease could buffer the acidic environment *in vitro* (Krishnamurthy *et al.*, 1998). Later studies, however, revealed that internal urease is also required for survival of the organism (Stingl *et al.*, 2002a; Stingl *et al.*, 2002b). UreI, a component of urease regulons is a pH-gated urea channel (Weeks *et al.*, 2000). The channel is fully open when environmental pH falls below 5.5 and permits influx of urea. Subsequently, urease is likely to maintain internal pH close to neutrality by catalysis of urea into ammonium (Stingl *et al.*, 2002b). In addition, acid resistance pathways independent of urease cluster are also possibly involved in survival and growth of *H. pylori* (Bijlsma *et al.*, 1998; Bijlsma *et al.*, 2000).

Enterobacterial species, such as *E.coli* and *Salmonella typhimurium*, encounter acid stress when passing through the human stomach. The major mechanism accounting for the acid resistance in these bacteria is to alter the expression of groups of genes (Foster, 1999). For *H. pylori*, a similar mechanism may be used. Identification of these acid responsive genes is required to understand such acid adaptation (refer to Chapter 3).

1.14.2. Genetic divergence and adaptation

Bacterial adaptation to hosts is a cycle of processes: invasion, persistence in the host and preparation for invading another host of the same or even a different species. A number of molecular mechanisms are involved in bacterial adaptation (Giraud *et al.*, 2001; Ochman, 2001). Through horizontal gene transfer, bacteria acquire novel properties to transmit to and persist in a new niche of the same host or a new host (Ochman, 2001). Transformation, transduction and conjugation are responsible for the horizontal transfer of genes (Ochman *et al.*, 2000). Conversely, loss of genes (a phenomenon termed ‘evolution by reduction’) occurs if the genes are not required for life within the host (Dobrindt and Hacker, 2001). Adaptive mutation confers new traits for bacteria to survive and grow in a new host or unfavourable conditions (Giraud *et al.*, 2001). Antibiotic resistance resulting from mutation exemplifies the importance of mutation in bacterial adaptation (Versalovic *et al.*, 1997). New traits generated from intra-genome recombination also contribute to adaptation (Suerbaum *et al.*, 1998b; Webb and Blaser, 2002). These molecular mechanisms ultimately lead to diversification of the bacterial genomes, displayed as chromosomal rearrangements, strain specific genes and allelic polymorphisms. For *H. pylori*, two unrelated strains are unlikely to have identical fingerprinting (Jiang *et al.*, 1996), indicating there is a

high degree of inter-strain genetic divergence. Comparison of the genomes of two strains also reveals similar results (Alm *et al.*, 1999; Tomb *et al.*, 1997). These inter-strain diversities in clinical isolates, however, have not been fully characterised. Analysis of genetic divergence among strains is important for understanding the adaptation mechanisms used by *H. pylori* (refer to Chapters 4 and 5).

1.14.3. Genetic diversity and pathogenesis

Through diversifying the genome, bacteria acquire the capacity to adapt to individual hosts with different genetic backgrounds. A major clinical concern is whether such adaptation leads to severe clinical outcomes. One assumption is that evolution selection favours the best adapted bacteria regardless of whether the ultimate consequence is harmful, benign or beneficial (Ochman and Moran, 2001). For *E. coli*, a normal resident of the mammalian intestinal flora, acquisition of a single pathogenicity island is sufficient to confer pathogenetic properties upon benign strains (McDaniel and Kaper, 1997). Another assumption is that the virulence of “well-adapted” bacteria is usually attenuated (Weiss, 2002). These assumptions are based on the analysis of inter-species genetic diversities. For a particular species, as in the case of *H. pylori* which is well adapted (estimated to colonise half the human population), the consequences of bacterial adaptation through genomic diversification are unclear. Most strains of *H. pylori* possess the *cag* pathogenicity island which is an important determinant for the consequences of infection. Investigation of the divergences of the *cag* pathogenicity island would aid our understanding of the complicated impacts of adaptive evolution.

1.15. Aims:

The aims of the studies described here were to investigate the transcription response of *H. pylori* to acid exposure and inter-strain genetic divergences, and thereby, to understand bacterial adaptation mechanisms and the impacts of adaptation on clinical outcomes.

Specifically, the objectives of the studies described here were as follows.

- (1) To identify *H. pylori* genes regulated by prolonged acid exposure using differential display-PCR.
- (2) To investigate the inter-strain differences in the genomic contents within *H. pylori* species and the relevance of the genomic diversities to the consequences of infection.
- (3) To analyse the composition diversity and sequence variations in the *cag* pathogenicity island.

CHAPTER 2
GENERAL MATERIALS AND METHODS

2.1 General Materials

2.1.1. Molecular reagents

All reagents used were analar grade and purchased from Sigma-Aldrich except where otherwise indicated. Water used in all experiments was purified by reverse osmosis, deionisation and UV treatment to attain an 18 M Ω quality (Maxima Four, Elga). The following solutions were prepared by standard methods as outlined by Sambrook *et al.* (1989).

0.5 M Ethylenediamine-tetraacetic acid (EDTA), pH 8.0: purchased in solution form.

Phosphate buffered saline (PBS): purchased in tablet form.

1M Tris (hydroxymethyl methylamine)-hydrochloride (Tris-HCl): purchased in solution form.

Tris acetate EDTA (TAE) buffer (1X): 40 mM Tris-acetate, 0.1 mM EDTA, pH 8.0.

Tris borate EDTA (TBE) buffer (0.5X): 45 mM Tris-HCl; 45 mM boric acid; 1mM EDTA, pH 8.0.

Tris EDTA (TE) buffer: 10 mM Tris-HCl, pH 7.4; 1mM EDTA, pH 8.

Agarose gel loading buffer: 0.25 (w/v) bromophenol blue and 50% glycerol in 0.5X TBE.

Ammonium acetate (7.5M): 57.8 g ammonium acetate was dissolved (BDH) in 80ml water, then brought to 100 ml volume with water and filter sterilised (Millex 0.45 μ m syringe filter unit, Millipore)

Ethidium bromide staining solution (0.5mg/l): Diluted ethidium bromide stock solution (10 mg ml⁻¹) was prepared at 1:20000 ratio with water to give a final concentration of 0.5 mg l⁻¹. Stored in the dark at room temperature (RT).

Proteinase K (100 mg): Dissolved 100 mg proteinase K in 5 ml water (20 mg ml⁻¹) and aliquoted in 500 µl volumes. Stored at -20 °C.

RNase (Dnase-free) (Boehringer Mannheim, 500 µg/ml): The ready-for-use solution was diluted 1:100 in sample DNA to give a final concentration of 5 µg ml⁻¹, in accordance with the manufacturer's instructions.

Sodium acetate (3M): Dissolved 40.8 g sodium acetate trihydrate in 80 ml water, adjusted to pH 5.2 with glacial acetic acid, brought to 100 ml with water and sterilised by autoclaving at 121 °C for 15 min.

2.1.2. *H. pylori* strains

The *H. pylori* strains used in this study included clinical isolates and a type strain NCTC 11638 that was obtained from the National Collection of Type Cultures, Colindale, London.

Primary isolation of *H. pylori* strains: Clinical *H. pylori* isolates were isolated from the gastric biopsies in this study or from the culture stocks. The gastric biopsy specimens were obtained from patients undergoing upper gastrointestinal endoscopy at the Meath and Adelaide Hospital, Dublin, Ireland. Biopsy specimens were inoculated with mucosal side downwards onto the surface of two fresh blood agar plates consisting of Columbia agar base (LAB M, UK) supplemented with 7% horse blood. Plates were

incubated at 37 °C for 4~7 days under microaerophilic conditions by flushing the anaerobic jars with CO₂ or using a CO₂ Generation Kit (Oxoid, UK). *H. pylori* was identified by Gram staining, typical colony morphology and urease activity (Christensen's agar slopes containing 40% urea, Oxoid, UK).

Storage: *H. pylori* strains were sub-cultured in duplicate onto blood agar and incubated under microaerophilic conditions for 3 days. The bacterial cells were collected and placed into a sterile 2 ml freezer vial containing 1 ml of storage medium. The medium comprised of Brain Heart Infusion Broth (BHI, Oxoid), 20% (v/v) glycerol and 10% (v/v) filtered horse serum. The vials were then vortexed briefly and stored at -80 °C.

Recovery of *H. pylori* strains: Vials containing the strains were defrosted at RT. The suspension was then removed from the vials with a sterile loop and inoculated onto two blood agar plates. Each plate was inoculated at 4 areas with 10 µl suspension each. The plates were then incubated for 4 to 7 days.

2.2. DNA Manipulation

2.2.1. Extraction of genomic DNA

Genomic DNA was extracted using the method described by Ausubel *et al* (1987). Using a sterile loop, bacteria were harvested from 3 to 5 days old cultures and washed twice in 1ml of sterile distilled water or PBS. The bacteria were pelleted using a bench-top microcentrifuge at 3000 g for 5 min. The pellet was resuspended in 567 µl of TE buffer and 30 µl of 10% sodium dodecyl sulphate (SDS, w/v) and 3 µl of proteinase K (20 mg ml⁻¹) was added. The suspension was incubated in a 37 °C water bath. After 1 h 100 µl of 5 M NaCl was added and mixed thoroughly. Then 80 µl of cetyltrimethy

ammonium bromide (CTAB)-NaCl solution [10 % (w/v) CTAB in 700 mM NaCl]] was added and the suspension was incubated for 10 min at 65 °C. An equal volume of chloroform: isoamyl alcohol (24:1) was then added to the solution and it was mixed thoroughly before centrifugation for 5 min. The upper aqueous layer was removed to a clean microcentrifuge tube and an equal volume of phenol: isoamyl alcohol (24:1) was added. The solution was mixed thoroughly and centrifuged as before. Again the aqueous layer was removed to a clean microcentrifuge tube and 0.6 volume of isopropanol was added to precipitate the genomic DNA. The DNA was pelleted by centrifugation for 30 min at 4 °C. The pellet was washed with 1 ml of 70 % (v/v) ethanol and then air-dried. The genomic DNA was resuspended in 50-100 μ l of TE buffer (pH 7.6).

The ultraviolet spectrophotometric method and/or agarose gel electrophoresis method as outlined by Sambrook *et al* (1989) were used to determine the concentration, quality, and relative purity of the DNA samples. Briefly, for spectrophotometrical quantification of DNA samples, appropriate volume of a DNA sample (5 μ l in most cases) was diluted with distilled water. The sample was then measured at 260 nm and 280 nm wavelength. The A_{260} at 1 was equivalent to 50 μ g ml⁻¹ DNA (40 μ g ml⁻¹ RNA or PCR generated DNA). For determining the purity, the A_{260}/A_{280} ratio greater than 1.8 for DNA (1.9 for RNA) suggests a pure DNA sample. For determining the DNA concentration with agarose gel analysis, a DNA sample was run on 1% agarose gel in parallel with a series of DNA standards of known amounts (*e.g.* 0, 2.5, 5, 10, 20, 30, 40, 50 μ g ml⁻¹). The concentration of the DNA samples was then estimated by comparing the intensity of the sample DNA band with the DNA standards.

2.2.2. Purification and concentration of DNA

The DNA solution was treated with equal volume of phenol:chloroform: isoamyl alcohol (25:24:1), mixed by inversion and then centrifuged at 12,000 g for 10 min. This step was repeated until the upper aqueous layer was clear. The upper aqueous layer was removed to a fresh microcentrifuge tube. An equal volume of chloroform:isoamyl alcohol (24:1, Sigma) was then added, mixed by inversion and then centrifuged at 12,000 g for 10 min. The upper aqueous layer was removed to a clean microcentrifuge tube and mixed with 1/10 volume of 3 M sodium acetate (pH 5.2) and cold absolute ethanol (the volume equal to twice of the total volume of upper aqueous layer plus sodium acetate). The mixture was then placed in -20 °C for 1 h to overnight. The sample DNA was pelleted at 4 °C for 30 min at 12000 g and washed twice in 70% ethanol. The DNA sample was air-dried and then resuspended in appropriate buffer.

2.2.3. Agarose gel electrophoresis and visualisation.

DNA was routinely separated on 0.8 to 2% agarose gels in 1 X Tris acetate-EDTA (TAE) buffer or 0.5 X TBE buffer. Gel loading buffer was added to all DNA samples at a ratio of 1:5. Electrophoresis of restriction endonuclease digested chromosomal DNA was performed using a Hoefer® HE 100 SuperSubTM horizontal tank (Hoefer Pharmacia Biotech) and a chiller unit (Mini Chiller Model 1000, Bio-Rad). All other electrophoresis was performed using Bio-Rad horizontal electrophoresis tanks. Gels were prepared using type II medium EEO agarose (Sigma) and freshly diluted TBE (0.5X). Electrophoresis was preformed under the following conditions:

DNA samples	Agarose concentration % (w/v)	Voltage (V)
PCR products <800 bp	2	75
PCR products >800 bp	1.5	75
Chromosomal DNA	0.8	75
Digested chromosomal DNA	0.8	30

Gels were stained in ethidium bromide (0.5 mg l^{-1}) for 10 min immediately after electrophoresis. Excess ethidium bromide was removed by washing in water for 10 min. Otherwise, gels were pre-stained by adding ethidium bromide (10 mg ml^{-1}) to the cooled molten agarose to a final concentration of 0.5 mg/ml . Gels were viewed on a transilluminator (2011 Macrovue, LKB) under UV light at 312 nm and photographed with an MP4 land camera using Polaroid 667 film and a 23A Wratten filter. When negatives were required, Polaroid 665 film (Sigma) was used. Photographs were scanned with an Epson PerfectionTM 1240U Colour Image Scanner to a Gateway 450 personal computer. Images were processed with Adobe Photo Shop post-production software (Version 5.0).

2.2.4. Polyacrylamide gel electrophoresis

For differential display PCR of RNA, 6%(w/v) denaturing gels were routinely used. The gels were made by mixing 30% (w/v) acrylamide solution [29% (w/v) acrylamide, 1% (w/v) *N, N'*-methylenebisacrylamide], urea (7M) in 1 X TBE, 0.1% (w/v) ammonium persulphate, 0.1% (v/v) *NN, N', N'*-tetra-methylene-bisacrylamide (TEMED). DNA samples were loaded onto the gels, run constantly at 42 W for 3 h.

After completion of electrophoresis, the gels were stained using a silver staining kit (Promega) following the manufacturer's instructions.

2.2.5 Labelling of DNA probes with Digoxigenin (DIG)

Double-stranded DNA was labelled with Digoxigenin-11-dUTP (DIG-11-dUTP) using DIG-High Prime reagents (Roche). The DNA template [$16\ \mu\text{l}$ (30-100 ng μl^{-1})] was denatured at $100\ ^\circ\text{C}$ for 10 min, cooled on ice immediately and $4\ \mu\text{l}$ of DIG-High prime reagent was added. The mixture was incubated at $37\ ^\circ\text{C}$ for 2 to 12 h and the reaction stopped by addition of $2\ \mu\text{l}$ EDTA solution (200 mM, pH 8.0). Then 0.1 volume of 4 M LiCl and 2.5 volume of ethanol were added to the reaction and it was placed at $-70\ ^\circ\text{C}$ for 30 min. The DNA was pelleted in a bench-top microcentrifuge at 12,000 g for 15 min at $4\ ^\circ\text{C}$ and then washed with $100\ \mu\text{l}$ of 70% (v/v) ethanol. The pellet was air-dried and resuspended in $50\ \mu\text{l}$ of sterile distilled water.

2.2.6. Southern hybridisation

Southern hybridisation was carried out as outlined by Sambrook *et al.* (Sambrook *et al.*, 1989). Genomic DNA was digested with an appropriate restriction endonuclease. The digested DNA was loaded onto a 0.8% agarose gel in 0.5 X TBE buffer. Electrophoresis was carried out at $4\text{-}6\ \text{V cm}^{-1}$ for 1 to 2 h. For denaturation of the sample DNA, the gels were submerged into the appropriate volume of denaturation solution (0.5 M NaOH, 1.5 M NaCl) with constant shaking for 30 min and this process was repeated. The DNA was then neutralised (1 M Tris-HCl, pH 8.0, 1.5 M NaCl) twice for 30 min with shaking. The DNA was transferred to a positively charged nylon membrane (BrightStar-plus, Amibon) using southern capillary transfer in 20 X SSC buffer (3 M NaOH, 0.3 M sodium citrate) overnight. The membrane was air-dried and

baked at 80 °C for 2 h to fix the DNA to the membrane. Prehybridisation was carried out by placing the membrane in the hybridisation solution [5 X SSC (750 mM NaCl, 75 mM Sodium citrate), 0.1% (w/v) N-lauroylsarcosine, 0.02% (w/v) SDS, 1% blocking reagent (Sigma)] for 2 to 4 hours at 68 °C. Hybridisation was then performed for 20 h with an appropriate volume of fresh hybridisation solution. The DIG-labelled probe was added into the hybridisation buffer immediately after denaturation at 100 °C for 10 min and immediately cooled on ice for 5 min. Following hybridisation, the membrane was washed twice in 2 X SSC buffer with 0.1% SDS at room temperature for 15 min (low-stringency wash). The membrane was then washed twice under high-stringency conditions (0.1 X SSC buffer with 0.1% SDS at 68 °C) for 15 min each. The hybridised signals were detected using DIG-High Prime DNA labelling and Detection Starter Kit (Roche) following the manufacturer's instructions.

2.2.7. Polymerase Chain Reaction (PCR)

The PCR amplification was carried out in a final volume of 50 μ l containing 1 μ l DNA template (20~100 ng μ l⁻¹), 1.5 mM MgCl, 0.4~0.8 μ M each primer, 1.5 U *Taq* polymerase (Qiagen or Sigma), 200 μ M each of dATP dCTP dGTP dTTP, and 1 X PCR buffer. The reaction solution was overlaid with mineral oil when appropriate. The thermal cycling parameters were 1 cycle of 94 °C followed by 30 to 35 cycles of denaturing at 94 °C (0.5 ~ 1 min), annealing at appropriate temperature (0.5 ~ 1 min) and extension at 72 °C (0.5 ~ 1 min). A final extension at 94 °C (5 ~ 30 min) was included in the amplification. The primers were designed using the primer design program available on the Internet (<http://www.ebi.ac.uk>). The primers were located at the most conserved DNA regions revealed by alignment (Clustal W, <http://www.ebi.ac.uk>) of the known sequences deposited in the GenBank

(<http://www.ncbi.nih.org>). Synthesis of the primers was provided by the service from MWG Biotech Ltd, London, UK.

2.2.8. Cloning and sequencing

For cloning the PCR products, the amplicons obtained from amplification with a final extension for 30 min were purified using a PCR Purification Kit (Qiagen) following the manufacturer's instructions. The purified PCR products were subsequently cloned into appropriate vectors using the corresponding cloning kits (Invitrogen) following the manufacturer's instructions.

Determination of the nucleotide sequences was provided by the service from MWG Biotech Ltd. Sequence analyses were then carried out using the appropriate programs available on the Internet (<http://www.ebi.ac.uk> or <http://www.ncbi.nih.org>). These programs included Blast N/P/X, Fasta 3, ORFinder, Clustl W.

2.3. RNA manipulation

2.3.1. Northern Hybridisation

Total RNA was extracted using Tri-reagent following the manufacturer's instructions. The concentration and the relative purity was determined using an ultraviolet spectrophotometric method as outlined by Sambrook *et al.* (1989) (see also section 2.2.1). The integrity of the RNA samples were determined using agarose gel electrophoresis. Samples of total RNA were loaded onto a 0.8% denaturing agarose gel. Electrophoresis was carried out at 4-6 V cm⁻¹ for 1 h. RNA was transferred to a positively charged nylon membrane (BrightStar-plus) using a capillary transfer method with a downward transfer apparatus. The transfer continued for approximately 1.5 h (15

min per millimetre of gel thickness). The membrane was rinsed briefly in 1 X TBE buffer, air-dried and baked at 80 °C for 2 hours to fix the RNA. Prehybridisation was carried out by placing the membrane in the hybridisation solution (Amibon) and incubating at 42 °C for 2 to 4 h. The heat-denatured Dig-labelled DNA probe was added into the hybridisation solution and incubated overnight. Following hybridisation, the membrane was washed twice in 2 X SSC buffer for 15 min and washed twice in 0.1 X SSC buffer at 42 °C for 15 min with agitation. The hybridised signals were detected using a DIG-High Prime DNA labelling and Detection Starter Kit (Roche) following the manufacturer's instructions.

CHAPTER 3

IDENTIFICATION OF *HELICOBACTER PYLORI* GENES REGULATED BY PROLONGED ACID EXPOSURE

3.1. Introduction

Bacteria usually survive and grow in rapidly changing and unfavourable environments. To cope with these harsh environments, bacteria evolve complex mechanisms to regulate transcription (Ishihama, 2000). Transcription in bacteria is completed by a DNA-dependent RNA polymerase (RNAP) in three sequential steps: initiation, elongation and termination (Wosten *et al.*, 1998). Transcriptional regulation occurs in each step to alter the expression of multiple genes. Varied gene expression in response to environmental signals has been observed in *H. pylori*, for example, elevated amounts of *fur* and *vacA* transcripts in response to low free iron (Szczepara *et al.*, 1999), altered expression of *hspR* regulons in response to heat shock (Huesca *et al.*, 1998; Spohn and Scarlato, 1999b) and varied *wbcJ* expression in response to acid stress (McGowan *et al.*, 1998).

Increased environmental acidity is the signal most frequently encountered by *H. pylori*. The organism experiences extremely low pH (pH 1-2) when it passes through the gastric lumen prior to reaching the protective mucosal surface (McGowan *et al.*, 1998). *In vitro* studies demonstrated that short exposure of *H. pylori* cells to acid altered the transcriptional level of an LPS-associated gene *wbcJ*, *cagA* and *hsp70* (Huesca *et al.*, 1998; McGowan *et al.*, 1998; Karita *et al.*, 1996). In addition, *H. pylori* encounters prolonged acid exposure *in vivo* (Lee *et al.*, 1993). Gastric acidity fluctuates under physiological conditions such as increased acid output at ingestion. Damage in the protective mucus layer may also decrease the local pH. Chronic *H. pylori* infection leads to increased gastric acid secretion (McColl. *et al.*, 1997). Such prolonged acid exposure may serve as an important environmental signal to regulate gene expression. Bacteria grown on acid (pH 6) media after 5 successive passages gave rise to colony

variation and enhanced ability to adhere and invade epithelial cells (Bukholm *et al.*, 1997). Moran and his colleagues (2002) observed acid-induced phase variation of Le^X and Le^Y in *H. pylori* cells grown in liquid culture. These observations suggest that prolonged acid exposure possibly regulates *H. pylori* gene expression (Moran, *et al.*, 2002).

Differential display PCR (DD-PCR) was initially used to identify differentially expressed genes in mammal cells between two different conditions (Liang and Pardee, 1992). The technique used random primers capable of frequently annealing to the reverse-transcribed cDNA. Consequently, PCR amplification could yield a range of products which could be resolved on a sequencing gel. The amplification conditions differed from regular PCR by using random (10 mer) primers, lower concentration of dNTP and shorter elongation time (30 sec). Total RNA or mRNA was reverse transcribed into cDNA using poly (dT) primers such as T₁₁CA. In bacteria, however, most RNA transcripts lack the poly (A) tails. In order to overcome this problem and apply DD-PCR to identify differentially expressed genes in bacteria, 11-mer random primers for reverse transcription were designed based on the computational analysis of *E. coli* genomic sequences (Fislagé *et al.*, 1997). Amplifications with the combination of 11-mer and 10-mer primers yielded reasonable binding patterns on the sequencing gels.

The aim of this section of the study was to investigate the influence of prolonged acid exposure on the transcriptional regulation of *H. pylori*. DD-PCR with different combinations of 11-mer and 10-mer primers was used to detect mRNA expressing profiles in *H. pylori* cells successively passaged on acidic (pH 5.5) and neutral media

(pH 7.4). The differentially displayed genes were then identified based on the comparison of the profiles between the two conditions.

3.2. Materials and Methods

3.2.1 Strain and culture conditions

H. pylori strain 98587 was isolated from the antrum of the stomach of a patient with duodenal ulcer disease. The strain was *cagA* positive/*vacA* s1/m1 detected by PCR using previously described conditions (Atherton *et al.*, 1995; van Doorn *et al.*, 1998a) (see appendix for detail). To prepare the acidic media, 1.35 ml of concentrated HCl was added into 1 litre Mueller-Hinton agar (Oxoid) supplemented with 7% horse blood immediately before pouring the plates. The pH of the agar plates was determined using a surface pH meter and the value ranged from 5.5 to 5.7. The neutral media (pH 7.4) were prepared as for preparation of the acidic media except for addition of HCl. *H. pylori* strain 98587 was recovered from the stock on neutral plates and subsequently grown on acidic (pH 5.5) and pH 7.4 Mueller-Hinton agar plates in parallel under microaerophilic conditions at 37°C for 3-5 days. Bacterial cells grown for 5 successive passages were collected for extraction of total RNA.

3.2.2 Differential display PCR

Total RNA from *H. pylori* cells that had been collected from acidic and neutral media was extracted using Tri-reagents (Sigma) according to the manufacturer's instructions. Residual DNA contaminations were removed by RNase-free DNase I (one unit per microgram total RNA) treatment. The concentration of RNA was determined by spectrophotometric analysis and the purity was demonstrated by agarose gel electrophoresis. For the synthesis of the first strand of cDNA, samples of total RNA were reverse transcribed in six independent reactions using six different non-anchored 11-mer primers with an AT-rich 3'-end (Ea1, Ea2, Ea3, Ea4, Ea6, Ea8) as previously described by Fislage, *et al.* (1997) (Table 3.1). Each reverse transcription reaction

Table 3.1. Sequences of the 11-mer and 10-mer primers for DD-PCR

RT primers	Sequence (5'→ 3')	G+C (%)	PCR Primers	Sequence (5'→ 3')	G+C (%)
Ea1	TTTTATCCAGC	36	Es1	GCTGGAAAAA	40
Ea2	ACTTTACGCAG	45	Es3	GAAGTGCTGG	60
Ea3	TTTATCCAGCG	45	Es5	TGGCGGCGGC	90
Ea4	TCAGCGTTTTA	36	Es7	ATGCGCTGGC	70
Ea6	TTTTTTCAGCA	27	Es8	TGCCGATGAA	50
Ea8	ATCATCCAGCA	45	Es13	ATGGCGATGA	50

contained 1 X RT buffer, 20 μ M dNTPs, 0.2 μ M of one primer, 100u M-MuLV reverse transcriptase (Roche) and 2 μ g of total RNA as template. A negative control (water) was always included for each reverse transcription. The reaction was incubated at 37 °C for 60 min followed by incubation at 75°C for 10 min to inactivate the enzyme. PCR was carried out in a final volume of 25 μ l containing 2 μ l cDNA template, 3 mM MgCl₂, 20 μ M of the same 11-mer primer used for the reverse transcription, 20 μ M of one of six 10-mer primers (Es1, Es3, Es5, Es7, Es8, Es13) (Table 3.1), 8 μ mol dNTP, 1.5U *Taq* polymerase, Stoffel fragment (Perkin-Elmer). The sequences of the Ea and Es primers are shown in Table 3.1. The thermal cycling parameters were as follows: 5 low stringency cycles of 94°C for 1 min, 40°C for 2 min and 72°C for 1 min, followed by 35 high stringency cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 30 sec, and a final extension cycle at 72°C for 5 min. A 3.5 μ l aliquot of PCR product was mixed with 1.5 μ l loading buffer, heated at 95°C for 5 min for denaturation and run on a 6% sequencing gel for 3 h. Gels were stained using a silver staining method, air-dried and exposed to APC film (Promega).

3.2.3 Re-amplification and cloning of putative candidates

The DD-PCR products for cells grown on acidic and neutral conditions were visually compared. Putative candidate differentially expressed genes were identified based on the difference in the intensities of the bands. Amplifications were repeated for those lanes in which putative candidate differentially expressed genes were identified. The candidates were only identified from the lanes in which the banding patterns were reproducible. Bands of interest were cut from the gels and eluted in 100 μ l distilled water overnight at 4°C. Five microliters of eluted DNA samples were re-amplified in a final volume of 50 μ l with an appropriate pair of primers using the high stringency

PCR conditions as described above. The re-amplified cDNA fragments were cloned into pCR2.1 vector (Invitrogen) according to the manufacturer's instructions. Recombinant plasmids were isolated using a Wizard miniprep plasmid kit (Promega) for preparing the probes and sequencing.

3.2.4. DNA sequencing and analysis

The nucleotide sequences of the candidate differentially displayed fragments were determined by the sequencing service at MWG-Biotech. Sequences were blasted against the *H. pylori* genome (<http://www.tigr.org>) and/or GenBank database. Multiple alignments were carried out using the Clustal W program. An open reading frame (ORF) was predicted using ORFinder.

3.2.5. Northern Blot

Northern Blot analyses were carried out using Northern plus Kits (Ambion) following the manufacturer's instructions. Briefly, aliquots of 10-15 µg total RNA each was separated by denaturing agarose gel electrophoresis and blotted onto BrightStar-plus nylon membrane. Membranes were probed with biotin-labelled probes prepared from cDNA clones or from PCR products obtained in the differential display. Hybridisations were performed at appropriate temperatures and the hybridisation signals were detected using a CSPD chemiluminescent method.

3.3. Results

3.3.1 Variation in morphology of colonies

Strain 98587 was grown on acidic media (pH 5.5) and pH 7.4 media in parallel for five passages. Compared to the colony appearance of culture from neutral media, colonies from acid culture were smaller, had an opaque appearance and showed a stronger zone of haemolysis. The difference in colony morphology suggested the gene expression profile of cells grown on acidic media was possibly different from that of cells grown on neutral media.

3.3.2. Identification of differentially expressed genes

To detect the expression profiles of *H. pylori* cells grown on acidic and neutral media, six Ea 11-mer primers (for reverse transcription) and six Es 10-mer primers (for PCR) were chosen, resulting in 36 possible combinations of Ea and Es primers. Each combination with Ea and Es primers displayed approximately an average of 50 bands. Hence, 1800 different mRNA species were screened. The expression profiles of cells grown on acidic and neutral media displayed on the sequencing gels were analysed. Eleven bands appeared to be differentially expressed (Figure 3.1). These bands were cut from the gels, recovered and re-amplified. The re-amplified products were subsequently cloned into pCR 2.1 vectors. Northern blot analysis demonstrated that the expression of seven candidates was elevated in cells grown on acidic media, while one was down regulated (Figure 3.2). The expressions of the remaining three candidates were not detectably affected by acidic exposure.

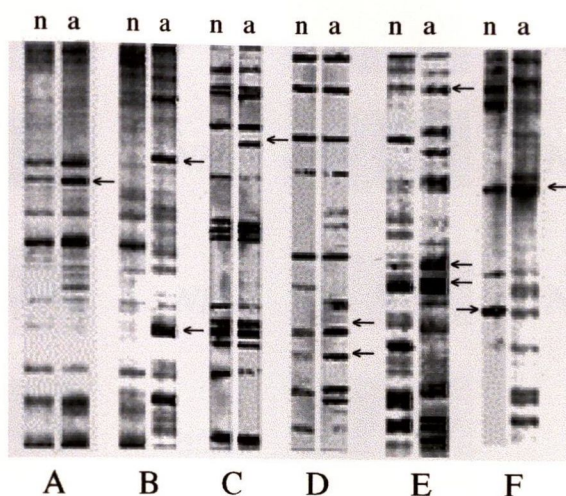


Figure 3.1. DD-PCR for *H. pylori* cells grown on pH 7.4 (n) and pH 5.5 (a) media for five passages using different pairs of primers: A, Ea3/Es13; B, Ea1/Es8; C, Ea8/Es13; D, Es8/Es7; E, Ea2/Es3; F, Ea6/Es8. Arrows indicate bands from which the candidate differentially expressed genes were identified. The candidates from each panel are as follows: A, *ureB*; B, *topA* (upper) and *tufB*; C, NH1; D, ND1 (upper) and *flaA*; E, ND2 (upper), *pap57* (middle) and ND3 (lower); F, NH2 (upper) and *atoE*.

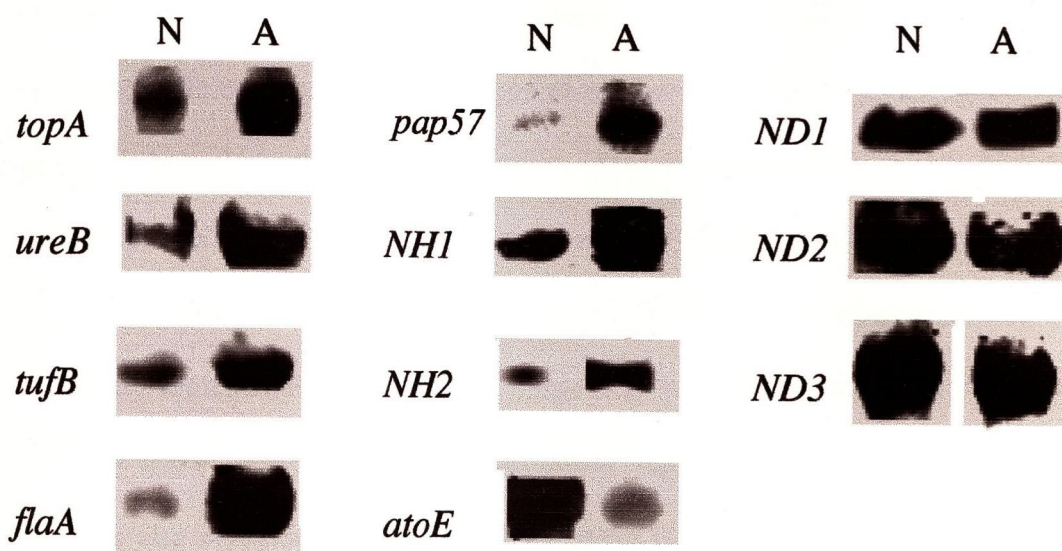


Figure 3.2. Northern blot analyses of total RNA from *H. pylori* cells grown on pH 7.4 (N) and pH 5.5 media (A) to confirm differential expression of the identified cDNA fragments. Total RNA was separated on 1% denaturing agarose gel, blotted onto nylon membrane and subjected to hybridisation with specific biotin-labelled probes. The expression of *topA*, *ureB*, *tufB*, *pap 57*, *NH1* and *NH2* was increased at pH 5.5. The expression of *atoE* was decreased. The expression of candidates ND1-3, however, was not detectably affected by acid exposure.

3.3.3. Nucleotide sequence analysis

Sequence analysis revealed that five differential cDNA fragments had homologues in the predicted coding sequences in *H. pylori* strain 26695 or J99 genomes with a minimum identity of 94%. These included *topA*, *tufB*, *ureB* and *fla* homologues whose expressions were increased in cells grown on acidic conditions. The only down-regulated differentially expressed fragment is an *atoE* (HP0693) homologue encoding a putative short-chain fatty acid transporter. One up-regulated fragment, designated as *pap57*, had no homologue in the two sequenced *H. pylori* strains. It, however, showed significant identity to a functionally unknown EAL domain existing in diverse bacterial signalling proteins (Figure 3.3). The remaining two cDNA fragments, NH1 and NH2, did not show any homology to previously reported gene sequences. Sequencing analysis also revealed that one cDNA fragment (*pap57*) was amplified by Es7 primer only, while others were amplified by a combination of an Ea primer with an Es primer.

atgcgctggcagcatccggaaaaaagcatgatttcaccggcagaatttatcccgtg
M R W Q H P E K S M I S P A E F I P L
gcagaagaaagcggactgatcatcccgattggcaactgggtcatgcgcacagcctgt
A E E S G L I I P I G N W V M R T A C
gcgcaactacgcgctggcaccaacaaaaagaaaccagccatctgaccatgtccatc
A Q L R A W H Q Q K E T S H L T M S I
aacgccagcgcat
N A S A

EAL consensus	48	IRWQHPDGGIISPDEFIPLAEETGLIVPLGRWVLE	83
Pap57	1	M.....E-KSM...A.....S...I...N..M	35
EAL consensus	84	QACQQLAEWQAAGPD---LRISINLSA	107
Pap57	36	RT..A..RA.HQQKETSH.TM...A..	61

Figure 3.3. Sequence of *pap57* and alignment with the consensus sequence of the EAL domain. Amino acid sequence of *pap57* was translated using Transeq DNA sequence translation tool (EMBOSS) available from the website of EBI (<http://www.ebi.co.uk>). The dots in the amino acid sequence of *pap57* represent the residues identical to those in the EAL domain. -, gap.

3.4. Discussion

H. pylori inhabits the surface of the human stomach. Infection by the bacterium results in increased gastric acid output especially in duodenal ulcer patients. Gastric acid secretion fluctuates under physiological conditions (Merrell and Camilli, 2002). Damage to the mucus layer is frequently observed in the stomach of patients infected by *H. pylori*. Therefore, *H. pylori* is likely to experience life-long acid attacks persistently or periodically, despite the protection offered by the mucus layer (McGowan *et al.*, 1996). To investigate the impact of prolonged acid exposure on the bacterium, strain 98587 isolated from the antrum of a duodenal ulcer patient was grown in pH 5.5 and pH 7.4 media in parallel. Bacterial cells grown on acidic media had remarkably different colony appearance compared to cells grown on neutral media, indicating that prolonged exposure to weak acid could change the life style of *H. pylori* cells, and even the expression of some bacterial pathogenic factors (Bukholm *et al.*, 1997; Moran *et al.*, 2002). Therefore, when analysing the effect of *H. pylori* on epithelial cells *in vitro*, one should take into account the impact of acid exposure on bacterial virulence.

To analyse the effect of prolonged acid exposure on the gene expression of *H. pylori*, DD-PCR was used to compare the expression profiles from cells grown on acidic media and neutral media. DD-PCR was originally used for eukaryotic cells where poly dT primers were used in the protocol (Liang and Pardee, 1992). However, the majority of *H. pylori* transcripts, as well as those in other bacteria, lack polyA tails at their 3' ends. This restricts the use of 3'-anchored poly dT primers for reverse transcription in *H. pylori*. In this study, Ea primers for reverse transcription were from those designed for the group of *Enterobacteriaceae* (Fislage *et al.*, 1997). Six Ea primers combined

with six Es primers (for PCR) were chosen in this study. Each combination of Ea and Es primers yielded a number of bands, suggesting that those primers are suitable for DD-PCR in *H. pylori*. A total of 1,800 bands were displayed. Since the predicted coding sequences of *H. pylori* are around 1,600, each mRNA transcript was likely to be screened if it is represented by one band. These results demonstrated that this DD-PCR protocol was efficient for identification of differentially expressed genes in *H. pylori*.

H. pylori produces large amounts of urease which liberates ammonia from urea that may contribute to neutralization of gastric acidity (Moran 1996) or prevention of cellular acidification (Stingl *et al.*, 2002a). Urease is crucial for survival of *H. pylori* at acidic pH *in vitro* (Stingl *et al.*, 2002c) and for successful colonisation of the gastric environment (Eaton *et al.*, 2002; ; Karita *et al.*, 1995; Tsuda *et al.*, 1994a). Mutants defective in *ureB* were not able to colonize animal models (Tsuda *et al.*, 1994b). Transcription of *ureB* is under control of a *H. pylori* σ^{80} factor. It is co-transcribed with *ureA* and uses the promoter sequences located within the upstream region of *ureA* (Shirai *et al.*, 1999). Increased levels of *ureB* mRNA in cells exposed to acid for a prolonged period observed in this study may result from the increased stability of the *ureB* transcript at acidic pH (Akada, *et al.*, 2000) and/or induction of *ureB* transcription (Slonczewski, *et al.*, 2000; Van Vliet, *et al.*, 2002). *H. pylori* produces high amounts of urease constituting up to 10% of the total protein content of cells grown in neutral media *in vitro*. However, an increased release of urease and an increased amount of urease in the outer membrane protein fraction have been found in *H. pylori* cells grown on acidic media (Fislage *et al.*, 1997). UreB also increased in cells grown in pH 5.5 liquid media (Slonczewski, *et al.*, 2000). These observations demonstrate that expression of urease is not constitutive. Therefore, it is likely that the expression of

urease in *H. pylori in vivo* is up-regulated to combat against frequent or persistent acid attacks.

In addition to urease, other genes are likely to be involved in acid resistance of *H. pylori*. A change in protein content was observed in response to a shift of extracellular pH from 7.5 to 3.0 (McGowan *et al.*, 1996). Urease positive mutants generated by random insertion mutagenesis were identical to the wild-type *H. pylori* strain on pH 7 plates, but did not grow on pH 5 plates. Transformation with wild-type DNA restored the growth (Bijlsma *et al.*, 1998). Therefore, a urease-independent acid resistance system, probably depending on the expression of more than one gene, is involved in the growth of *H. pylori* at low pH. In agreement with the hypothesis, increased expressions of *tufB*, *flaA* and *topA* in cells grown on acidic media were observed in this study. All three genes are likely to play important roles in the stress encountered by *H. pylori*.

Elongation factor Tu (EF-Tu) is involved in the binding and transport of the appropriate codon-specified aminoacyl-tRNA to the aminoacyl site of the ribosome (LaRiviere *et al.*, 2001). In addition to its role in polypeptide elongation, EF-Tu may regulate synthesis of proteins which respond to stress at the level of transcription (Young and Furano, 1981). It has also been revealed that *E. coli* EF-Tu interacts with unfolded and denatured proteins, as do molecular chaperones that are involved in protein folding and protein renaturation after stress (Caldas, *et al.*, 1998). Only one copy of the gene encoding EF-Tu is present in *H. pylori*, in contrast to two identical copies in *E. coli*. This gene is designated as *tufB* in strain 26695. Data presented here demonstrate that *H. pylori tufB* is acid-regulated. Differential expression of *tufB* in response to environmental signals has also been observed in *E. coli* and *Coxiella burnetii* (Seshadri

et al., 1999). The role of *tufB* in acid resistance of *H. pylori* is unclear. Cytoplasmic acidification (Stingl *et al.*, 2002 a) and influx of urea (Weeks, *et al.*, 2000) resulting from low external pH may influence the folding and denaturation of proteins. Therefore, one could speculate that *H. pylori* might up-regulate EF-Tu to promote functional folding of proteins and thus maintain the appropriate activity of enzymes as does its *E. coli* counterpart.

H. pylori cells possess a unipolar bundle of sheathed flagella. The flagellum is composed of two substructures: a hook-basal-body that spans the bacterial membranes and an external filament (Josenhens *et al.*, 2002). The *H. pylori* filament contains two different flagellin species of similar molecular mass that are present in different amounts: the major flagellin FlaA and the minor flagellin FlaB (Suerbaum *et al.*, 1995). FlaA is located at the distal end of the filament (Suerbaum *et al.*, 1995). Both flagellin proteins are required for full motility of *H. pylori* cells and are essential for the establishment of persistent infection of *H. pylori* in animal models (Eaton *et al.*, 1996; Leying *et al.*, 1992). A positive correlation between *flaA* mRNA level and colonisation efficiency has been observed (Mankoski *et al.*, 1999). Acid regulation of *flaA* supports the proposed hypothesis that *H. pylori* may alter the *flaA* expression in response to environmental signals (Suerbaum, 1995).

The *topA* gene encodes DNA topoisomerase I (TopA) in bacteria which is required for maintenance of negative supercoiling of DNA. DNA topology in the promotor region is an important factor that contributes to transcriptional regulation. TopA affects the transcriptional rate by altering the conformation of promotor DNA. Differential expression of *topA* in response to high temperature, osmolarity, and availability of

oxygen (Ni-Bhriain and Dorman, 1993; Qi *et al.*, 1999; Tse-Dinh, 2000; Yoshizawa and Yamamoto 1989) has been observed in *E. coli* and plays an important role in bacterial adaptation to these stresses. In *H. pylori*, both HP0116 and HP0440 have been predicted to encode TopA. The product of HP0440, which has not been characterised, is 30.5% identical to that of HP0116. The identified differentially expressed gene is a homologue of HP0116 which is essential for the survival of *H. pylori* (Suerbaum *et al.*, 1998a). Minor changes in cytoplasmic pH (pH 6.5-7.5) significantly alter DNA conformational structure (Neidhardt, 1996). Acid-induced expression of *topA* observed in this study might help prevent the DNA conformation alteration resulting from cytoplasmic acidification (Stingl *et al.*, 2002a). It has been demonstrated that *flaA* has a short spacer between -10 and -35 promoter elements, rendering their transcription sensitive to local DNA conformational changes (Suerbaum *et al.*, 1998a). The up-regulation of *flaA* and *topA* genes raises the possibility that induction of *flaA* is mediated by *topA* through altering the topology of the promoter DNA.

The *atoE* gene encodes a putative short-chain fatty acid transporter. It is located at the cytoplasmic membrane and transports C₄-C₆ fatty acid into the cytoplasm (Neidhardt, 1996). These fatty acid compounds are converted to acetyl-CoA in the cytoplasm. Uptake of short-chain fatty acid is sensitive to suppression by metabolites produced by the bacteria (Neidhardt, 1996). Down-regulation of *atoE* in response to prolonged acid exposure possibly results from the accumulation of metabolites when *H. pylori* cells are grown continuously in acidic conditions. Alternatively, it may reflect the bacterium's attempt to actively reduce the uptake of weak acid and hence maintain internal pH close to neutral.

The amino acid sequence of *pap57* showed a high degree of identity with an EAL domain existing in the diverse bacterial signalling proteins. The EAL domain is named after the conserved amino acid residues and its function is unknown. The EAL-containing proteins include sensor kinases, regulators and hypothetical proteins, suggesting that *pap57* may be involved in the sensing of environmental signals in *H. pylori*. The two sequenced *H. pylori* strains lack the EAL-containing protein, indicating that *pap57* is a strain-specific gene (Alm *et al.*, 1999; Tomb *et al.*, 1997). Generally, strain-specific genes are not essential for survival. The observed increased expression of *pap57* in response to prolonged acid exposure, therefore, is probably not directly associated with acid survival of the bacterium. Studies designed to determine the entire sequence of *pap57* and phenotypes of *pap57* mutants would be helpful for investigating the functions and roles of this gene in acid survival or signal transduction.

Two differentially expressed sequences identified in this study, *NH1* and *NH2*, showed no homology to previously reported gene sequences. Most probably these sequences represent the untranslated region of mRNA which are commonly less conserved than the coding regions between homologues in different organisms (Colonna-Romano, *et al.*, 1998). Only a few transcriptional regulators have been predicted in *H. pylori* (Tomb *et al.*, 1997). The bacterium lacks the global regulator for acid stress. A panel of genes, however, has been shown to be regulated by a shift to acidic pH (Allan *et al.*, 2001; Ang *et al.*, 2001). Data presented here also revealed differential expressions of at least eight genes in response to prolonged acid exposure. However, the transcriptional regulation pathway and the functions of most acid responsive genes have not been characterized and should be further investigated.

CHAPTER 4

DIVERGENCE IN THE GENOMIC CONTENTS AMONG

***HELICOBACTER PYLORI* STRAINS**

4.1. Introduction

Diverse clinical outcomes occur in individuals infected by *H. pylori*. Host, bacterial and environmental factors play a role in determining outcome (El-Omar *et al.*, 2000; Fox *et al.*, 2000; Graham *et al.*, 1997). *H. pylori* CagA, VacA and LPS are important virulence factors that contribute to the determination of clinical outcomes (Atherton *et al.*, 1995; Covacci *et al.*, 1993; Heneghan and Moran, 2001; Moran *et al.*, 1996; Moran, 1999; Moran and Prendergast, 2001).

H. pylori is a species with a plastic genome (refer to Chapter 1.10). Strains isolated from unrelated individuals rarely have identical fingerprinting as defined by PFGE, RFLP and RAPD-PCR (Akopyants, *et al.*, 1992; Gibson, *et al.*, 1998). Allelic polymorphism occurs frequently in *H. pylori*. Furthermore, comparison of the whole genomic sequences demonstrated the existence of several groups of strain-specific genes: *H. pylori*-specific proteins, conserved hypothetical proteins, DNA restriction-modification (R-M) systems and genes involved in cell-envelope synthesis and cellular processes (Alm, *et al.*, 1999). The strain-specific genes confer distinctive properties that may determine strain-specific traits, such as virulence (for example, *cag* PAI), adaptation and antibiotic resistance (Boucher and Doolittle, 2001; Jain, *et al.*, 2002).

In some individuals, *H. pylori*-associated gastritis may stay unchanged through their lifetime, while in others, may progress divergently into gastric cancer (GC) or duodenal ulcer (DU) (Blaser and Berg, 2001). Factors affecting the divergence of these pathologies are poorly understood. Variation in the genomic contents among strains of *H. pylori* may be a contributing factor. In this study, an attempt was made to identify differences in genomic content of *H. pylori* strains from patients with different

pathologies and to investigate the relevance of genomic divergence to clinical outcomes. A new subtraction hybridisation technique, namely rapid subtraction hybridisation (RaSH), was recently developed by Jiang *et al.* (2000) to identify differentially expressed genes by comparing RNA from human melanoma cells treated with interferon with RNA from untreated control. This technique utilizes cloning to screen the subtraction hybridization library and simplifies the procedures of subtraction hybridisation. The procedures of RaSH are illustrated in Figure 4.1. In this study, RaSH was applied to compare the genomic contents of two Irish strains of *H. pylori*: one isolated from a patient with GC and the other from a patient with DU. The distribution of the strain-specific genes identified from the comparison was subsequently determined in a panel of clinical isolates.

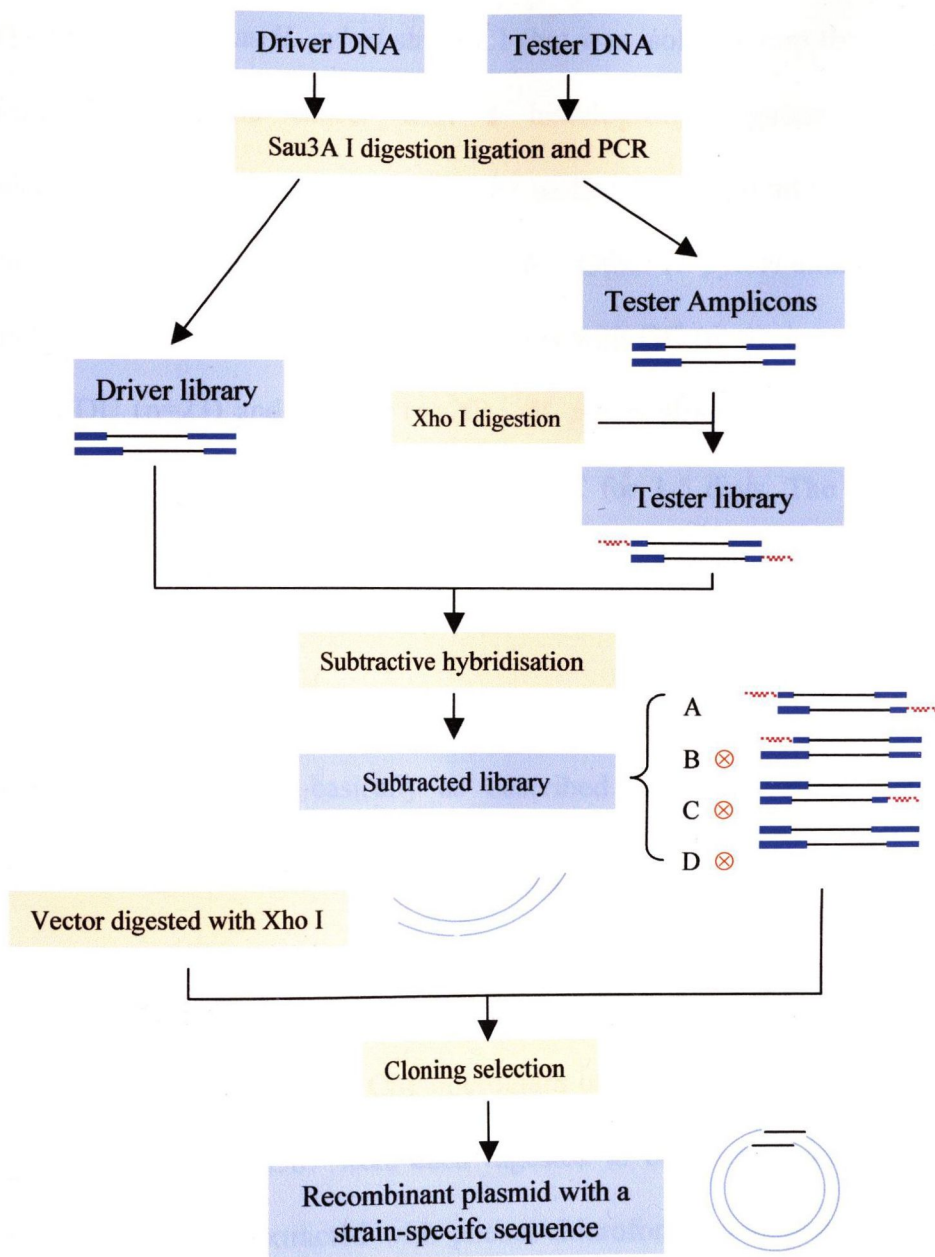


Figure 4.1. Schematic illustration of the procedures of RaSH according to Jiang *et al.*, 2000. The subtracted library contains 4 types of DNA fragments. Type A fragments, which are specific to the tester, have clonable ends. Type B and C fragments are hybridised DNA between the tester and the driver. Type D fragments are excessive driver DNA. Type B, C and D are not clonable since they lack the Xho I compatible ends.

4.2. Materials and Methods

4.2.1 Bacterial culture

The tester strain was *H. pylori* strain C1 that was isolated from the antrum of a patient with distal gastric cancer with a histological diagnosis of intestinal type adenocarcinoma. *H. pylori* strain 98587 isolated from the antrum of a patient with a duodenal ulcer was used as a driver strain. Other *H. pylori* strains used in this study included strains from the antrum of patients with GC (n=4), intestinal metaplasia (IM, n=8), DU (n=23) and gastritis (n=20). *H. pylori* strains were grown on blood agar plates under microaerobic conditions at 37°C for 3-5 days. The details of the strains used in this section of the study can be found in the appendix.

4.2.2 RaSH

RaSH was conducted basically as described by Jiang *et al.* (2000) with some modifications.

Preparation of the driver and tester DNA libraries: Genomic DNA was extracted as described in Chapter 2.2.1. One microgram of tester DNA from strain C1 and driver DNA from strain 98587 were each digested to completion with 10 units Sau3 AI (Roche) for 16 h, extracted with phenol/chloroform, precipitated with ethanol, and resuspended in 10 mM Tris·HCl, pH 7.5. The digested DNA was mixed with primers XDPN-14/XDPN-12 (Table 4.1) (final concentration, 12 µM) in 20 µl of 1× ligation buffer (Sigma), heated at 55°C for 1 min, and cooled down to 14°C within 1 h. After adding 3 µl of T4 ligase (5 units/µl) to the mixtures individually, ligation was carried out at 14°C overnight. The ligated DNA mixtures were purified using a PCR Cleanup Kit (Qiagen) and resuspended in 100 µl of 10 mM TE buffer (pH 7.0). The PCR

Table 4.1. The sequences of oligonucleotides used for RaSH and PCR

Oligonucleotides	Sequences (5' -3')	Usages (expected size, bp)
XDPN-18*,	CTGATCACTCGAGAGAGATC	Amplification of the digested genomic DNA
XDPN-14*, XDPN-12*,	CTGATCACTCGAGA GATCTCTCGAGT	
F5f F5r	ATATCGTGTCTGCTAAAGAGC ATCGCTGATGAAATCAAGTC	Clone F5 (483)
F10f F10r	GATAGTGCATAATAAAAGCGAG CCACCACAAAAAAGCCTTATTC	Clone F10 (181)
F17f F17r	ACGCATGGCAAAGTTTTACG TCTATCAATTCAAACCAGCCC	Clone F17 (154)
P18f P18r	CCTCTTTTAAAATTTGAAC ACCCTTTCTGTCGCTATC	Clone P18 (97)
P32f P32r	CCTAAAATTGCAAAAAATGC CCATTGATACAAAAATTACG	Clone P32 (135)
F1 [#] B1 [#]	ATGCATGTGTCTGACTGTCTGA ATTTGCGGTAATATGTCGTGC	<i>cagA</i> (349)
CAG7A CAG7B	TAACGCCACTAATGTTGATG TTCTTAAGAGCGATTTCACTG	<i>cag7</i> (273)
CAGWF [§] CAGTR [§]	AAAAATTGACCRCTCAATC GCTTTAGCTTCTGAYACYGC	<i>cagA</i> 3' region
VA1F [¶] VA1XR [¶]	ATGGAAATACAACAAACACAC CCTGARACCGTTTCCTACAGC	<i>vacA</i> s region (176, 203)
HPMGF [¶] HPMGR [¶]	CACAGCCACTTTCAATAACGA CGTCAAAATAATTCCAAGGG	<i>vacA</i> m region (401, 476)
Cag2F Cag2R	GTGGTTTATCACTCACTGCATG TATGAGCGATACAGCGGTTGCTAG	<i>cag2</i> (431) [□]

The sequences of the oligonucleotides are from *, Jiang *et al.*, 2001; §, Yamaoka, *et al.*, 1998; ¶, van Doorn *et al.*, 1998; # Tummuru, *et al.*, 1993. □ The expected product length is 767 bp and 440 bp in strain J99 and NCTC 11638, respectively.

reaction mixtures were set up as follows: 1 µl of the ligated DNA mixture, 1 µM dNTPs, 10 µM XDPN-18 (Table 4.1), and 1.5 units of *Taq* polymerase (Qiagen). The parameters for PCR were 1 cycle for 5 min at 72°C, followed by 35 cycles for 1 min at 94°C, 1 min at 55°C, 1 min at 72°C, and 1 cycle for 5 min at 72°C. The PCR products were pooled and purified. The amplicons from driver DNA were directly used as the driver DNA library. Tester PCR products (10 µg) were further digested with *Xho*I, purified and resuspended in 10 mM Tris·HCl. The tester DNA library was then constructed.

Subtractive hybridisation: DNA (100 ng) from the tester DNA library was mixed with 6 µg DNA of the driver DNA library in 10 µl of hybridisation buffer [0.5 M NaCl/50 mM Tris, pH 7.5/0.2% SDS/40% (v/v) formamide]. Following boiling for 5 min, the mixture was hybridised at 42 °C overnight. The hybridised mixture was then phenol/chloroform extracted, ethanol precipitated and resuspended in 20 µl of 10 mM TE buffer (pH 7.0). The subtracted library was then generated.

Cloning and sequencing Four microliters of the hybridised products was ligated with 100 ng of *Xho*I-digested pCRII vectors (Invitrogen) for 1 h at RT and transformed into TOPO 10F *E. coli* bacterial cells. Bacterial colonies were picked randomly and grown in liquid media. Plasmids were extracted from bacterial cells from 14-hour cultures and digested with *Xho*I. The inserts were purified from the agarose gels and spotted onto BrightStar-plus nylon membranes (Amibon). The membranes were probed with DIG-Labelled tester or driver genomic DNA digested with *Sau*3 AI and detected using DIG-high prime DNA labelling and detection starter kit (Roche) following the manufacturer's instructions. The candidate clones specific for tester strain C1 were

selected for determination of their sequences. Sequences were determined by the service provided by MWG-Biotech Co. Sequence analysis was performed as described in Chapter 2.2.8.

4.2.3. Southern blot and PCR amplification

The correct identification of the strain-specific sequences for tester strain C1 was verified by Southern blot and/or PCR analyses. Southern blot was carried out as described in Chapter 2.2.6. The distribution of these sequences in strains of *H. pylori* was determined by PCR amplification. The sequences of the primers used are shown in Table 4.1.

PCR was carried out as described in Chapter 2.2.7 to detect *cagA*, *cag7* and *cag2*. The sequences of the primers used were shown in Table 4.1. The annealing temperature was 55 °C for *cagA*, 58 °C for *cag2* or 54 °C for *cag7*. The genotypes of *vacA* were determined as described in a previous report by van Doorn *et al* (1998a). Briefly, the PCR mixture was comprised of 10 X PCR buffer, 1.5 mM MgCl₂, 200 µM each of deoxynucleotides, 0.2 µM of the forward and reverse primers, and 1.5 Unit of Taq DNA polymerase in a final volume of 50 µl. DNA templates (20-50 ng) were included. The PCR program comprised 5 min of predenaturation at 94°C, followed by 40 cycles of 30 s at 94°C, 45 s at 50°C, and 45 s at 72°C and a final incubation at 72°C for 5 min. PCR products were separated on 2% agarose gels and visualised with ethidium bromide staining.

4.2.4. Statistical analysis

Statistical analysis was performed using SPSS V/O statistical analysis software. The chi-square test or the Fisher's exact test was used where applicable. A p value lower than 0.05 was considered as a statistically significant difference.

4.3. Results

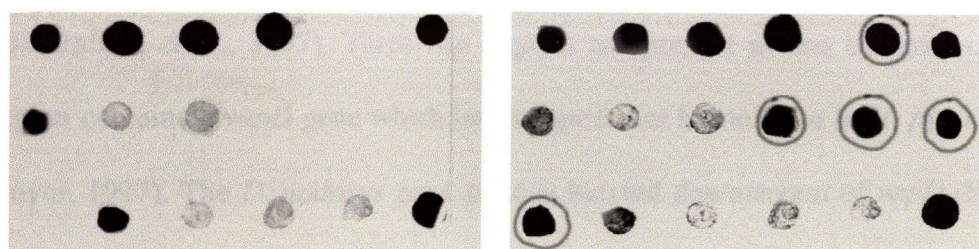
4.3.1. Genotyping the tester and the driver strains

PCR analysis demonstrated both tester strain C1 and driver strain 98587 possessed *cagA*, *cag7* and *cag2* genes, indicating that they had an entire *cag* pathogenicity island. The *vacA* genotypes determined by PCR were *s1/m1*, therefore, the tester and the driver strains were both considered to be virulent strains (Blaser and Berg, 2001).

4.3.2. Identification of sequences specific for the tester strain C1

RaSH employs a cloning based strategy to screen a subtracted hybridisation library and subsequently identify target sequences. The genomic contents of the tester and the driver strains were compared using RaSH to identify sequences specific to strain C1 from a cancer patient. To simplify the procedures, attempts were made to construct both tester and driver DNA libraries without PCR amplification. However, in a total of three attempts, no recombinant colonies were produced after subtraction hybridisation (data not shown). Therefore, the DNA libraries were constructed with PCR amplifications.

Following subtraction hybridisation, the hybridised mixture was cloned into pCR II vectors. A total of 53 colonies was randomly selected to screen putative sequences specific to the tester strain by Dot blot analyses (Figure 4.2). Based on the results of Dot blot analyses, thirty-three clones were judged to be specific to the tester strain C1 and the sequences of 11 clones were determined. Southern blot and/or PCR demonstrated that these 11 clones were specific to tester strain C1.



A

B

Figure 4.2. Representative dot-blot analysis of products of subtractive hybridisation by using genomic DNAs of *H. pylori* strains C1 as tester and 98587 as driver. Plasmid DNAs from candidate clones were digested. The inserts were gel purified and spotted onto nylon membranes, probed with Dig-labelled genomic DNAs from driver (A) or tester (B). The clones that were specific to tester strain C1 are circled.

4.3.3. Analysis of the identified sequences specific to strain C1

The identified sequences specific to tester strain C1 were analysed against the genome sequences of *H. pylori* 26695 and J99 and the public database available on the Internet. These sequences were divided into three groups (Table 4.2). Group I included four genes specific to strain C1. These genes specific to strain C1 showed no significant DNA homology to the sequenced strains 26695 and J99 or only had DNA homologues in strain J99. Clones P32 and F5 had no homologues in the two sequenced strains. Figure 4.3 shows their deduced amino acid sequences and the alignment of Clone P32 with GepA of *Dichelobacter nodosus*, a bacterium that causes ovine footrot (Bloomfield *et al.*, 1997). Three *vap* regions are present in this bacterium. Each *vap* region contains a *vapD* gene which has a significant homologue in *H. pylori* (Cao and Cover, 1997). The *D. nodosus* *gepA* gene is located downstream of *vapD* (Bloomfield *et al.*, 1997). The G+C contents of Clone F5 and P32 were 34% and 30%, respectively, which were lower than the average G+C content (39%) of the *H. pylori* genome.

Group II consisted of three clones. Sequence analysis using the ORFinder program revealed that they were non-coding sequences specific to strain C1. These clones showed DNA matches to the intergenic regions present in either strain 26695 or J99 (Table 4.2). Figure 4.4 shows the location and flanking sequences of their homologous DNA in the sequenced strains. The homologous sequences of these three clones are located in the rearrangement loci in the sequenced strains.

Group III included four clones which had protein homologues in both sequenced strains (Table 4.2). For Clone F1 and P18, only patches of their sequences showed

Table 4.2. Characteristics of the sequences specific to *H. pylori* strain C1

Sequences	Size (bp)	Protein Homologues	DNA match in	
			26695	J99
Group I				
P32	178	GepA, accessory genetic element, <i>Dichelobacter nodosus</i>	-	-
F5	502	None	-	-
P29	154	<i>H. pylori</i> specific hypothetical protein,	-	JHP0704
P1	216	ResI, type III restriction enzyme, <i>H. pylori</i>	-	JHP1297
Group II				
F3*	171	None	-	+
F10*	216	None	+	-
F17*	185	None	+	-
Group III				
F1	139	YdiE, putative O-sialoglycoprotein endopeptidase, <i>H. pylori</i>	HP1584 (86/89)	JHP1491 (82/89)
F11	311	RfaD, ADP-L-glycero-D-mannoheptose-6-epimerase, <i>H. pylori</i>	HP0859 (275/286) HP0858 (26/28)	JHP0793 (277/289) JHP0792 (27/28)
F18	165	RpoB, RNA polymerase beta subunit, <i>H. pylori</i>	HP1198 (162/165) 0054 (84/136)	JHP1121 (150/165)
P18	116	MobA, molybdopterin-guanine dinucleotide biosynthesis protein A, <i>H. pylori</i>	HP0769	JHP0706

* Clone F3 showed 93% identity to the intergenic region between jhp919-920, which is located at rearrangement locus 3 (Alm *et al.*, 1999); Clone F10 showed 91% identity to the intergenic region between HP0313-HP0314 which is located within rearrangement locus 1; Clone F17 showed 96% identity of 105 bp to the intergenic region between HP1432-HP1433 which is located within rearrangement locus 7.

A. Clone F5

```
1 gatccattcctatggccaatatcgtgtcgcgctaagagcaaaccaaa
  I H S Y G Q Y R V A K E Q T K
47 caactagaaataatacgcaatactttaagaaaacaatatgaaaat
  Q L E I I R N T L R K Q Y E N
92 cttttgatagaattaagattagacaaacagaagctcagattacag
  L L I E L R L D K Q K L R L Q
137 cttgccccagacttagaaaaaatagattctagaatacgcaagaat
  L A Q D L E K I D S R I R K N
182 atggacaaaatgcatctgcttaaattagcttatataaagtagcttt
  M D K M H L L K L A Y K S S F
227 agggtttttaaagtgtattaaggagcatttagatgaatatgagaaa
  R V L K C I K E H L D E Y E K
272 aaattcccttatgacaatgcccaaagaattgtagccttgtgccaa
  K F P Y D N A Q R I V A L C Q
317 caataccatgaagcgctcacagcgcttgccaagcgctcttaaac
  Q Y H E A L T A H C Q A S L N
362 tttataggaggttaa 376
  F I G G *
```

B. Clone P32

```
2 gatcaattttaagtttagataatcaaaaacctataattaacctaata
  I N L S L D N Q K P I I N L K
47 ttgcaaaaaatgctatatctatatattgctcaggcagaaggcgat
  L Q K M L Y L Y I A Q A E G D
92 tctaaactcattaagaagattttcaagcttgggattatgggcct
  S K L I K E D F Q A W D Y G P
137 gttattcctgatgtttatcgtaatttttgatcaatggatc 177
  V I P D V Y R N F C I N G
```

C. Alignment of Clone P32 with GepA

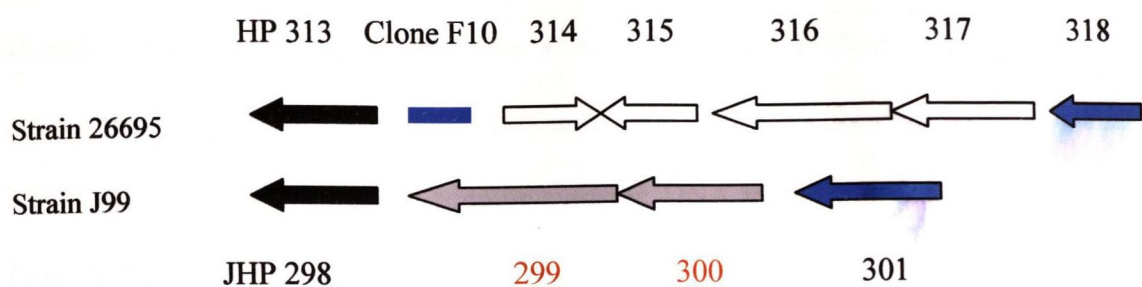
```
P32: 2 INLSLDNQKPIINLKLQKMLYLYIAQAEG--DSK 99
GepA: 16 VRLALADKKPLTNLKVQKLTYLVYGCWLGFNFQ 34
      * * ** *** ** ** *
```



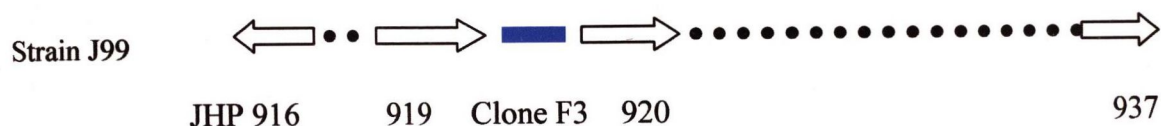
```
P32: 100 LIKEDFQAWDYGPVIPDVY 177
GepA: 35 MFQDRVEAWKYGPVPLVY 68
      ** * * * * *
```

Figure 4.3. Nucleotide sequence and the deduced amino acid sequence of Clone F5 (A), Clone P32 (B) and alignment of the amino acid sequence of Clone P32 with *Dichelobacter nodosus* GepA (C). Amino acid sequences were deduced using ORFinder available from NCBI website. The frames were created at -2 as indicated by the highlighted letters. The identity between Clone P32 and GepA was 39% and the identical amino acid residues between the two aligned sequences are indicated with *.

A.



B.



C.

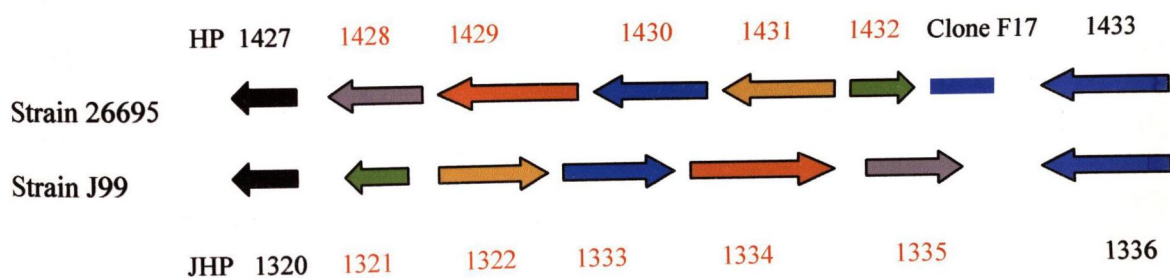


Figure 4.4. Schematic illustration of the locations of clone F10, F3 and F17 in the sequenced strains 26695 and j99 (not to proportion). The identified clones are shown in blue bars; the gene names are shown in the upper lane (for strain 26695) and lower lane (for strain J99). The gene names in red font represent the rearrangement genes. The empty bars represent genes which are strain-specific. The solid bars represent conserved genes between the two strains. The homologues between two strains are shown in the same colour. Arrows indicate the orientation of the gene. Panels A, B and C show the chromosomal rearrangement loci 1, 3 and 7 (Alm, *et al.*, 1999)

significant matches with their corresponding homologues at nucleotide acid level. Multiple alignment of Clone P18 with *mobA* in strain 26695 and J99 revealed that the shared identity was 44.0%. For Clone F18, which was a protein homologue of *H. pylori* RpoB, its DNA sequence significantly matched with not only *rpoB* but also HP0054. The latter encodes an adenine/cytosine DNA methyltransferase and lacks a homologue in strain J99 (Alm *et al.*, 1999).

4.3.4. Distribution of the strain-specific sequences among H. pylori strains.

To determine whether the two novel sequences in group I were common to strains of *H. pylori*, the prevalence of Clones F5 and P32 were determined in 55 strains (including the tester and the driver strains) from patients with GC, IM, DU and gastritis. Clones P32 and F5 were detected in 36.5% (20/55) and 65.5% (36/55) of strains, respectively (refer to the appendix). Interestingly, no strains from patients with IM possessed Clone P32, a homologue of GepA. The prevalence of Clone P32 in strains from DU was significantly higher than those from IM (Figure 4.5). There was no significant difference in the prevalence of Clone F5 in strains from different diseases (Figure 4.5).

A subgroup (19 strains) of these 55 strains was further analysed with respect to other identified strain-specific sequences, *cag* genes and the *vacA* genotypes. The prevalence of Clone F10, F17 and Clone P18 was determined in this subgroup of strains from patients with GC, IM and DU diseases (Table 4.3.). Clones F10 and F17, which had DNA matches with the intergenic fragments, were detected in 63.2% and 10.5%, respectively. A pair of primers was designed to detect *mobA* with sequence resembling Clone P18. Clone P18-like sequence was detected in 83.2% (16/19) strains. The status

Table 4.3. Distribution of the strain-specific sequences and the *vacA* genotypes among strains of *H. pylori*

Strains	P32	F5	F10	F17	P18	<i>cagA</i>	<i>cag7</i>	<i>cag2</i>	<i>vacA</i> <i>s</i>	<i>m</i>	Groups
C1	+	+	+	+	+	+	+	+	1	1	GC
41289	-	-	-	-	+	+	+	+	1	1	
7109	+	+	-	-	+	+	+	+	1	2	
863	-	-	+	-	+	+	-	+	1	1	
77	-	-	+	-	+	+	-	+	1	2	IM
5612	-	+	+	-	+	+	+	+	1	1	
8248	-	-	+	+	+	+	+	+	1	1	
5536	-	-	+	-	+	+	+	+	1	1	
5283	-	+	-	-	+	+	+	+	1	1	
7231	-	-	+	-	+	+	+	+	1	1	
3517	-	+	+	-	-	+	+	+	1	1	
8260	-	-	-	-	+	+	+	+	1	1	
3655	-	+	+	-	-	+	+	+	1	2	DU
3519	+	+	-	-	+	+	+	+	1	1	
3612	-	-	+	-	+	+	+	+	1	2	
3582	+	+	+	-	+	+	+	+	1	2	
3568	+	-	-	-	+	+	+	+	1	2	
3526	-	-	+	-	+	+	+	+	1	2	
98587	-	-	-	-	-	+	+	+	1	1	

+, presence of the sequences; -, absence of the sequences; GC, gastric cancer, IM, intestinal metaplasia; DU, duodenal ulcer.

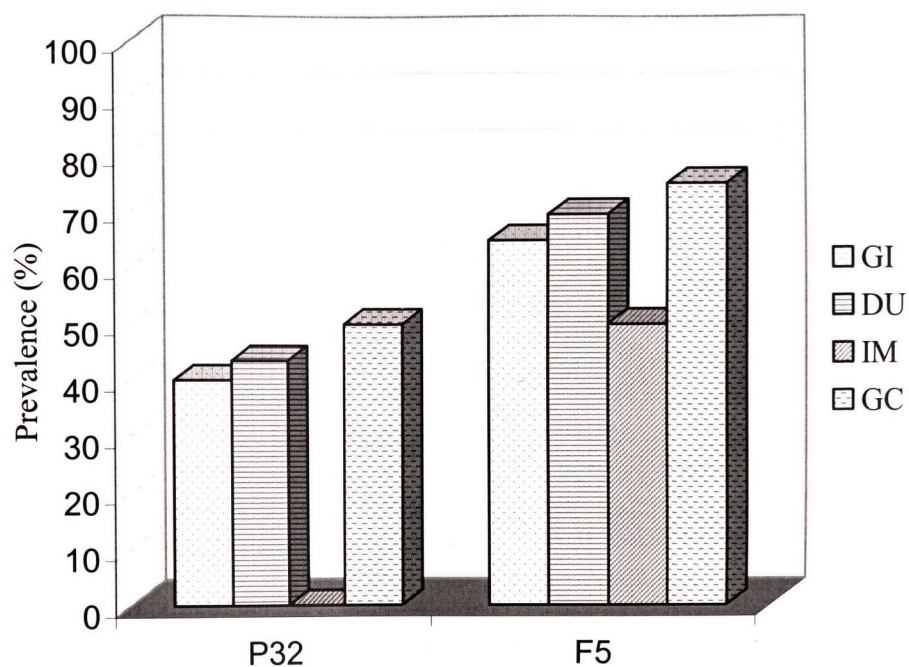


Figure 4.5. Prevalence of clone P32 and F5 in strains from different diseases.

The prevalence of clone P32 in strains from DU was significantly higher than that in strains from IM. GI gastritis; DU duodenal ulcer; IM intestinal metaplasia; GC gastric cancer.

4.4. Discussion

RaSH, originally used for RNA comparison (Jiang *et al.*, 2000), was applied here to identify genomic differences between two strains of *H. pylori*. This technique differs from other subtractive hybridisation approaches not only in the principles, but also in its simplified procedures as only one round of PCR amplification and one round of hybridisation is required. Attempts to construct both tester and driver DNA libraries without PCR amplification failed. This may be attributed to the high level of complexity of the genomic DNA which can be significantly reduced by PCR amplification (Hampson, *et al.*, 1996). Therefore, amplification is required for the construction of a DNA library in this technique.

Genotyping demonstrated that the tester strain C1 and the driver strain 98587 both were virulent strains. The genomic contents of strain C1 were then compared with driver strain 98587. Three groups consisting of 11 sequences were identified to be specific to *H. pylori* strain C1 isolated from a patient with gastric cancer. The sequences from groups I and II showed no homology in one or both sequenced strains of *H. pylori*, indicating that they represent genes or DNA regions that are strain-specific in their distributions (Alm, *et al.*, 1999; Tomb, *et al.*, 1997). The nature of the identified sequences from group III is complex. They are most likely to represent genes that are carried in all strains of *H. pylori* (including the driver strain) but are substantially divergent in their sequences among strains (Akopyants *et al.*, 1998b). For example, Clone P18 and F1 contained only patches of sequence matched with their homologues. The identified sequences from group III may also reflect the complicated genomic diversity between the tester and the driver strains. Clone F18 was homologous to RpoB which is undoubtedly present in all strains of *H. pylori*. The sequence of

Clone F18, however, also showed significant DNA matches with HP0053. HP0053 coding for an adenine/cytosine DNA methyltransferase is absent from strain J99 (Alm, *et al.*, 1999). This suggests that the identification of Clone P18 reflects the differences in HP0053 or those imperfectly matched genes, rather than *rpoB*, between the tester and the driver strains.

4.4.1. Strain-specific genes

Comparison of the genomic contents between the tester strains C1 and the driver strain 98587 demonstrated that 4 genes of group I were specific to strain C1 which was isolated from a patient with adenocarcinoma. Clones F5 and P32 were novel sequences. The lower G+C contents of Clone F5 and P32 indicate that both sequences were possibly acquired through horizontal gene transfer (Ochman *et al.*, 2000). Both clones were highly prevalent in the clinical isolates, with Clone P32 in about one third and F5 in two thirds of the detected strains. These results indicate that, for the species of *H. pylori*, additional genes exist that are absent from the sequenced strains. In agreement with this, *pap57* encoding a protein containing an EAL domain (described in Chapter 3) and IS607 have been identified (Kersulyte, *et al.*, 2000). These facts suggest that, for a given bacterial species, particularly *H. pylori*, genome sequencing of one or a few strains is unlikely to determine all the coding genes. A “species genome”, proposed by Lan and Reeves (2000) and further explained by Boucher *et al.*, (20001), is required to accommodate those new sequences. A “species genome” comprises all the genes present in 99% of the characterised strains of a species, whereas, genes less frequent than 1% should be considered “foreign” or on the decline (Boucher *et al.*, 2001). The “species genome” consists of a core set (>95%) and an auxiliary set of genes (1-95%). The core set are phylogenetically stable and constant in gene content

because these genes are essential for survival or for the interaction with products of other genes. These genes are usually not acquirable. On the other hand, the auxiliary set are usually acquired from other species or generated through significant intragenomic recombination conferring new traits for a specific strain (Boucher *et al.*, 2001). Clone F5 and Clone P32 were frequently seen in the clinical isolates of *H. pylori*, indicating that they belong to the auxiliary set and may confer new adaptive traits for those strains.

Clone P32 is a homologue of GepA (genetic element protein A) in *D. nodosus*. It has been suggested that *gepA*, together with *intB*, *regA*, *gepB* and *gepC* genes probably constituted a conjugative transposon (Bloomfield, *et al.*, 1997). *H. pylori* is naturally competent and possesses a number of genes involved in the acquisition of foreign DNA and thus, diversification of the *H. pylori* genome (Hofreuter, *et al.*, 2001). These include the *comB* genes that constitute a type IV secretion system-like structure (Chang *et al.*, 2001; Hofreuter *et al.*, 2001), type II restriction-modification systems (Ando *et al.*, 2000; Donahue *et al.*, 2000), insertion elements (Kersulyte *et al.*, 2000) and individual genes (Chang *et al.*, 2001; Smeets *et al.*, 2000). It will be interesting to determine whether *H. pylori* contains homologues of *intB*, *regA*, *gepB* and *gepC* that are potentially involved in conjugative gene transfer. An intriguing observation is that the Clone P32-positive strains from patients with intestinal metaplasia were less common than those with duodenal ulceration, gastritis and gastric cancer, although a statistically significant difference was only found between IM and DU. The interpretation of the data include (1) Gain and/or loss of genes frequently occur in *H. pylori* when colonising the human stomach (Janssen *et al.*, 2001; Israel *et al.*, 2001). Clone P32 would be removed from the genome to avoid a fatal event if the presence of

this sequence is detrimental for the bacterium colonising environments created by mucosal metaplasia; (2) To maintain the appropriate size of the genome, *H. pylori* may delete Clone P32 since it is no longer required for adaptation to IM associated environments (Fitzgerald & Musser, 2001); (3) Relatively small numbers of IM samples were examined. Therefore, further studies are indicated.

Clone F5 showed no homologue in the public databases and, therefore, belongs to a class of genes coding for *H. pylori*-specific hypothetical proteins (Alm *et al.*, 1999; Janssen *et al.*, 2001). Clone P29, a homologue of JHP 704, also belongs to this class. This class of genes is usually introduced into *H. pylori* by lateral gene transfer (Janssen *et al.*, 2001). These genes facilitate adaptation of the bacterium to the particular host individual or changing environment (Jain *et al.*, 2002; Ochman *et al.*, 2000).

Clone P1 was homologous to ResI that is classified as a type III restriction enzyme. The results are in agreement with the fact that the DNA restriction-modification (R-M) systems are extensively diverse among strains of *H. pylori* (Akopyants *et al.*, 1998; Lin *et al.*, 2001; Salama *et al.*, 2000). These variable R-M loci may be related to (1) virulence such as *iceA* and *hsdA* (Bjorkholm *et al.*, 2002; Peek *et al.*, 1995) (2) facilitation of recombination by generation DNA ends through cutting DNA at specific sites (Salama *et al.*, 2000); (3) defence against foreign DNA by digestion of DNA; or to be “selfish” because they do not provide any advantage to the host strains (Akopyants *et al.*, 1998).

4.4.2. Chromosomal rearrangement

The group II sequences specific to strain C1, Clones F3, F10 and F17, had DNA matches with the intergenic regions in one of the sequenced strains. These intergenic regions are located within the chromosomal rearrangement loci (Figure 4.4) (Alm *et al.*, 1999). Therefore, these strain-specific sequences may represent the occurrence of rearrangements in the tester and the driver strains.

Clone F3 had a DNA match with the intergenic region between jhp919 and 920. The flanking genes (jhp916-937) of ~ 24kb are located in the rearrangement locus 3 which is probably generated by integration of a large fragment of foreign DNA (Alm *et al.*, 1999). This region was known as a “plasticity zone” which contains 46% and 48% of genes that are specific to 26695 and J99, respectively (Alm *et al.*, 1999; Occhialini *et al.*, 2000). The results demonstrated that this rearrangement also occurred in clinical isolates.

Clone F10 and Clone F17 had DNA matches with the intergenic regions between HP0313-0314 and between HP1432-HP1433, respectively. These intergenic regions are located within the rearrangements 1 (for Clone F10) and 7 (for Clone F17), respectively. The generation of both arrangements is associated with repeat elements (Alm *et al.*, 1999). Clones F10 and F17 were present in a fraction of the examined strains (Table 4.3). These results demonstrated that rearrangements 1 and 7 frequently occur in the clinical isolates of *H. pylori*.

4.4.3. Allelic polymorphism

Strain-specific clones from group III were partially matched to the DNA sequences of genes in the two sequenced strains. They are most likely to represent genes with significantly divergent sequences. These divergences in the sequences are potentially important in the evolution of bacteria since they may facilitate the formation of new genotypes by recombination in the case of mixed infection and sometimes generate dramatic changes in bacterial phenotypes (Falush *et al.*, 2001; Fitzgerald & Musser, 2001; Joyce *et al.*, 2002). For *H. pylori* which inhabits the hostile human stomach and colonises genetically diverse hosts, these changes may speed adaptation of the bacterium to the individual host.

4.4.4. Extensive genomic divergence

Comparison of the genomic contents between a strain from a patient with gastric cancer and one from duodenal ulcer demonstrated the three types of genetic diversity. The diversity is manifested as variable status of some genetic loci (strain-specific genes), chromosome rearrangement, and divergence in the sequences of certain genes (allelic polymorphism). To further analyse the genetic contents in the clinical isolates, five identified strain-specific clones representing these different types of diversity, the status of *cagA*, *cag7* and *cag2* along with the *vacA* genotypes were determined in 19 strains from patients with GC, IM and DU (Table 4.3). Each strain exhibited a distinct distribution profile of these variables. This indicates that the *H. pylori* genome is highly divergent. The substantial genetic diversity reflects the high rate of mutation, free recombination, frequent acquisition of foreign DNA in *H. pylori* (Chang *et al.*, 2001; Hofreuter *et al.*, 2001; Suerbaum, *et al.*, 1999; Wang, *et al.*, 1999). The extensively

divergent genome is important for the bacterium to rapidly adapt to the changing environments and the individual hosts.

CHAPTER 5
HETEROGENEITY IN THE *cag* PATHOGENICITY ISLAND OF
HELICOBACTER PYLORI

5.1. Introduction

H. pylori produces a protein of varied size of 120, 128, or 130 kDa (Covacci *et al.*, 1993). This protein is highly immunogenic and can be recognised by the sera of infected individuals (Tummuru *et al.*, 1993). Strains which produce this antigen usually exhibit a cytotoxic phenotype causing vacuole formation in HeLa cells. Covacci and colleagues (1993) used a pool of patient sera to screen an expression library constructed from the digested genomic DNA of strain CCUG 17874 (or NCTC 11638). They found the antigen is a putative 1147-aa protein with an estimated molecular mass of 128,012.73 Dalton (Da). The gene encoding this antigen is named as *cagA* (cytotoxin associated gene A) since it does not have vacuolating activity (Tummuru *et al.*, 1993).

Further studies on *cagA* led to the identification of the *cag* pathogenicity island (PAI), a large DNA region containing *cagA*, which is inserted between the glutamate racemase (*glrA*) gene and HP0519 (Tomb *et al.* 1997). The sequences of the *cag* PAI have been determined in strains 26695, J99 and NCTC 11638 (Akopyants *et al.* 1998; Alm *et al.*, 1999; Censini *et al.*, 1996; Tomb *et al.* 1997). The *cag* island is flanked by a 31 bp direct repeat (TTACAATTTGAGCCATTCTTTAGCCTGTTTT), one copy of which is present at the empty site of *cag*⁻ strains. The presence of the direct repeats suggests that the *cag* PAI might be gained or lost by site-specific or generalised recombination involving this terminal repeat. The length of the *cag* PAI is approximately 40 Kb. The island consists of 27 genes, in addition to (1) insertion elements (IS 605) which are not common and not related to *cag* associated pathogenicity; (2) the predicted ORFs including HP0533 (Tomb *et al.*, 1997), *cagR* and *cagB* (Censini *et al.*, 1996), which are unlikely to encode proteins (Alm *et al.*, 1999); and (3) HP0548/0549, a homologue of

DNA helicase, which contains a frameshift and is probably inactive (Tomb *et al.*, 1997). The orientation and organisation of the *cag* genes are shown in Figure 5.1.

The *cag* PAI encodes a type IV secretion system, a conjugal pilus-like structure, responsible for uptake or export of DNA or injection of proteins into host cells (Baron *et al.*, 2002). In *H. pylori*, the system translocates CagA into the host cells and induces growth factor-like changes in the epithelial cells (Chapter 1.9.1). The *cag* PAI induces transcription of IL-8 in epithelial cells via a CagA-independent pathway.

The *cag* genes may be organised as a contiguous structure or may be separated into two regions: *cag* I (the left part of *cag* including *cagA*) and *cag* II (the right part), by insertion element (IS) 605 or a large DNA region (Alm *et al.*, 1999; Censini *et al.*, 1996). Certain genes in the *cag* I region or the whole *cag* I region may be absent from the *cag* PAI (Censini *et al.*, 1996). A partial *cag* PAI, or a *cag* PAI lacking some genes, is deficient in the ability to induce IL-8 and translocate CagA (Akopyants *et al.*, 1998; Fischer *et al.*, 2001). The virulence of strains containing a partial *cag* PAI, therefore, may be reduced. Whether the reduced virulence in these strains will affect the clinical outcomes of infected individuals requires examination.

Proper phosphorylation of the translocated CagA at tyrosine (Y) residues is essential for induction of growth factor-like phenotype changes in the host cells. Three tyrosine residues, Y918, Y972 and Y1039 (the positions refer to these on the CagA sequence of strain 26695) which are located at the C-terminus of CagA, have been demonstrated to be phosphorylated (Backert *et al.*, 2001; Odenbreit *et al.*, 2000; Puls *et al.*, 2002; Stein *et al.*, 2002). The consensus sequence of a motif to be phosphorylated is EPIYA or RX-

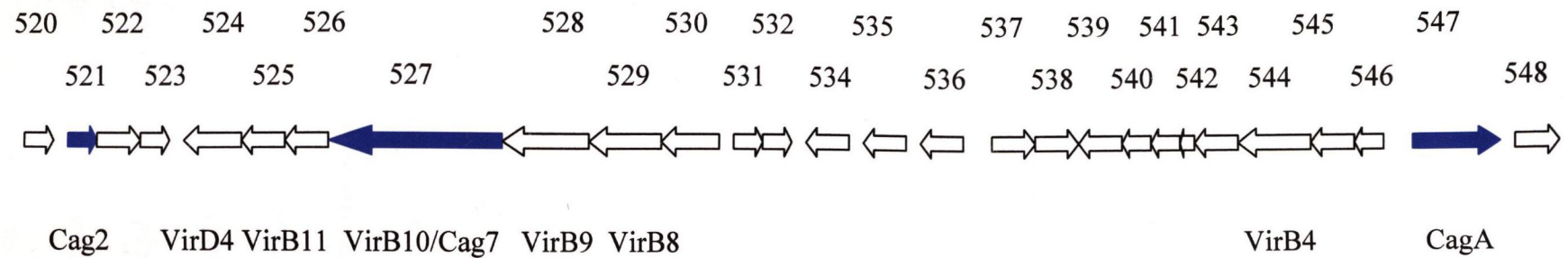


Figure 5.1. Schematic illustration of the orientation and organisation of the *cag* PAI according to Tomb *et al.*, 1997. The gene numbers as annotated in strain 26695 are displayed at the upper part. Arrows indicate the orientation of genes. Genes examined in this study are represented with solid bars. Vir B homologues and the name of the genes examined are displayed at the lower part.

K (2x3) R- [D/E](2x3) R-Y. Efficiency in CagA phosphorylation and the resultant activity of CagA varies among strains. Variation of the nucleotide sequence at the *cagA* 3' region was initially observed when *cagA* was first characterised (Covacci, *et al.*, 1993). Duplication or deletion of an element of 102 bp (named as the γ segment in this study) at the *cagA* 3' region was observed among *H. pylori* strains (Covacci, *et al.*, 1993). Another repetitive element of 57 bp (named as the β segment in this study) is also responsible for variation at the 3' region (Yamaoka, *et al.*, 1999). Since the tyrosine motifs to be phosphorylated are located within this region, this variation may alter numbers of the available motifs and thus, the activity of CagA (Evans *et al.*, 2001).

Cag7, a homologue of *Agrobacterium tumefaciens* VirB10, is essential for translocation of CagA and induction of IL-8 release from the host cells. The *cag7* (HP0527) gene is located at the *cag* II region (Censini *et al.*, 2001; Fischer *et al.*, 2001). In strain 26695, the 5' region (from nucleotide position 25 to 804) of *cag7* is composed of two copies of a repetitive element of 390 bp (Tomb, *et al.*, 1997). In strain J99 and NCTC 11638, however, the corresponding region of *cag7* contains only one copy of the repetitive element (Akopyants *et al.*, 1998; Alm *et al.*, 1999). Therefore, it is very likely that expansion and contraction of the repetitive elements would generate variation at the 5' end of *cag7* in clinical isolates of *H. pylori*.

The aims of this section of the study were (1) to investigate the heterogeneity in the composition of the *cag* PAI by detection of *cagA*, *cag7* and *cag2*; (2) to analyse the variations at the *cagA* 3' region and the *cag7* 5' region among strains from different gastroduodenal diseases: gastritis, duodenal ulcer (DU), intestinal metaplasia (IM) and distal gastric adenocarcinoma (GC).

5.2. Materials and Methods:

5.2.1. *H. pylori* strains

A total of 86 clinical isolates of *H. pylori* were examined in the study (refer to appendix). These isolates were from four groups: gastritis alone (n=34), IM (n=8), DU (n=40) and GC (n=4). *H. pylori* strain NCTC 11638 was used as a reference strain. These strains were recovered from glycerol stocks and subcultured on blood agar plates under microaerobic conditions at 37 °C for 3-5 days.

5.2.2. Polymerase chain reaction (PCR)

Genomic DNA was extracted using the method described in Chapter 2.2.1, or using Dneasy® tissue kit (Qiagen) in accordance with the manufacturer's instructions. The quantity and quality of the DNA samples were analysed as describe in Chapter 2.2.1. The *cagA*, *cag7* and *cag2* genes were detected using PCR conditions described in Chapter 4.2.3. NCTC 11638 served as a positive control. To analyse the variations of the *cagA* 3' region and the *cag7* 5' region, the variable regions were amplified. Primers CAG7F/CAG7R were designed based on multiple alignment of the three available *cag7* sequences in the public databases and used to amplify the 5' region of *cag7*. The sequence of CAG7F was 5'-AGGCCTGGCCAATTTTACGC-3', and the sequence of CAG7R was 5'-GATCGTTTAAAGCTGTGGAC-3'. The cycling conditions included one cycle of 94 °C for 5 minutes; 5 cycles of 94 °C for 45 seconds, 63 °C for 60 seconds and 72 °C for 60 seconds; 5 cycles of 94 °C for 45 seconds, 60 °C for 60 seconds and 72 °C for 60 seconds; 20 cycles of 94 °C for 45 seconds, 58 °C for 60 seconds and 72 °C for 60 seconds; and finally one cycle of 5 minutes at 72 °C. The *cagA* 3' region was amplified as described in the report by Yamaoka *et al.* (1998).

5.2.3. Southern hybridisation

Southern hybridisation was performed mainly as described in Chapter 2.2.6. The probes were generated from the amplicons of the conserved regions of *cag7* and *cagA* from strain NCTC11638. The amplicons were purified using according to the manufacturer's instructions and subsequently labelled with digoxingenin (DIG). The blots were detected using a DIG-high prime DNA labelling and detection starter kit according to the manufacturer's instructions (Roche).

5.2.4. Cloning and sequencing

The appropriate amplicons were purified, cloned into pCR 4.0 sequencing vectors (Invitrogen) and sequenced as described in Chapter 2.2.8. The sequences were deposited in the GenBank and the accession numbers for *cagA* are AF427098, AF427099 and AF427100, for *cag7* the accession number is AF300828.

5.2.5. Sequence analysis and statistical analysis.

Sequence analysis was performed as described in Chapter 2.2.8. A tyrosine phosphorylation site (TPS) was identified if a motif with sequence EPIYA or RX-K (2x3) R- [D/E](2x3) R-Y was present around position Y918, Y972 or Y1039. Statistical analysis was performed using SPSS V/O statistical analysis software. The chi-square test or the Fisher's exact test where applicable was used. A *p* value lower than 0.05 was considered as a statistically significant difference.

5.3. Results:

5.3.1. Prevalence of *cagA*⁺ strains

The presence of *cagA* was tested in all 86 strains using PCR with primers F1 and B1. In 71 strains, PCR amplification yielded expected amplicons of 349 bp. To exclude possible false negative results, Southern blot was used complementarily to detect *cagA* in those 15 *cagA*⁻ strains determined by PCR. The presence of *cagA* in five of them was confirmed (refer to appendix for details). Thus, *cagA* was present in 76 (88.4%) isolates of *H. pylori*. As shown in Figure 5.2, the prevalence of *cagA*⁺ strains in the DU group was significantly higher than that in the gastritis group ($p < 0.005$). All strains from IM and GC were *cagA*⁺. The prevalence of *cagA*⁺ in IM or GC groups, however, was not statistically significantly different from that in the gastritis group ($p > 0.05$) (Figure 5.2).

5.3.2. Composition of the *cag* PAI

To analyse the composition of the *cag* PAI, the status of *cagA* at the right end of the *cag* PAI, *cag2* at the left end and *cag7* at the middle part were analysed in all 86 strains. Absence of each gene was confirmed by repeated PCR and Southern blot. Deletion of *cagA* and/or *cag7* was observed (Table 5.1). Of the 10 *cagA*⁻ strains, 5 strains possessed *cag2* or *cag7*, or both. Thus, a total of 81 (94.2%) strains possessed the *cag* PAI. For details of the genotyping results, refer to the appendix.

A partial *cag* PAI was identified if one or more of the three genes (*cagA*, *cag2* and *cag7*) were absent. Among the 81 strains containing the *cag* PAI, 10 (12.3%) possessed a partial *cag* PAI. The prevalence of the strains carrying a partial *cag* PAI in DU (5.0%), IM (12.5%) and GC (25.0%) groups was lower than that in gastritis group

Table 5.1. Status of *cagA*, *cag7* and *cag2* in strains from different disease groups

Status of <i>cag</i> genes			Gastritis (n=34)	DU (n=40)	IM (n=8)	GC (n=4)	Total (n=86)
<i>cagA</i>	<i>cag7</i>	<i>cag2</i>					
+	+	+	23	38	7	3	71
+	-	+	2	1	1	1	5
-	+	+	4	1	0	0	5
-	-	-	5	0	0	0	5

DU duodenal ulcer; IM intestinal ulcer; GC gastric cancer.

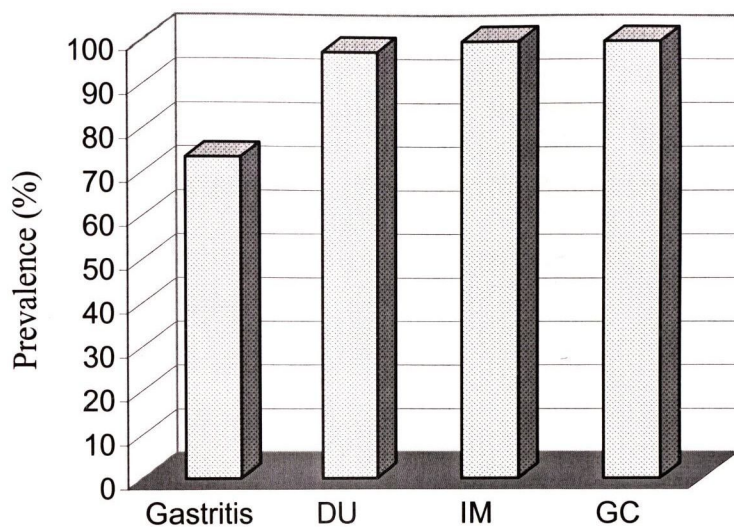


Figure 5.2. Distribution of *cagA*⁺ strains among different diseases. DU, duodenal ulcer; IM, intestinal metaplasia; GC, gastric cancer.

(33.4%). Statistic analysis, however, showed that the prevalence of the strains carrying a partial *cag* PAI in DU, IM or GC were not significantly different than that in gastritis (all *p* values were greater than 0.05).

5.3.3. Variation at the 3' region of *cagA*

To analyse the diversity of the *cagA* 3' region, the 3' end was amplified from those 76 *cagA*⁺ strains. Amplifications failed in two strains and they were excluded from further analysis. The size of the amplicon from strain NCTC 11638 was 531 bp as expected (Censini *et al.*, 1996). The sizes of amplicons from clinical isolates ranged from approximately 530bp to 1350bp (Figure 5.3).

Structure and types of *cagA* 3' region: Twelve amplicons were cloned into the pCR 4.0 sequencing vectors and sequenced at both strands. Multiple alignment of their deduced amino acid sequences showed that the variable length of the *cagA* 3' region was mainly caused by deletions of the α segment, deletion of the β segment and duplication or deletion of the γ segment (Figure 5.4). The β segment was a newly identified segment, while the involvement of both α (57bp) and γ (102 bp) segments was previously suggested (Covacci, *et al.*, 1993; Yamaoka, *et al.*, 1999). The β segment was 105 bp long and located immediately downstream of the α segment.

Based on the name and the number of the segments responsible for the size variation, the *cagA* 3' regions were divided into 6 types. The Greek letters represent the segment causing size variation and the lower case digits following the letters indicate the numbers of the corresponding segments present in the *cagA* 3' region. They were

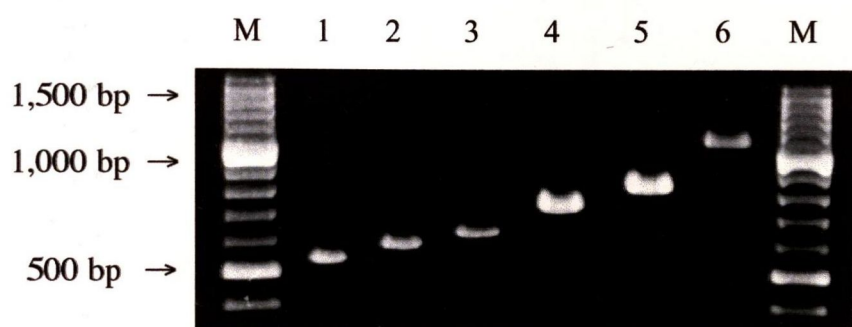


Figure 5.3. Representative agarose gel analysis of the PCR products of the *cagA* 3' region. Lane M, 100 bp DNA markers, Lanes 1-6 show the different sizes of the amplicons from six individual strains which resulted from a varied number of the repetitive elements.

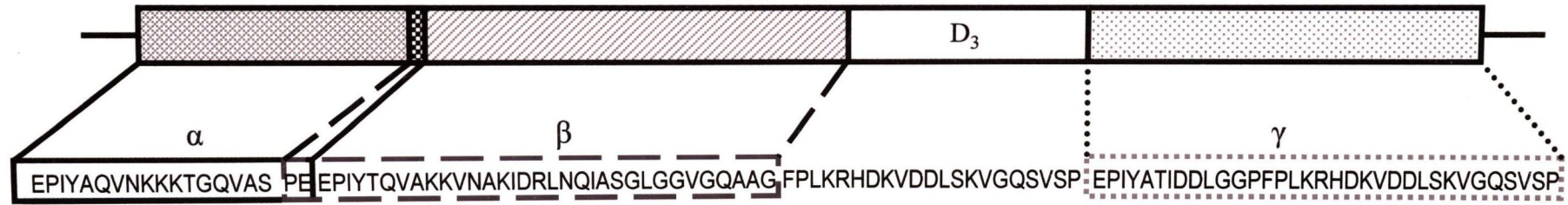


Fig. 5.4. Primary structure of the *cagA* 3' region. The *cagA* 3' region was comprised of three variable segments (α , β and γ) and a conserved segment D_3 . The α and β segments were overlapped with two aa residues. The segments are not represented on a proportional scale. Cross-hatched, α segment; hatched, β segment; dotted, γ segment; open, D_3 segment.

considered as γ_{3+} in cases where 3 or more γ segments were present in the *cagA* 3' region. The structure of each type is illustrated below.

$$\alpha_0 = 582\text{bp} \quad -\beta-D_3-\gamma-$$

$$\beta_0 = 528\text{bp} \quad -\alpha-D_3-\gamma-$$

$$\gamma_0 = 531\text{bp} \quad -\alpha-\beta-D_3-$$

$$\gamma_1 = 633\text{bp} \quad -\alpha-\beta-D_3-\gamma-$$

$$\gamma_2 = 735\text{bp} \quad -\alpha-\beta-D_3-\gamma-\gamma-$$

$$\gamma_{3+} = 837\text{bp} + n(102)\text{bp} \quad -\alpha-\beta-D_3-\gamma-\gamma-\gamma-$$

As the β segment had a similar size to the γ segment, the types β_0 and γ_0 , which resulted from the deletion of β and γ segments respectively, was difficult to discriminate by the amplicon size.

TPS sites at the *cagA* 3' region: The putative TPS sites were analysed in these sequenced strains. The positions corresponding to Y918, Y972 and Y1039 in strain 26695 were named as TPS I, II and III, respectively.

Strains with type β_0 *cagA* lacked TPS I due to deletion of β segment. For the remaining strains, an EPIYT motif instead of EPIYA, was present around the Y918 position. Considering that the +1 and -1 positions of Y918 are required for proper phosphorylation, these strains were considered not to possess TPS I. Therefore, all the sequenced strains lacked TPS I.

TPS II is located within the γ segment and thus may be presented as multiple copies if duplication of the γ segment occurs. Sequence analysis demonstrated that all strains, except for γ_0 type strains, possessed TPS II. The number of TPS II was equal to the number of γ segments excluding strain 94092 (type γ_3). The Y residue in the first γ segment in this strain was replaced with H, but the Y residue was present in the second and third γ segments.

For TPS III, sequence analysis demonstrated six strains possessed this site. The consensus sequence of the motif at this site was DKLKDSTKY. The presence of TPS III was not correlated with any specific *cagA* type. Absence of this site resulted from the replacement of Y with K or N.

Association of the cagA 3' region variation with the clinical outcomes: The predominant genotype among strains was γ_1 and it was seen in 70.3% (52/74) of strains (Table 5.2). The distribution of the *cagA* γ_1 genotype in the duodenal ulcer (76.3%) or intestinal metaplasia group (75.0%) was slightly higher than that in the gastritis group (58.3%), but the differences were not statistically significant ($p>0.05$) (Table 5.2).

5.3.4. Variation at the 5' region of cag7

To analyse the variation at the 5' region of *cag7*, a pair of primers was designed to amplify this region in reference strain NCTC 11638 and the 76 *cag7*⁺ strains. The 5' primer was located at the intergenic region upstream of *cag7*, while the 3' primer was located at the conserved region. The amplicon was 655 bp as expected in the reference strain (Akopyants, *et al.*, 1998). Amplification of the *cag7*⁺ strains yielded products of approximately 650, 1050 and 1500 or 1800 bp (Figure 5.5).

Table 5.2. Prevalence of the *cagA* 3' region genotypes in different patient groups

Genotypes	DU ^a (n=38)	IM ^b (n=8)	Gastritis (n=24)	GC (n=4)	Total (n=74)
α_0	1	0	0	0	1
β_0	1	0	1	0	2
γ_0	1	0	2	0	3
γ_1	29	6	14	3	52
γ_2	3	1	4	1	9
γ_{3+}	3	1	3	0	7

The proportion of each of γ_1 , γ_2 and γ_3 genotypes in duodenal ulcer group ^a and in intestinal metaplasia group ^b were not significantly different from that in gastritis group ($p>0.05$).

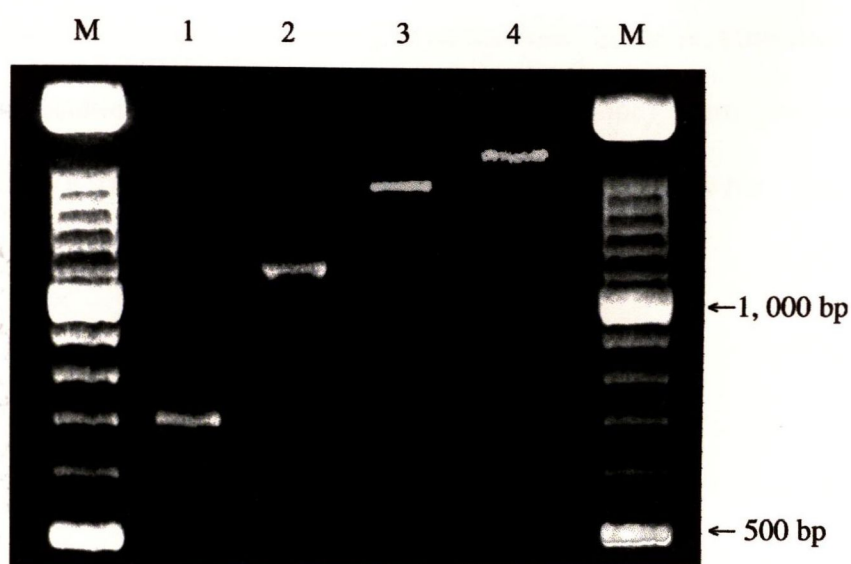


Figure 5.5. Representative agarose gel analysis of the PCR products of the *cag7* 5' region. Lane M, 100 bp DNA ladder; Lane 1-4, the amplicons from individual strains of *H. pylori* represent variable size of the *cag7* 5' region.

Twelve amplicons of the 5' region of *cag7* were sequenced. Analyses of the sequences demonstrated that the variable amplicon sizes were caused by duplication of a repetitive element inserted at amino acid position 9. Typically the repetitive element was composed of 130 aa. The element, however, consisted of 128 aa in two strains (536 and 1174), and 136 aa in one strain (625). The amino acid sequence of the duplicated repetitive elements differed by 0 to 3 amino acid residues. High usage of glutamate (E) (20/130) was observed in the repetitive element. Multiplet count refers to the number of homooligopeptide runs of two or more aa residues. In the repetitive element, at least 5 runs of doublets EE and 3 of doublets KK (lysine) were present. The element resembled a range of proteins with high content of E when blasted against the Swissport protein database.

Based on the number of repetitive elements in the variable region, *cag7* was divided into four types, A (1R, 655 bp), B (2R, 1045 bp), C (3R, 1435 bp) and D (4R or 5R, about 1800 or 2200). None of the amplicons from type D was sequenced and the number of repetitive elements was deduced from the amplicon size. Type B was the predominant *cag7* type, which was found in 73.7% strains. The prevalence of each type did not differ significantly in different groups (Figure 5.6).

5.3.5. The Combined *cagA-cag7* genotypes

Of the 74 strains in which the *cagA* 3' region genotypes have been analysed, *cag7* was absent in 5 strains (Table 5.3). Therefore, a total of 69 strains were analysed for the combined genotypes of *cagA* and *cag7*. The predominant combined genotype was *cagA*

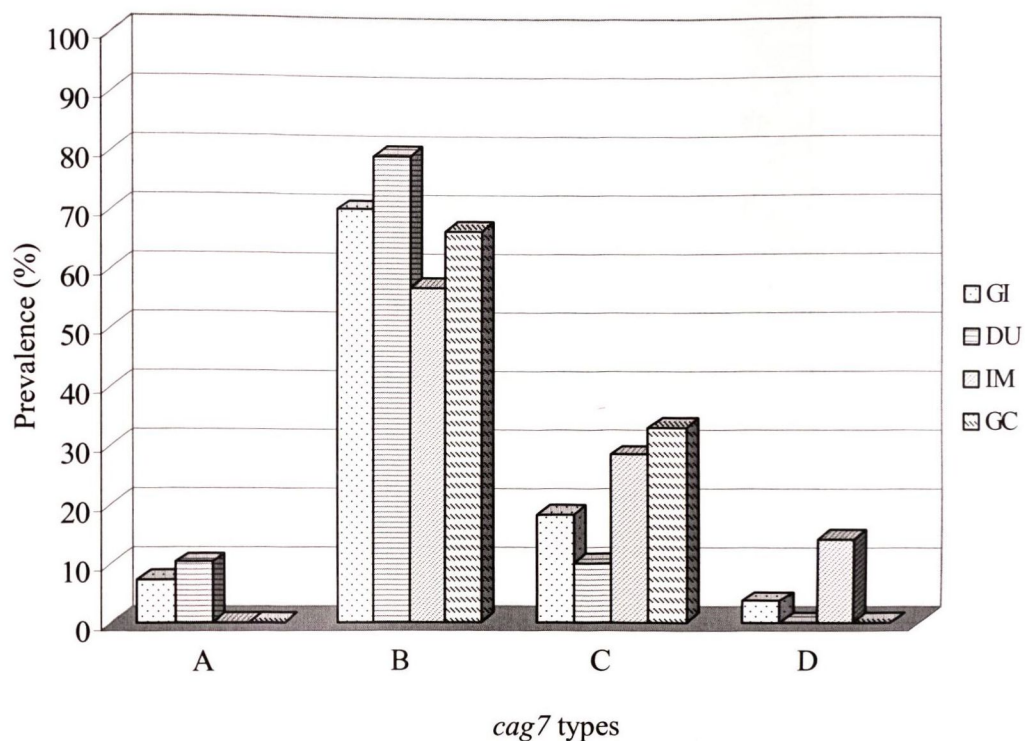


Figure 5.6. Distribution of the *cag7* genotypes in different disease groups. GI gastritis; DU duodenal ulcer; IM intestinal metaplasia; GC gastric cancer. Statistical analyses revealed that the prevalence of each genotype (*cag7* A, B, C or D) in strains from GI patients was not significantly different from that from DU, IM or GC patients (All *p* values were greater than 0.05).

Table 5.3. Distribution of the combined types of *cagA* and *cag7* in different groups.

<i>CagA</i> types	<i>cag7</i> types in GI					<i>cag7</i> types in DU					<i>cag7</i> types in IM					<i>cag7</i> types in GC				
	A	B	C	D	n	A	B	C	D	n	A	B	C	D	n	A	B	C	D	n
α_0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
β_0	0	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
γ_0	0	1	1	0	2	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
γ_1	1	8	3	0	12	1	24	3	0	28	0	3	1	1	5	0	1	1	0	2
γ_2	0	4	0	0	4	0	3	0	0	3	0	1	0	0	1	0	1	0	0	1
γ_{3+}	1	1	1	0	3	2	1	0	0	3	0	0	1	0	1	0	0	0	0	0
Total	2	15	5	0	22	4	29	4	0	37	0	4	2	1	7	0	2	1	0	3

Compared to gastritis group, the proportion of the *cag7* B-*cagA* γ_1 type in duodenal ulcer group was significantly higher ($p < 0.05$). GI gastritis; DU duodenal ulcer; IM intestinal metaplasia; GC gastric cancer.

γ_1 -*cag7* B, which occurred in 36 strains (52.2%). The combined genotype *cagA* γ_1 -*cag7* C occurred in 8 strains (11.6%). The prevalence of *cagA* γ_1 -*cag7* B in strains from the duodenal ulcer group (64.9%) was significantly higher than that in strains from the gastritis group (36.4%) ($p < 0.05$).

5.3.6. The mixed population

In some cases, the amplicons for *cagA* and *cag7* had a major band and one or more faint bands when resolved on an agarose gel. However, the size of a faint band was always equal to the size of a corresponding type of *cagA* or *cag7*. It thus appeared that the multiple binding patterns were most likely generated from a mixed population of strains. To further confirm this, each band from strain 97551 for *cagA* and strain 94536 for *cag7* was eluted from the gel, cloned and sequenced. The nucleotide sequence was identical between each band (data not shown). A mixed population was observed in 22 of 74 (29.7%) *cagA*⁺ strains and in 11 of 76 (14.5%) *cag7*⁺ strains.

5.4. Discussion

The *cagA* gene is a major virulence gene and is considered as a marker for the presence of the *cag* PAI (Censini *et al.*, 1996). The results showed that the majority of strains (88.4%) were *cagA* positive, which is consistent with previous studies on Irish strains (Marshall *et al.*, 1999; Ryan *et al.*, 2000). The prevalence of the *cagA*⁺ strains in DU was higher than that in gastritis. This fact implies that *cagA* may be associated with the development of DU (Blaser and Berg, 2001; Ryan *et al.*, 2000). The prevalence of the *cagA*⁺ strains from GC or IM groups was also higher than that from gastritis group, but the differences were not statistically significant.

5.4.1. Heterogeneity in the compositions of the *cag* PAI

Heterogeneity in the composition of the *cag* PAI in Irish strains was investigated by detection of *cagA*, *cag7* and *cag2*. The results demonstrated that the *cag* PAI in Irish strains underwent deletion of certain genes, such as *cagA* and *cag7*, indicating that *cag* PAI exhibits heterogeneous compositions among strains. It appears that there is no preferential deletion of the *cag* genes, since absence of any *cag* gene has been observed (Audibert *et al.*, 2001; Kidd *et al.*, 2001). An exception is found in Japanese strains that are all *cagA*⁺. While *cagA* is well conserved in all strains, other *cag* genes are likely to be deleted (Ikenoue *et al.*, 2001; Maeda *et al.*, 1999).

A partial *cag* PAI was present in 12.3% of the examined strains. This is in contrast to a report based on UK strains in which all strains were found to carry an intact *cag* PAI (Slater *et al.*, 1999). Other studies showed that a partial *cag* PAI was present in 4.6 to 6.3 % of strains (Audibert *et al.*, 2001; Maeda *et al.*, 1999; Slater *et al.*, 1999). The

results presented here suggest that a larger proportion of *cag*⁺ Irish strains contain the partial *cag* PAI.

5.4.2. Variation at the *cagA* 3' region

Heterogeneous size of CagA was mainly caused by variation at the *cagA* 3' region (Covacci, *et al.*, 1993; Yamaoka, *et al.*, 1998). This variation is biologically important since the tyrosine phosphorylation motifs are located here (Odenbreit *et al.*, 2000). Therefore, the variation at the *cagA* 3' region was investigated. The amplicons for this region differed by approximately 800 bp, demonstrating that this region is highly variable. Further analysis of the sequences demonstrated contraction and/or expansion of an α segment of 57 bp and a γ segment of 102 bp contributes to the size variation. Importantly, a novel sequence segment of 105 bp, namely β segment, was demonstrated to be involved in the size variation. The β segment is located between the α segment and D₃ segment. Deletion of the β segment was found in 2 of 74 (2.7%) strains. Therefore, the *cagA* 3' region can be denoted as $\alpha_n\text{-}\beta_n\text{-D}_3\text{-}\gamma_n$, in which n represents the copy number of a variable segment.

Phosphorylation of CagA occurs at three tyrosine residues (Backert *et al.*, 2001; Odenbreit *et al.*, 2000; Puls *et al.*, 2002; Stein *et al.*, 2002). The impact of *cagA* 3' region variation on these TPS sites was analysed in this study. The results demonstrated that the total numbers of the TPS sites were affected by two factors: (1) the copy numbers of the β or γ segment; and (2) nonsynonymous mutations. TPS I was absent from CagA of type β_0 due to the deletion of the β segment, and the number of TPS II sites was usually equal to the number of the duplicated γ segments. Each of the three TPS sites, however, was also affected by nonsynonymous mutations: replacement

of EPIYA at TPS I with EPIYT, replacement of EPIYA at TPS II with EPIHT in strain 94092 (type γ_3) and replacement of Y at TPS III with K or N.

Yamaoka *et al.* (1999) reported that the presence of 3 or more γ segments was associated with IM and GC in America, but not in Italy. The results of the present study showed that numbers of the repetitive segments were not associated with clinical outcomes. Therefore, genotyping of the *cagA* 3' region is not useful for predicting clinical outcomes of *H. pylori* infection in Ireland.

5.4.3. Variation at the 5' region of *cag7*

The results showed that the 5' region of *cag7* exhibited extensive size variation among clinical isolates. The variation was due to the duplication of a repetitive element which is 128 to 136 amino acid (aa) long. This is in contrast to the previous analysis of three sequenced reference strains which had a repetitive element of 130 aa long (Liu *et al.*, 1999), indicating the heterogeneous size of the element among strains. The results showed that the intra-strain duplicated elements were nearly identical, suggesting that the changes in numbers of the element originated recently. The changes may have arisen from replication strand slippage or recombination. Sequence analysis also demonstrated that in the repetitive element, at least 5 runs of doublets EE (glutamine) and 3 of doublets KK (lysine) were present, and duplication of the element increased the numbers of these doublets accordingly. The biological roles of the variation are probably involved in: (1) Regulating the expression of the GTPase homologues in the *cag* PAI (VirB4, VirB11, VirD4) since the element showed low degree of homology to a GTPase regulator; and/or (2) facilitating self-assembly and binding to other homologues of the *A. tumefaciens* Vir proteins. The *A. tumefaciens* VirB10, a homolog

of Cag7, is essential for assembling the core channel of type IV secretion machinery by interacting with VirB9, VirB8 and itself (Das and Xie, 2000). Increased numbers of the multiple runs of doublets of EE and KK would conceivably provide opportunities for multiple salt bridges contributing to protein-protein interaction (Liu *et al.*, 1999). However, the possibility of benign variation could not be excluded. Further studies such as mutagenesis analysis could elucidate the roles. In this study, *cag7* was typed into four A, B, C and D types based on the numbers of the duplicated repetitive elements. Type B was the predominant type and the distribution of the *cag7* types in the different disease groups was similar, indicating that variation does not play a significant role in the determination of clinical outcomes of *H. pylori* infection.

5.4.4. The combined genotypes of *cagA* and *cag7*

The results demonstrated that the prevalence of the *cagA* γ_1 -*cag7B* strains was higher in the duodenal ulcer group (64.9%) than in the gastritis group (36.4%). The results may have two explanations. Firstly, the *cagA* γ_1 -*cag7B* strains may have higher virulence determined by *cagA*, *cag7* or other bacterial virulence genes such as VacA (Atherton, *et al.*, 1995, Censini, *et al.*, 2001). Secondly, the *cagA* γ_1 and the *cag7B* types were co-selected during evolution, perhaps because strains with this genotype may be more suited to gastric environments (e.g. low pH) of duodenal ulcer patients. Yamaoka *et al.* (1999) reported that *cagA* γ_1 strains were more resistant to acid exposure than *cagA* γ_3 strains, which supports the second explanation.

5.4.5. The mixed populations

The variation at the *cagA* 3' region and the *cag7* 5' region share a common feature of duplication and contraction of one or more long repetitive fragments. The feature

resembles the homopolymeric tracts (multiple runs of a nucleotide) present in phase variation genes in bacteria including *H. pylori* (Alm, *et al.*, 1999; Tomb, *et al.*, 1997; Wang, *et al.*, 2000). Replication strand slippage generates multiple lineages or derivatives with different length of the tract from a single cell. These lineages have phenotypes featured with or without expression of a phase variation gene. Multiple PCR products for the *cagA* 3' region and the *cag7* 5' end from a single strain were observed. They are most likely to arise from multiple lineages originating from a single cell because their nucleotide sequences are identical. This suggests that a strain of *H. pylori* will potentially generate multiple lineages with different *cagA* 3' and/or *cag7* variants, constituting a mixed population. Intrastrain recombination or replication strand slippage in a manner resembling phase variation are probably involved in the generation of the mixed population.

5.4.6. Conclusions

The results of this study first demonstrated: (1) In addition to the α and β segments, a newly identified γ segment was involved in *cagA* 3' region variation; (2) The number of tyrosine phosphorylation sites of CagA is affected by the numbers of the three repetitive segments and nonsynonymous mutations at the *cagA* 3' region; (3) The prevalence of the *cagA* γ_1 -*cag7B* strains was higher in duodenal ulcer group (64.9%) than that in gastritis group (36.4%). This combined genotype may be used as a marker for predicting the clinical outcomes of infected individuals. The results presented here also suggest that the *cag* PAI is highly divergent among strains: heterogeneous gene compositions, varied size of the individual genes and divergent sequences. These divergences could potentially modify bacterial virulence (Blaser and Berg, 2001) or influence bacterial adaptation (Bjorkholm *et al.*, 2001 b).

CHAPTER 6
GENERAL DISCUSSION

6.1. Transcriptional response of *H. pylori* to acid exposure

Transcriptional regulation of gene expression is a major microbial strategy to cope with acid stress (Foster, 1999). The study described here investigated the transcriptional response of *H. pylori* to prolonged acid exposure using DD-PCR. A total of eight genes were found to be regulated by prolonged acid exposure. Expressions of seven genes, including *ureB*, *flaA*, *topA*, *tufB*, *topA* (a novel sequence) and two fragments which showed no significant homologues in the public databases, were up-regulated in cells grown on acidic media. The mRNA level of *atoE*, however, was decreased. Two other studies also investigated acid-responsive genes in *H. pylori*. Using microarray analysis Ang *et al.* compared the expression profiles of *H. pylori* cells grown on pH 5.5 and pH 7.2 Columbia agar plates for 48 h (2001). Allan *et al.* (2001) also used microarray to analyze the differentially expressed genes of cells of *H. pylori* grown in pH 4 and usual BHI broth.

6.1.1 Groups of acid-regulated genes

The results presented here demonstrated that *H. pylori* is likely to vary expression of groups of genes to cope with acid stress. To maintain the internal pH (pH_i) close to neutral or at an acceptable level, the bacterium requires up-regulation of urease expression. This was supported by the observation in this study that *ureB* expression was increased in response to acid exposure, and by another report that production of UreB was increased in cells exposed to acid (Slonczewski *et al.*, 2000). The observed down-regulation of *atoE* at low external pH (pH_o) suggests that the bacterium could decrease uptake of weak acid to struggle for a neutral pH_i . Genes responsible for ion flux across cytoplasmic membranes may be involved in maintenance of pH_i homeostasis (Stingl *et al.*, 2002 a). Transcriptional regulation of these genes plays an

important role in acid resistance of enterobacteria (Foster, 1999). The results showed that none of these genes were differentially expressed in acid exposure, consistent with the results from other studies (Allan, *et al.*, 2001; Ang, *et al.*, 2001). Therefore, it appears that regulation of expression of genes for ion flux is not required by the bacterium to maintain internal pH.

At low external pH, cytoplasmic acidification (drops to pH 4.9, Stingl *et al.*, 2002a) and urea influx (Weeks *et al.*, 2000) occur in *H. pylori*. Acid exposure can directly damage cellular surface structures. The results of this study demonstrated *H. pylori* regulates sets of genes to counteract damage caused by these acid-related factors. To maintain DNA stability and conformation at low external pH, the bacterium may alter the expression of *topA* (observed in this study) and genes involved in DNA repair (Ang, *et al.*, 2001). In order to maintain correct folding, assembly, and maturation of newly synthesised proteins at acidic external pH, the bacterium may regulate the expression of genes coding for chaperon proteins. These genes include *tufB* (this study) and heat shock genes (Allan, *et al.*, 2001; Slonczewski *et al.*, 2000). To protect the surface structures from direct acid damage, the bacterium may vary expression of genes involved in the synthesis of these structures. This group of genes is comprised of *flaA* (this study), LPS synthesis genes, and genes coding for outer membrane proteins (Allan *et al.*, 2001; Ang *et al.*, 2001; McGowan *et al.*, 1998; Moran *et al.*, 2002).

The results of this study showed that some acid-regulated genes lacked apparent evidence for involvement in acid resistance, such as *pap57*. This gene is not present in the two sequenced strains, indicating it is a gene unique to certain *H. pylori* strains.

Groups of such genes were also found by other studies (Ang *et al.*, 2001; Allan *et al.*, 2001). Further investigations on their implications in acid adaptation are required.

6.1.2. Future studies on acid adaptation

The identified acid-regulated genes provide useful information for our further understanding of the mechanisms of acid resistance. Several issues, however, are still unclear. (1) Which signals are sensed by *H. pylori* to modulate the expressions of acid-responsive genes? It could be considered that external pH is sensed by UreI (Weeks *et al.*, 2001). However, there is no evidence that UreI regulates the acid-responsive genes. Cytoplasmic acidification and urea might be candidate signals sensed by the bacterium. (2) Does *H. pylori* possess one or more master-like regulatory genes? If so, what are these genes? In *Salmonella typhimurium*, several regulatory genes play a role in log phase acid tolerance response including *rpoS*, *fur* (iron regulator) and *phoPQ* (two-component signal transduction system) (Foster, 1999). So far, there is no evidence that the major regulatory genes in *H. pylori*, including two-component systems and quorum-sensing genes, are involved in acid resistance. The promoters of several acid-regulated genes are likely to be sensitive to the alterations in local DNA conformation. Therefore, it would be of interest to know if *topA*, a differential expression gene identified in this study could function in regulation of acid-responsive genes. (3) What roles do these acid-regulated genes play in acid resistance? Excluding urease genes, few genes have been characterized and, therefore, further studies are necessary. (4) Only a few identified acid-regulated genes from one study have been confirmed by others. This may be attributed to different pH values, varied exposure times, different compositions of media and different techniques used. These factors should be taken into account in future studies.

6.2. Generation of genetic divergence in *H. pylori*

6.2.1. Features of genetic diversity

In this study, inter-strain genetic diversities were investigated at a genome scale and at a large horizontal transferred DNA region, the *cag* pathogenicity island, which is a major virulence determinant (Blaser and Berg, 2001). The results demonstrated that *H. pylori* exhibits several types of genetic diversity. (1) Strain-specific genes: These strain-specific genes included two novel sequences (clone F5 and P32) and homologues of *resI* and JHP 704; (2) Chromosomal rearrangements: three rearrangement loci including a locus of the “plasticity zone” randomly occurred in the clinical isolates; (3) Allelic polymorphism including size variation of *cagA* and *cag7* and divergent sequences between alleles (*cagA*, *cag7* and *mobA*); (4) Instability of the *cag* pathogenicity island. Deletion of *cagA* and/or *cag7* was observed. Combined, these types of diversities build up a highly divergent *H. pylori* genome. The status of the genotypes of nine genes/sequences, representing these identified types of genetic diversity, were determined in 19 strains related to duodenal ulcer diseases, intestinal metaplasia or gastric cancer. Each strain exhibited a unique distribution profile of these genes/sequences, supporting the conclusion.

6.2.2. Generation of genetic divergence in *H. pylori*

Generation of these identified types of genetic diversities may involve a number of mechanisms. (1) *Mutation*. *H. pylori* shows a higher mutation rate than *E. coli* (Bjorkholm *et al.*, 2001a). Approximately a quarter of strains possess a mutator-like phenotype, which usually results from deficiency in MutS in *E. coli*. The high mutation rate in *H. pylori* may be attributed to lack of a functional DNA repair system

(Bjorkholm *et al.*, 2001a; Wang, *et al.*, 1999). The observed allelic polymorphisms in this study reflect the high mutation rate in *H. pylori*. Notably, nonsynonymous mutation occurred in *cagA* and *cag7* which possibly affected their activities. (2) *Recombination*. Recombination may also be responsible for the observed allelic polymorphisms. Recombination in *H. pylori* is more frequent than in any other organism studied to date (Achtman and Suerbaum, 1998). (3) *Slipped-strand replication*. Variation in the *cag7* and *cagA* sizes involved single or multiple long repetitive elements. The nearly identical sequences between duplicated elements suggest a recent origination of the duplication. Strand-slippage during replication is a major cause of size variation. Slipped-strand replication may generate multiple lineages from an individual cell (parent cell), which manifest as a mixed population observed in this study. (4) *Horizontal gene transfer*. Transformation, transduction and conjugation mediate the horizontal transfer of genes. *H. pylori* is naturally competent for transformation. Genes with a possible origin from plasmids have been demonstrated in the bacterium (Alm, *et al.*, 1999; Garcia-Vallve, *et al.*, 2002). In *H. pylori*, a type IV secretion machinery resembling a conjugative apparatus plays a major role in uptake of foreign DNA. A GepA homologue (Clone P32) was identified in this study, suggesting a putative conjugative gene transfer mechanism may exist in *H. pylori*. Strain-specific genes identified in this study are possibly acquired through horizontal gene transfer due to their lower G+C contents. (5) *Reduction of the genome*. *H. pylori* acquires genes horizontally *in vivo* demonstrating a continuous state of gene influx occurs in the bacterium (Israel, *et al.*, 2001). A mechanism of gene efflux (reduction of the genome), however, is required to maintain the appropriate genomic size of the organism (Garcia-Vallvé *et al.*, 2002). This is exemplified by the genetic diversity of the *cag* PAI. Elimination of the *cag* PAI from the genome has been demonstrated to result from

recombination between a *cag* positive strain and a *cag* negative strain (Kersulyte *et al.*, 1999). Inactivation of a *cag* gene (HP0532) is an example of gene degradation during reduction (Fischer *et al.*, 2001). Absence of *cagA* and/or *cag7* from certain examined strains in this study is another manifestation of reduction of the genome. Lower frequency of clone P32 in strains from intestinal metaplasia patients than from duodenal ulcer may indicate the elimination of this gene by genome reduction.

6.3. Implications of genome divergence

Transmission to a new host imposes a large selection force on *H. pylori*. Rapidly changing and fluctuating environments, detrimental factors in ingestion, and periodic gastric acid secretion challenge the bacterium frequently. This requires the bacterium to be fully prepared for these challenges. The bacterium persists in its host for decades. The gastric mucosa may undergo significant changes during the inhabitation and thus, pose a threat to the bacterium. The aforementioned mechanisms allow the bacterium to diversify its genome, vary the phenotype traits and therefore to efficiently adapt to the hostile environment. The bacterium acquires favourable traits through horizontal gene transfer (Ochman and Moran, 2001), improves competitiveness through elimination of neutral or detrimental genes (Mira, *et al.*, 2001), and evades the host defence system through allelic polymorphism generated by mutation, recombination or strand-slippage replication (Slev and Potts, 2002).

This study demonstrated that individual genetic loci (for example, strain-specific genes) or allele variants were only found in a fraction of strains isolated from unrelated individual hosts. These genetic traits are possibly selected if they are beneficial for the bacterium to inhabit particular environments. The selection does not “consider”

whether it is beneficial, neutral or harmful to the hosts (Ochman and Moran, 2001). Severe consequences will occur if a virulent genetic trait is selected. In agreement with this, this study observed higher prevalence of *cagA*⁺ strains or *cag7* B-*cagA* γ_1 strains in duodenal ulcer compared to gastritis. Individuals infected by strains with these genetic traits may have increased risk for development of duodenal ulcer disease.

BIBLIOGRAPHY

Achtman M. and Suerbaum S. (2000) Sequence variation in *Helicobacter pylori*. Trends in Microbiology 8: 57-58.

Akada J.K., Shirai M., Takeuchi H., Tsuda M. and Nakazawa T. (2000) Identification of the urease operon in *Helicobacter pylori* and its control by mRNA decay in response to pH. Molecular Microbiology 36: 1071-1084.

Akopyants N.S., Fradkov A., Diatchenko L., Hill J.E., Siebert P.D., Lukyanov S.A., Sverdlov E.D. and Berg D.E. (1998) PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. Proceedings of the National Academy of Sciences of the United States of America 95: 13108-13113.

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

Akopyants N.S., Eaton K.A. and Berg D.E. (1995) Adaptive mutation and cocolonization during *Helicobacter pylori* infection of gnotobiotic piglets. Infection and Immunity 63: 116-121.

Akopyants N., Bukanov, N., Westblom, T., Kresovich, S. and D. Berg. (1992) DNA diversity among clinical isolates of *Helicobacter pylori* detected by PCR-based RAPD fingerprinting. Nucleic Acids Research (Online) 20: 5137-5142.

Allan E., Clayton C.L., McLaren A., Wallace D.M. and Wren B.W. (2001) Characterization of the low-pH responses of *Helicobacter pylori* using genomic DNA arrays. *Microbiology* 147: 2285-2292.

Alm R.A., Ling L.S., Moir D.T., King B.L., Brown E.D., Doig P.C., Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G., Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C., Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F. and Trust T.J. (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397: 176-180.

Andersen A.P., Elliott, D.A., Lawson, M., Barland, P., Hatcher, V.B. and Puszkun, E.G. (1997) Growth and Morphological transformations of *Helicobacter pylori* in broth media. *Journal of Clinical Microbiology* 35: 2918-2922.

Ando T., Xu Q., Torres M., Kusugami K., Israel D.A. and Blaser M.J. (2000) Restriction-modification system differences in *Helicobacter pylori* are a barrier to interstrain plasmid transfer. *Molecular Microbiology* 37: 1052-1065.

Andrutis K.A., Fox J.G., Schauer D.B., Marini R.P., Li X., Yan L., Josenhans C. and Suerbaum S. (1997) Infection of the ferret stomach by isogenic flagellar mutant strains of *Helicobacter mustelae*. *Infection and Immunity* 65: 1962-1966.

Ang S., Lee C.Z., Peck K., Sindici M., Matrubutham U., Gleeson M.A. and Wang J.T. (2001) Acid-induced gene expression in *Helicobacter pylori*: study in genomic scale by microarray. *Infection and Immunity* 69: 1679-1686.

Amarri S. and Celinska-Cedro D. (2001) Paediatrics. *Helicobacter* 17 (suppl 1): S32-S37.

Ansorg R., Von-Recklinghausen G., Pomarius R. and Schmid E.N. (1991) Evaluation of techniques for isolation, subcultivation, and preservation of *Helicobacter pylori*. *Journal of Clinical Microbiology* 29: 51-53.

Appelmelk B.J., Monteiro M.A., Martin S.L., Moran A.P. and Vandenbroucke-Grauls C.M. (2000) Why *Helicobacter pylori* has Lewis antigens. *Trends in Microbiology* 8: 565-570.

Appelmelk B.J., Simoons-Smit I., Negrini R., Moran A.P., Aspinall G.O., Forte J.G., De-Vries T., Quan H., Verboom T., Maaskant J.J., Ghiara P., Kuipers E.J., Bloemena E., Tadema T.M., Townsend R.R., Tyagarajan K., Crothers J.M., Monteiro M.A., Savio A. and De-Graaff J. (1996) Potential role of molecular mimicry between *Helicobacter pylori* lipopolysaccharide and host Lewis blood group antigens in autoimmunity. *Infection and Immunity* 64: 2031-2040.

Aras R.A., Takata T., Ando T., van der Ende A. and Blaser M.J. (2001) Regulation of the HpyII restriction-modification system of *Helicobacter pylori* by gene deletion and horizontal reconstitution. *Molecular Microbiology* 42: 369-382.

Arents N.L., van-Zwet A.A., Thijs J.C., Kooistra-Smid A.M., van-Slochteren K.R., Degener J.E., Kleibeuker J.H., and Van-Doorn L.J. (2001) The importance of *vacA*, *cagA*, and *iceA* genotypes of *Helicobacter pylori* infection in peptic ulcer disease and gastroesophageal reflux disease. The American Journal of Gastroenterology 96: 2603-2608.

Asaka M., Sugiyama T., Nobuta A., Kato M., Takeda H. and Graham D.Y. (2001) Atrophic gastritis and intestinal metaplasia in Japan: results of a large multicenter study. *Helicobacter* 6: 294-299.

Atherton J.C., Cao P., Peek R.M., Tummuru M.K., Blaser M.J. and Cover T.L. (1995) Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. The Journal of Biological Chemistry 270: 17771-17777.

Audibert C., Burucoa C., Janvier B. and Fauchere J.L. (2001) Implication of the structure of the *Helicobacter pylori* *cag* pathogenicity island in induction of interleukin-8 secretion. Infection and Immunity 69: 1625-1629.

Austin J.W., Doig P., Stewart M. and Trust T.J. (1991) Macromolecular structure and aggregation states of *Helicobacter pylori* urease. Journal of Bacteriology 173: 5663-5667.

Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Smith J.A., Stedman J.G. and Struhl K. (1987) Current protocol in Molecular Biology. New York: Greene Publishing Associates Inc. and J Wiley and Sons Inc.

Axon A. (1995) Is *Helicobacter pylori* transmitted by the gastro-oral route? Alimentary Pharmacology & Therapeutics 9: 585-589.

Backert S., Moese S., Selbach M., Brinkmann V. and Meyer T.F. (2001) Phosphorylation of tyrosine 972 of the *Helicobacter pylori* CagA protein is essential for induction of a scattering phenotype in gastric epithelial cells. Molecular Microbiology 42: 631-644.

Backert S., Muller E.C., Jungblut P.R. and Meyer T.F. (2001) Tyrosine phosphorylation patterns and size modification of the *Helicobacter pylori* CagA protein after translocation into gastric epithelial cells. Proteomics 1: 608-617.

Backert S., Ziska E., Brinkmann V., Zimny-Arndt U., Fauconnier A., Jungblut P.R., Naumann M. and Meyer T.F. (2000) Translocation of the *Helicobacter pylori* CagA protein in gastric epithelial cells by a type IV secretion apparatus. Cellular Microbiology 2: 155-164.

Banatvala N., Mayo K., Megraud F., Jennings R., Deeks J.J. and Feldman R.A. (1993) The cohort effect and *Helicobacter pylori*. Journal of Infectious diseases 168: 219-221.

Baron C., O'Callaghan D. and Lanka E. (2002) Bacterial secrets of secretion: EuroConference on the biology of type IV secretion processes. *Molecular Microbiology* 43: 1359-1365.

Bauerfeind P., Garner, R., Dunn, B.E. and Mobley, H.L. (1997) Synthesis and activity of *Helicobacter pylori* urease and catalase at low pH. *Gut* 40: 25-30.

Bearson S., Bearson B. and Foster J.W. (1997) Acid stress responses in *enterobacteria*. *FEMS Microbiology Letters* 147: 173-180.

Beier D. and Frank R. (2000) Molecular characterization of two-component systems of *Helicobacter pylori*. *Journal of Bacteriology* 182: 2068-2076.

Beier D., Spohn G., Rappuoli R., and Scarlato V. (1998) Functional analysis of the *Helicobacter pylori* principal sigma subunit of RNA polymerase reveals that the spacer region is important for efficient transcription. *Molecular Microbiology* 30: 121-134.

Bijlsma J.J., Lie A., Ling M., Nootenboom I.C., Vandenbroucke Grauls C.M. and Kusters J.G. (2000) Identification of loci essential for the growth of *Helicobacter pylori* under acidic conditions. *The Journal of Infectious Diseases* 182: 1566-1569

Bijlsma J.J., Gerrits M.M., Imamdi R., Vandenbroucke-Grauls C.M. and Kusters J.G. (1998) Urease-positive, acid-sensitive mutants of *Helicobacter pylori*: urease-independent acid resistance involved in growth at low pH. *FEMS Microbiology Letters* 167: 309-313.

Bjorkholm B.M., Guruge, J.L., Oh, J.D., Syder, A.J., Salama, N., Guillemin, K., Falkow, S., Nilsson, C, Falk, P.G., Engstrand, L. and Gordon, J.I. (2002) Colonization of Germ-free Transgenic Mice with Genotyped *Helicobacter pylori* Strains from a Case-Control Study of Gastric Cancer Reveals a Correlation between Host Responses and HsdS Components of Type I Restriction-Modification Systems. The Journal of Biological Chemistry: 34191-34197.

Bjorkholm B., Lundin A., Sillen A., Guillemin K., Salama N., Rubio C., Gordon J.I., Falk P. and Engstrand L. (2001a) Comparison of genetic divergence and fitness between two subclones of *Helicobacter pylori*. Infection and Immunity 69: 7832-7838.

Bjorkholm B., Sjolund M., Falk P.G., Berg O.G., Engstrand L. and Andersson D.I. (2001b) Mutation frequency and biological cost of antibiotic resistance in *Helicobacter pylori*. Proceedings of the National Academy of Sciences of the United States of America 98: 14607-14612.

Blaser M.J. (1995) The role of *Helicobacter pylori* in gastritis and its progression to peptic ulcer disease. Alimentary Pharmacology & Therapeutics 9 (Suppl 1): 27-30.

Blaser M.J. and Berg D.E. (2001) *Helicobacter pylori* genetic diversity and risk of human disease. The Journal of Clinical Investigation 107: 767-773.

Blattner F.R., Plunkett G., Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick

H.A., Goeden M.A., Rose D.J., Mau B. and Shao Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277: 1453-1474.

Bloomfield G.A., Whittle, G., McDonagh, M.B., Katz, M.E. and Cheetham, B.F. (1997) Analysis of sequences flanking the *vap* regions of *Dichelobacter nodosus*: evidence for multiple integration events, a killer system, and a new genetic element. *Microbiology* 143: 553-562.

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

Boucher Y., Nesbo C.L. and Doolittle W.F. (2001) Microbial genomes: dealing with diversity. *Current Opinion in Microbiology* 4: 285-289.

Breslin N.P., Lee J.M., Buckley M.J., Balbirnie E., Rice D. and O'Morain C.A. (2000) Validation of serological tests for *Helicobacter pylori* infection in an Irish population. *Irish Journal of Medical Science* 169: 190-194.

Buckley M.J and O'Morain C.A. (1998) *Helicobacter* biology-discovery. *British Medical Bulletin* 54: 7-16.

Bukholm G., Tannaes T., Nedenskov P., Esbensen Y., Grav H.J., Hovig T., Ariansen S. and Guldvog I. (1997) Colony variation of *Helicobacter pylori*: pathogenic potential is

correlated to cell wall lipid composition. *Scandinavian Journal of Gastroenterology* 32: 445-454.

Burgess R.R. and Antony L. (2001) How sigma docks to RNA polymerase and what sigma does. *Current Opinion in Microbiology* 4: 126-131.

Calam J. (1996) *Clinician's Guide to Helicobacter pylori*. London : Chapman & Hall Medical.

Caldas T.D., Yaagoubis, A.E. and Richarme, G. (1998) Chaperon properties of bacterial elongation factor EF-Tu. *The Journal of Biological Chemistry* 273: 11478-11482.

Cao P. and Cover T.L. (1997) High-level genetic diversity in the *vapD* chromosomal region of *Helicobacter pylori*. *Journal of Bacteriology* 179: 2852-2856.

Censini S., Lange C., Xiang Z., Crabtree J.E., Ghiara P., Borodovsky M., Rappuoli R. and Covacci A. (1996) *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proceedings of the National Academy of Sciences of the United States of America* 93: 14648-14653.

Censini S., Stein M. and Covacci A. (2001) Cellular responses induced after contact with *Helicobacter pylori*. *Current Opinion in Microbiology* 4: 41-46.

Chang K.C., Yeh Y.C., Lin T.L. and Wang J.T. (2001) Identification of genes associated with natural competence in *Helicobacter pylori* by transposon shuttle random mutagenesis. *Biochemical and Biophysical Research Communications* 288: 961-968.

Chey W.D., Murthy U., Shaw S., Zawadski A., Montague J., Linscheer W. and Laine L. (1999) A comparison of three fingerstick, whole blood antibody tests for *Helicobacter pylori* infection: a United States, multicenter trial. *The American Journal of Gastroenterology* 94: 1512-1516.

Chey W.D., Woods M. and Scheiman J.M. (1997) Lansoprazole and ranitidine affect the accuracy of the ¹⁴C-urea breath test by a pH-dependent mechanism. *American Journal of Gastroenterology* 92: 446-450.

Chow W.H., Blaser M.J., Blot W.J., Gammon M.D., Vaughan T.L., Risch H.A., Perez-Perez G.I., Schoenberg J.B., Stanford J.L., Rotterdam H., West A.B. and Fraumeni J.F. (1998) An inverse relation between *cagA*⁺ strains of *Helicobacter pylori* infection and risk of esophageal and gastric cardia adenocarcinoma. *Cancer Research* 58: 588-590.

Colland F., Rain J.C., Gounon P., Labigne A., Legrain P. and De Reuse H. (2001) Identification of the *Helicobacter pylori* anti-sigma28 factor. *Molecular Microbiology* 41: 477-487.

Colonna-Romano S., Porta A., Franco A., Kobayashi G.S. and Maresca B. (1998) Identification and isolation by DDRT-PCR of genes differentially expressed by

Histoplasma capsulatum during macrophages infection. Microbial Pathology 25: 55-66.

Correa P., Fontham E.T., Bravo J.C., Bravo L.E., Ruiz B., Zarama G., Realpe J.L., Malcom G.T., Li D., Johnson W.D. and Mera R. (2000) Chemoprevention of gastric dysplasia: randomized trial of antioxidant supplements and anti-*Helicobacter pylori* therapy. Journal of the National Cancer Institute 92: 1881-1888.

Covacci A., Censini S., Bugnoli M., Petracca R., Burroni D., Macchia G., Massone A., Papini E., Xiang Z. and Figura N. (1993) Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. Proceedings of the National Academy of Sciences of the United States of America 90: 5791-5795.

Cover T.L. and Blaser M.J. (1992) *Helicobacter pylori* and gastroduodenal disease. Annual Review of Medicine 43: 135-145.

Czajkowsky D.M., Iwamoto H., Cover T.L. and Shao Z. (1999) The vacuolating toxin from *Helicobacter pylori* forms hexameric pores in lipid bilayers at low pH. Proceedings of the National Academy of Sciences of the United States of America 96: 2001-2006.

Das A. and Xie Y. (2000) The *Agrobacterium* T-DNA transport pore proteins VirB8, VirB9 and VirB10 interact with one another. Journal of Bacteriology 182: 758-763.

Debets-Ossenkopp Y.J., Brinkman A.B., Kuipers E.J., Vandenbroucke-Grauls C.M. and Kusters J.G. (1998) Explaining the bias in the 23S rRNA gene mutations associated with clarithromycin resistance in clinical isolates of *Helicobacter pylori*. *Antimicrobial Agents and Chemotherapy* 42: 2749-2751.

Dietz P., Gerlach G. and Beier D. (2002) Identification of target genes regulated by the two-component system HP166-HP165 of *Helicobacter pylori*. *Journal of Bacteriology* 184: 350-362.

Dobrindt U. and Hanker J. (2001) Whole genome plasticity in pathogenic bacteria. *Current Opinion in Microbiology* 4: 550-557.

Doig P., de-Jonge B.L., Alm R.A., Brown E.D., Uria Nickelsen M., Noonan B., Mills S.D., Tummino P., Carmel G., Guild B.C., Moir D.T., Vovis G.F. and Trust T.J. (1999) *Helicobacter pylori* physiology predicted from genomic comparison of two strains. *Microbiology and Molecular Biology Reviews* 63: 675-707.

Donahue J.P., Israel D.A., Peek R.M., Blaser M.J. and Miller G.G. (2000) Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori*. *Molecular Microbiology* 37: 1066-1074.

Dubois A., Fiala N., Heman-Ackah L.M., Drazek E.S., Tarnawski A., Fishbein W.N., Perez-Perez G.I. and Blaser M.J. (1994) Natural gastric infection with *Helicobacter*

pylori in monkeys: a model for spiral bacteria infection in humans. *Gastroenterology* 106: 1405-1417.

Dundon W., Nishioka H., Montecucco C., Montemurro P., Rappuoli R. and Del Giudice G. (2001) Problems assaying neutrophil activators. *Trends in Microbiology* 9: 314-315.

Dunn B.E., Cohen, H. and Blaser M.J. (1997) *Helicobacter pylori*. *Clinical Microbiology Reviews* 10: 720-741.

Dwyer B.K.J., Tee W., Marakowski E. and Raios K. (1988) Antibody response to *Campylobacter pylori* in diverse ethnic groups. *Scandinavian Journal of Infectious Diseases* 20: 349-350.

Eaton K.A., Gilbert J.V., Joyce E.A., Wanken A.E., Thevenot T., Baker P., Plaut A. and Wright A. (2002) *In vivo* complementation of *ureB* restores the ability of *Helicobacter pylori* to colonize. *Infection and Immunity* 70: 771-778.

Eaton K.A., Suerbaum S., Josenhans C. and Krakowka S. (1996) Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. *Infection and Immunity* 64: 2445-2448.

Eaton K.A., Morgan D.R. and Krakowka S. (1992) Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori*. *Journal of Medical Microbiology* 37: 123-127.

Eaton K.A., Brooks C.L., Morgan D.R. and Krakowka S. (1991) Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infection and Immunity* 59: 2470-2475.

Eaton K.A., Morgan D.R. and Krakowka S. (1989) *Campylobacter pylori* virulence factors in gnotobiotic piglets. *Infection and Immunity* 57: 1119-1125.

Elitsur Y S.J. and Neace C. (1998) Prevalence of *Helicobacter pylori* infection in children from urban and rural West Virginia. *Digestive Diseases and Sciences* 43: 773-778.

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

Evans D.J. and Evans D.G. (2001) *Helicobacter pylori* CagA: analysis of sequence diversity in relation to phosphorylation motifs and implications for the role of CagA as a virulence factor. *Helicobacter* 6: 187-198.

Evans D.J., Queiroz D.M., Mendes E.N. and Evans D.G. (1998) Diversity in the variable region of *Helicobacter pylori* *cagA* gene involves more than simple repetition of a 102-nucleotide sequence. *Biochemical and Biophysical Research Communications* 245: 780-784.

Evans D.J., Evans D.G., Takemura T., Nakano H., Lampert H.C., Graham D.Y., Granger D.N. and Kviety P.R. (1995) Characterization of a *Helicobacter pylori* neutrophil-activating protein. *Infection and Immunity* 63: 2213-2220.

Evans D.J., Evans, D.G., Kirkpatrick, S.S. and Graham, D.Y. (1991) Characterization of the *Helicobacter pylori* urease and purification of its subunits. *Microbial Pathogenesis* 10: 15-26.

Fabret C., Feher V. and Hoch J.A. (1999) Two-component signal transduction in *Bacillus subtilis*: How one organism sees its world. *Journal of Bacteriology* 181: 1975-1983.

Falk P., Roth K.A., Boren T., Westblom T.U., Gordon J.I. and Normark S. (1993) An in vitro adherence assay reveals that *Helicobacter pylori* exhibits cell lineage-specific tropism in the human gastric epithelium. *Proceedings of the National Academy of Sciences of the United States of America* 90: 2035-2039.

Falke J.J., Bass R.B., Butler S.L., Chervitz S.A. and Danielson M.A. (1997) The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annual Review of Cell and Developmental Biology* 13: 457-512.

Fallone C.A., Loo V.G., Lough J. and Barkun A.N. (1997) Hematoxylin and eosin staining of gastric tissue for the detection of *Helicobacter pylori*. *Helicobacter* 2: 32-35.

Falush D., Kraft C., Taylor N.S., Correa P., Fox J.G., Achtman M. and Suerbaum S. (2001) Recombination and mutation during long-term gastric colonization by *Helicobacter pylori*: estimates of clock rates, recombination size, and minimal age. *Proceedings of the National Academy of Sciences of the United States of America* 98: 15056-15061.

Fiedorek SC, Evans DL. (1991) Factors influencing the epidemiology of *Helicobacter pylori* infection in children. *Paediatrics* 88: 578-582.

Figueiredo C., Quint W.G., Sanna R., Sablon E., Donahue J.P., Xu Q., Miller G.G., Peek R.M., Blaser M.J. and van Doorn L.J. (2000) Genetic organization and heterogeneity of the *iceA* locus of *Helicobacter pylori*. *Gene* 246: 59-68.

Fischer W., Puls J., Buhrdorf R., Gebert B., Odenbreit S. and Haas R. (2001) Systematic mutagenesis of the *Helicobacter pylori* *cag* pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. *Molecular Microbiology* 42: 1337-1348.

Fislage R., Berceanu, M., Humboldt, Y., Wendt, M. and Oberenfer, H. (1997) Primer design for a prokaryotic differential display RT-PCR. *Nucleic Acids Research* 25: 1830-1835.

Fitzgerald J.R. and Musser J.M. (2001) Evolutionary genomics of pathogenic bacteria. *Trends in Microbiology* 9: 547-553.

Forbes K.J., Fang Z. and Pennington T.H. (1995) Allelic variation in the *Helicobacter pylori* flagellin genes *flaA* and *flaB*: its consequences for strain typing schemes and population structure. *Epidemiology and Infection* 114: 257-266.

Forsyth M.H. and Cover T.L. (2000) Intercellular communication in *Helicobacter pylori*: *luxS* is essential for the production of an extracellular signalling molecule. *Infection and Immunity* 68: 3193-3199.

Forsyth M.H. and Cover T.L. (1999) Mutational analysis of the *vacA* promoter provides insight into gene transcription in *Helicobacter pylori*. *Journal of Bacteriology* 181: 2261-2266.

Forsyth M.H., Atherton J.C., Blaser M.J. and Cover T.L. (1998) Heterogeneity in levels of vacuolating cytotoxin gene (*vacA*) transcription among *Helicobacter pylori* strains. *Infection and Immunity* 66: 3088-3094.

Foster J.W. (1999) When protons attack: microbial strategies of acid adaptation. *Current Opinion in Microbiology* 2: 170-174.

Fox J.G. (2002) The non-*Helicobacter pylori* *Helicobacters*: their expanding role in gastrointestinal and systemic diseases. *Gut* 50: 273-283.

Fox J., Beck, P., Dangler, C., Whary, M., Wang, T., Shi, H. and Nagler-Anderson, C. (2000) Concurrent enteric helminth infection modulates inflammation and gastric immune responses and reduces *Helicobacter*-induced gastric atrophy. *Nature Medicine* 6: 536-542.

Fox J.G., Batchelder M., Marini R., Yan L., Handt L., Li X., Shames B., Hayward A., Campbell J. and Murphy J.C. (1995) *Helicobacter pylori*-induced gastritis in the domestic cat. *Infection and Immunity* 63: 2674-2681.

Foyes S., Dorrell N., Ward S.J., Stabler R.A., McColm A.A., Rycroft A.N. and Wren, B.W. (2000) *Helicobacter pylori* possesses two *cheY* response regulators and a histidine kinase sensor, *cheA*, which are essential for chemotaxis and colonisation of the gastric mucosa. *Infection and Immunity* 68: 2016-2023.

Fresnadillo-Martinez M.J., Rodriguez-Rincon M., Blazquez-de-Castro A.M., Garcia-Sanchez E., Garcia-Sanchez J.E., Trujillano-Martin I., Cordero-Sanchez M., Alvarez-Alvarez P., Paz-Bouza J. and Garcia-Rodriguez J.A. (1997) Comparative evaluation of selective and nonselective media for primary isolation of *Helicobacter pylori* from gastric biopsies. *Helicobacter* 2: 36-39.

Fujinaga R., Nakazawa T. and Shirai M. (2001) Allelic exchange mutagenesis of *rpoN* encoding RNA-polymerase sigma54 subunit in *Helicobacter pylori*. 7: 148-155.

Fung W.P., Papadimitrou, J.M. and Matz, L.R. (1979) Endoscopic, histological and ultrastructural correlations in chronic gastritis. The American Journal of Gastroenterology 71: 269-279.

Furuta T., Shirai N., Takashima M., Xiao F., Hanai H., Nakagawa K., Sugimura H., Ohashi K. and Ishizaki T. (2001) Effects of genotypic differences in CYP2C19 status on cure rates for *Helicobacter pylori* infection by dual therapy with rabeprazole plus amoxicillin. Pharmacogenetics 11: 341-348.

Garcia-Vallve S., Janssen P. and Ouzounis C. (2002) Genetic variation between *Helicobacter pylori* strains: gene acquisition or loss? Trends in Microbiology. 10: 1016-1018.

Gerhard M., Lehn N., Neumayer N., Boren T., Rad R., Schepp W., Miehke S., Classen M., and Prinz C. (1999) Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. Proceedings of the National Academy of Sciences of the United States of America 96: 12778-12783.

Gibson J.R., Slater E., Xerry, J., Tompkins D. S. and Owen R. J. (1998) Use of an amplified-fragment length polymorphism technique to fingerprint and differentiate isolates of *Helicobacter pylori*. Journal of Clinical Microbiology 36: 2580-2585.

Giraud A.M., Tenaillon O., Clara A., Radman M., Fons M. and Taddei F. (2001) Costs and benefits of high mutation rates: adaptive evolution of bacteria in mouse gut. Science 291: 2606-2608.

Goodwin C.S., Armstrong, J.A. (1990) Microbiological aspects of *Helicobacter pylori* (*Campylobacter pylori*). European Journal of Clinical Microbiology and Infectious Diseases 9: 1-13.

Goodwin C.S., Armstrong J.A. and Chilvers T. (1989) Transfer of *Campylobacter pylori* and *Campylobacter mustelae* to *Helicobacter* genus nov. as *Helicobacter pylori* comb. nov. and *Helicobacter mustelae* comb. nov. , respectively. International Journal of Systematic Bacteriology 39: 397-405.

Graham D. (1997) *Helicobacter pylori* infection in the pathogenesis of duodenal ulcer and gastric cancer: a model. Gastroenterology 113: 1983-1991.

Guruge J.L., Falk P.G., Lorenz R.G., Dans M., Wirth H.P., Blaser M.J., Berg D.E. and Gordon J.I. (1998) Epithelial attachment alters the outcome of *Helicobacter pylori* infection. Proceedings of the National Academy of Sciences of the United States of America 95: 3925-3930.

Ha N.C., Oh S.T., Sung J.Y., Cha K.A., Lee M.H. and Oh B.H. (2001) Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. Nature Structural Biology 8: 505-509.

Hampson I.N., Hampson L. and Dexter T.M. (1996) Directional random oligonucleotide primed (DROP) global amplification of cDNA: its application to subtractive cDNA cloning. Nucleic Acids Research 24: 4832-4835.

Han S.R., Zschausch H.C., Meyer H.G., Schneider T., Loos M., Bhakdi S. and Maeurer M.J. (2000) *Helicobacter pylori*: clonal population structure and restricted transmission within families revealed by molecular typing. *Journal of Clinical Microbiology* 38: 3646-3651.

Han S.W., Flamm R., Hachem C.Y., Kim H.Y., Clarridge J.E., Evans D.G., Beyer J., Drnec J. and Graham D.Y. (1995) Transport and storage of *Helicobacter pylori* from gastric mucosal biopsies and clinical isolates. *European Journal of Clinical Microbiology & Infectious Diseases* 14: 349-352.

Handt L.K., Klein, H.J., Rozmiarek H., Fox J.G., Pouch W.J. and Motzel S.L. (1995) Evaluation of two commercial serologic tests for the diagnosis of *Helicobacter pylori* infection in the rhesus monkey. *Laboratory Animal Science* 45: 613-617.

Hendricks J.K. and Mobley H.L. (1997) *Helicobacter pylori* ABC transporter: effect of allelic exchange mutagenesis on urease activity. *Journal of Bacteriology* 179: 5892-5902.

Heneghan M.A. and Moran A.P. (2001) *Helicobacter pylori*, Lewis antigens, and inflammation. *Gut* 48: 869.

Higashi H., Tsutsumi R., Muto S., Sugiyama T., Azuma T., Asaka M. and Hatakeyama M. (2002) SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. *Science* 295: 683-686.

Hirata Y., Maeda S., Mitsuno Y., Akanuma M., Yamaji Y., Ogura K., Yoshida H., Shiratori Y. and Omata M. (2001) *Helicobacter pylori* activates the cyclin D1 gene through mitogen-activated protein kinase pathway in gastric cancer cells. *Infection and Immunity* 69: 3965-3971.

Hirschl A.M., Schulze K. and Hentsche E., (1994) Serological, microbiological and clinical results of a 43 months follow-up after successful eradication of *H. pylori* in duodenal ulcer disease. *The American Journal of Gastroenterology* 89: 1374-.

Hoch J.A. (2000) Two-component and phosphorelay signal transduction. *Current Opinion in Microbiology* 3: 165-170

Hofreuter D., Odenbreit S. and Haas R. (2001) Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. *Molecular Microbiology* 41: 379-391.

Huesca M., Goodwin A., Bhagwansingh A., Hoffman P. and Lingwood C.A. (1998) Characterization of an acidic-pH-inducible stress protein (hsp70), a putative sulfatide binding adhesin, from *Helicobacter pylori*. *Infection and Immunity* 66: 4061-4067.

Hulten K. Han S.W. and Enroth H. (1996) *Helicobacter pylori* in the drinking water in Peru. *Gastroenterology* 110: 1031-1035.

Ikenoue T., Maeda S., Ogura K., Akanuma M., Mitsuno Y., Imai Y., Yoshida H., Shiratori Y. and Omata M. (2001) Determination of *Helicobacter pylori* virulence by simple gene analysis of the *cag* pathogenicity island. Clinical and Diagnostic Laboratory Immunology 8: 181-186.

Ishihama A. (2000) Functional modulation of Escherichia coli RNA polymerase. Annual Review of Microbiology 54: 499-518.

Israel D.A., Salama N., Arnold C.N., Moss S.F., Ando T., Wirth H.P., Tham K.T., Camorlinga M., Blaser M.J., Falkow S. and Peek R.M. (2001) *Helicobacter pylori* strain-specific differences in genetic content, identified by microarray, influence host inflammatory responses. The Journal of Clinical Investigation 107: 611-620.

Iwamoto H., Czajkowsky D.M., Cover T.L., Szabo G. and Shao Z. (1999) VacA from *Helicobacter pylori*: a hexameric chloride channel. FEBS Letters 450: 101-104.

Jain R., Rivera M., Moore J. and Lake J. (2002) Horizontal gene transfer in microbial genome evolution. Theoretical Population Biology 61: 489-495.

Janssen P.J., Audit B. and Ouzounis C.A. (2001) Strain-specific genes of *Helicobacter pylori*: distribution, function and dynamics. Nucleic Acids Research 29: 4395-4404.

Jiang H., Kang D.C., Alexandre D. and Fisher P.B. (2000) RaSH, a rapid subtraction hybridization approach for identifying and cloning differentially expressed genes.

Proceedings of the National Academy of Sciences of the United States of America 97: 12684-12689.

Jiang Q., Hiratsuka K. and Taylor D.E. (1996) Variability of gene order in different *Helicobacter pylori* strains contributes to genome diversity. *Molecular Microbiology* 20: 833-842.

Jonkers D., Stobberingh E., de Bruine A., Arends J.W. and Stockbrugger R. (1997) Evaluation of immunohistochemistry for the detection of *Helicobacter pylori* in gastric mucosal biopsies. *The Journal of Infection* 35: 149-154.

Josenhans C., Niehus E., Amersbach S., Horster A., Betz C., Drescher B., Hughes K.T. and Suerbaum S. (2002) Functional characterization of the antagonistic flagellar late regulators FliA and FlgM of *Helicobacter pylori* and their effects on the *H. pylori* transcriptome. *Molecular Microbiology* 43: 307-322.

Josenhans C., Eaton K.A., Thevenot T. and Suerbaum S. (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. *Infection and Immunity* 68: 4598-4603.

Josenhans C., Labigne A. and Suerbaum S. (1995) Comparative ultrastructural and functional studies of *Helicobacter pylori* and *Helicobacter mustelae* flagellin mutants: both flagellin subunits, FlaA and FlaB, are necessary for full motility in *Helicobacter* species. *Journal of Bacteriology* 177: 3010-3020.

Joyce E.A., Chan K., Salama N.R. and Falkow S. (2002) Redefining bacterial populations: a post-genomic reformation. 3: 462-473.

Joyce E.A., Gilbert J.V., Eaton K.A., Plaut A. and Wright A. (2001) Differential gene expression from two transcriptional units in the *cag* pathogenicity island of *Helicobacter pylori*. Infection and Immunity 69: 4202-4209.

Joyce E.A., Bassler B.L. and Wright A. (2000) Evidence for a signaling system in *Helicobacter pylori*: detection of a *luxS*-encoded autoinducer. Journal of Bacteriology 182: 3638-3643.

Karita M. and Blaser M.J. (1998) Acid-tolerance response in *Helicobacter pylori* and differences between *cagA*⁺ and *cagA*⁻ strains. The Journal of Infectious Diseases 178: 213-219.

Karita M., Tummuru M.K., Wirth H.P. and Blaser M.J. (1996) Effect of growth phase and acid shock on *Helicobacter pylori cagA* expression. Infection and Immunity 64: 4501-4507.

Karita M., Tsuda M. and Nakazawa T. (1995) Essential role of urease *in vitro* and *in vivo* *Helicobacter pylori* colonization study using a wild-type and isogenic urease mutant strain. Journal of Clinical Gastroenterology 21 Suppl 1: S160-163.

Kawamata O., Yoshida H., Hirota K., Yoshida A., Kawaguchi R., Shiratori Y. and Omata M. (1996) Nested-polymerase chain reaction for the detection of *Helicobacter pylori* infection with novel primers designed by sequence analysis of urease A gene in clinically isolated bacterial strains. *Biochemical and Biophysical Research Communications* 219: 266-272.

Kelly S.M., Farmery S.M. and Gibson G.R. (1994) Isolation of *H. pylori* from faeces of patients with dyspepsia in the UK. *Gastroenterology* 107: 1671-1674.

Kersulyte D., Mukhopadhyay A.K., Shirai M., Nakazawa T. and Berg D.E. (2000) Functional organization and insertion specificity of IS607, a chimeric element of *Helicobacter pylori*. *Journal of Bacteriology* 182: 5300-5308.

Kersulyte D., Chalkauskas H. and Berg D.E. (1999) Emergence of recombinant strains of *Helicobacter pylori* during human infection. *Molecular Microbiology* 31: 31-43.

Kidd M., Lastovica A.J., Atherton J.C. and Louw J.A. (2001) Conservation of the *cag* pathogenicity island is associated with *vacA* alleles and gastroduodenal disease in South African *Helicobacter pylori* isolates. *Gut* 49: 11-17.

Kjoller M., Fischer A. and Justesen T. (1991) Transport conditions and number of biopsies necessary for culture of *Helicobacter pylori*. *European Journal of Clinical Microbiology & Infectious Diseases* 10: 166-167.

Klein P.D., Gaillour A., Opekun A.R., Smith E.O. (1991) Water source as risk factor for *Helicobacter pylori* infection in Peruvian children, Gastrointestinal Physiology Working Group. Lancet 337: 1503-1506.

Krishnamurthy P., Parlow M., Zitzer J.B., Vakil N.B., Mobley H.L., Levy M., Phadnis S.H. and Dunn B.E. (1998) *Helicobacter pylori* containing only cytoplasmic urease is susceptible to acid. Infection and Immunity 66: 5060-5066.

Kusters J.G., Gerrits M.M., van Strijp J.A. and Vandenbroucke Grauls C.M. (1997) Coccoid forms of *Helicobacter pylori* are the morphological manifestation of cell death. Infection and Immunity 65: 3672-3679.

Lan R. and Reeves P.R. (2000) Intraspecies variation in bacterial genomes: the need for a species genome concept. Trends in Microbiology 8: 396-401.

LaRiviere F.J., Wolfson A.D. and Uhlenbeck O.C. (2001) Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. Science 294: 165-168.

Latifi A., Winson K., Foglino M., Bycroft B. W., Stewart G. S., Lazdunski A., and Williams P. (1995) Multiple homologues of LuxR and LuxI control expression of virulence determinants and secondary metabolites through quorum sensing in *Pseudomonas aeruginosa* PAO1. Molecular Microbiology 17: 333-343.

Lee A., Dixon M.F., Danon S.J., Kuipers E., Megraud F., Larsson H. and Mellgard B. (1995) Local acid production and *Helicobacter pylori*: a unifying hypothesis of gastroduodenal disease. *European Journal of Gastroenterology & Hepatology* 7: 461-465.

Lee A., Fox, J. and Hazell, S. (1993). Pathogenicity of *Helicobacter pylori*: a perspective. *Infection and Immunity* 61: 1601-1610.

Lee A., Otto G, Dick E.H. and Krakowka S. (1991) Transmission of *Helicobacter spp.* A challenge to the dogma of faecal-oral spread. *Epidemiology of Infection* 107: 99-109.

Lee J.M., Breslin N.P., Fallon C. and O'Morain C.A. (2000a) Rapid urease tests lack sensitivity in *Helicobacter pylori* diagnosis when peptic ulcer disease presents with bleeding. *The American Journal of Gastroenterology* 95: 1166-1170.

Lee J.M., Deasy E. and O'Morain C.A. (2000b) *Helicobacter pylori* eradication therapy: a discrepancy between current guidelines and clinical practice. *European Journal of Gastroenterology & Hepatology* 12: 433-437.

Leodolter A. and Megraud F. (1999) Diagnosis of *Helicobacter pylori* infection. *Current Opinion in Gastroenterology* 17 (Suppl 1): S19-S23.

Leying H., Suerbaum S., Geis G. and Haas R. (1992) Cloning and genetic characterization of a *Helicobacter pylori* flagellin gene. *Molecular Microbiology* 6: 2863-2874.

Liang P. and Pardee A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257: 967-971.

Lin L.F., Posfai J., Roberts R.J. and Kong H. (2001) Comparative genomics of the restriction-modification systems in *Helicobacter pylori*. *Proceedings of the National Academy of Sciences of the United States of America* 98: 2740-2745.

Lindkvist P., Wadstrom T. and Giesecke J. (1995) *Helicobacter pylori* infection and foreign travel. *Journal of Infectious Diseases* 172: 1135-1136.

Liu G., McDaniel T.K., Falkow S. and Karlin S. (1999) Sequence anomalies in the Cag7 gene of the *Helicobacter pylori* pathogenicity island. *Proceedings of the National Academy of Sciences of the United States of America* 96: 7011-7016.

Loy C.T., Irwig L.M. and Katelaris P.H. (1996) Do commercial serological kits for *Helicobacter pylori* infection differ in accuracy? A meta-analysis. *American Journal Gastroenterology* 94: 1138-1144.

Lu J.J., Perng C.L., Shyu R.Y., Chen C.H., Lou Q., Chong S.K. and Lee C.H. (1999) Comparison of five PCR methods for detection of *Helicobacter pylori* DNA in gastric tissues. *Journal of Clinical Microbiology* 37: 772-774.

Luo Z.Q. and Farrand S.K. (2001) The *Agrobacterium tumefaciens* rnd homolog is required for TraR-mediated quorum-dependent activation of Ti plasmid tra gene expression. *Journal of Bacteriology* 183: 3919-3930.

Macnab R.M. (1996) Flagella and motility. In *Escherichia coli* and *Salmonella*: Cellular and molecular biology. Washington, DC: American Society for Microbiology Press.

Madan E., Kemp J., Westblom T.U., Subik M., Sexton S. and Cook J. (1988) Evaluation of staining methods for identifying *Campylobacter pylori*. *American Journal of Clinical Pathology* 90: 450-453.

Maeda S., Yoshida H., Ikenoue T., Ogura K., Kanai F., Kato N., Shiratori Y. and Omata M. (1999) Structure of *cag* pathogenicity island in Japanese *Helicobacter pylori* isolates. *Gut* 44: 336-341.

Malaty M G.D., Klein PD, Evans DG, Adam E, and Evans DJ. (1991) Transmission of *Helicobacter pylori* infection. Studies in families of healthy individuals. *Scandinavian Journal of Gastroenterology* 26: 927-932.

Malfertheiner P., Megraud F., O'Morain C., Hungin A.P., Jones R., Axon A., Graham D.Y., Tytgat G. and (2002) Current concepts in the management of *Helicobacter pylori* infection--the Maastricht 2-2000 Consensus Report. *Alimentary Pharmacology & Therapeutics* 16: 167-180.

Malfertheiner P., O'Connor H.J., Genta R.M., Unge P. and Axon A.T. (2002) Symposium: *Helicobacter pylori* and clinical risks--focus on gastro-oesophageal reflux disease. *Alimentary Pharmacology & Therapeutics* 16 Suppl 3: 1-10.

Mankoski R., Hoepf, T., Krakowka, S. and Eaton, K.A. (1999) The *flaA* mRNA transcription level correlates with *Helicobacter pylori* colonisation efficiency in gnotobiotic piglets. *Journal of Medical Microbiology* 48: 395-399.

Marshall B.J. and Goodwin C.S. (1987) Revised nomenclature of *campylobacter pyloridis*. *International Journal of Systematic Bacteriology* 37: 68.

Marshall B.J. and Warren J.R. (1984) Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* i: 1311-1315.

Marshall B.J. (1983) Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet* i: 1273-1275.

Marshall D.G., Hynes S.O., Coleman D.C., O'Morain C.A., Smyth C.J. and Moran A.P. (1999) Lack of a relationship between Lewis antigen expression and *cagA*, *CagA*, *vacA* and *VacA* status of Irish *Helicobacter pylori* isolates. *FEMS Immunology and Medical Microbiology* 24: 79-90.

Matsuoka M., Yoshida Y., Hayakawa K., Fukuchi S. and Sugano K. (1999) Simultaneous colonisation of *Helicobacter pylori* with and without mutations in the

23S rRNA gene in patients with no history of clarithromycin exposure. Gut 45: 503-507.

McColl K.L.E., El-Omar E. and Gillen D. (1997) Pathogenesis and host response in *Helicobacter pylori* infections: Normed Verlag, Bad Homburg, Germany.

McDaniel T.K. and Kaper J.B. (1997) A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. Molecular Microbiology 23: 399-407.

McGowan C.C., Cover T.L. and Blaser M.J. (1996) *Helicobacter pylori* and gastric acid: biological and therapeutic implications. Gastroenterology 110: 926-938.

McGowan C.C., Necheva A., Thompson S.A., Cover T.L. and Blaser M.J. (1998) Acid-induced expression of an LPS-associated gene in *Helicobacter pylori*. Molecular Microbiology 30: 19-31.

McNamara D. and O'Morain C. (1999) Gastro-oesophageal reflux disease and *Helicobacter pylori*: an intricate relation. Gut 45 Suppl 1: 13-17.

Megraud F., Bonnet F., Garnier M. and Lamouliatte H. (1985) Characterisation of "*campylobacter pyloridis*" by culture, enzymatic profiles and protein content. Journal of Clinical Microbiology 22: 1007-1010.

Megraud F. and Broutet N. (2000) Have we found the source of *Helicobacter pylori*?
Alimentary Pharmacology & Therapeutics 14 (Suppl 3): 7-12.

Mendall MA, Goggin P. and Molineaux N. (1992) Childhood living conditions and
Helicobacter pylori seropositivity in adult life. Lancet 339: 892-897.

Merrell D.S. and Camilli A. (2002) Acid tolerance of gastrointestinal pathogens.
Current Opinion in Microbiology 5: 51-55.

Meyer T., Covacci A., Kist M. and Pahl H.L. (2000) *Helicobacter pylori* activates
mitogen-activated protein kinase cascades and induces expression of the proto-
oncogenes *c-fos* and *c-jun*. The Journal of Biological Chemistry 275: 16064-16072.

Midolo P. and Marshall B.J. (2000) Accurate diagnosis of *Helicobacter pylori*: Urease
tests. Gastroenterology Clinics of North America 29: 871-878.

Miehlke S., Kirsch C., Agha-Amiri K., Gunther T., Lehn N., Malfertheiner P., Stolte
M., Ehninger G. and Bayerdorffer E. (2000) The *Helicobacter pylori vacA s1, m1*
genotype and *cagA* is associated with gastric carcinoma in Germany. International
Journal of Cancer 87: 322-327.

Mira A., Ochman H. and Moran N.A. (2001) Deletional bias and the evolution of
bacterial genomes. Trends in Genetics 17: 589-596.

Mitchell H.M., Li Y.Y. and Hu P.J. (1992) Epidemiology of *Helicobacter pylori* in southern China: identification of early childhood as the critical period for acquisition. *Journal of Infectious Diseases* 166: 149-153.

Miwa H., Yamada T., Sato K., Ohta K., Ohkura R., Murai T., Nagahara A., Takei Y., Ogiwara T. and Sato N. (2000) Efficacy of reduced dosage of rabeprazole in PPI/AC therapy for *Helicobacter pylori* infection: comparison of 20 and 40 mg rabeprazole with 60 mg lansoprazole. *Digestive Diseases and Sciences* 45: 77-82.

Miyaji H., Azama T. and Ito S. (2000) *Helicobacter pylori* infection occurs via close contact with infected individuals in early life. *Journal of Gastroenterology & Hepatology* 15: 257-262.

Mobley H.L., Cortesia, M.J., Rosenthal, L.E and Jones, B.D. (1988) Characterization of urease from *Campylobacter pylori*. *Journal of Clinical Microbiology* 26: 831-836.

Molinari M., Galli C., Norais N., Telford J.L., Rappuoli R., Luzio J.P. and Montecucco C. (1997) Vacuoles induced by *Helicobacter pylori* toxin contain both late endosomal and lysosomal markers. *The Journal of Biological Chemistry* 272: 25339-25344.

Molloy R.M. and Sonnenberg A. (1997) Relation between gastric cancer and previous peptic ulcer disease. *Gut* 40: 247-252.

Monteiro M.A., Chan K.H., Rasko D.A., Taylor D.E., Zheng P.Y., Appelmelk B.J., Wirth H.P., Yang M., Blaser M.J., Hynes S.O., Moran A.P. and Perry M.B. (1998)

Simultaneous expression of type 1 and type 2 Lewis blood group antigens by *Helicobacter pylori* lipopolysaccharides: Molecular mimicry between *H. pylori* lipopolysaccharides and human gastric epithelial cell surface glycoforms. The Journal of Biological Chemistry 273: 11533-11543.

Moran A.P., Knirel Y.A., Senchenkova S.N., Widmalm G., Hynes S.O. and Jansson P. (2002) Phenotypic variation in molecular mimicry between *Helicobacter pylori* lipopolysaccharides and human gastric epithelial cell surface glycoforms. The Journal of Biological Chemistry 277: 5785-5795.

Moran A.P. and Prendergast M.M. (2001) Molecular mimicry in *Campylobacter jejuni* and *Helicobacter pylori* lipopolysaccharides: contribution of gastrointestinal infections to autoimmunity. Journal of Autoimmunity 16: 241-256.

Moran A.P. (1999) *Helicobacter pylori* lipopolysaccharide-mediated gastric and extragastric pathology. Journal of Physiology and Pharmacology. 50: 787-805.

Moran A.P. (1997) Pathogenesis of gastric *Helicobacter pylori*. Trends in Microbiology 5: 262-263.

Moran A.P. (1996) Pathogenic properties of *Helicobacter pylori*. Scandinavian Journal of Gastroenterology 31: 22-31.

Moran A.P., Prendergast M.M. and Appelmelk B.J. (1996) Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease. *FEMS Immunology and Medical Microbiology* 16: 105-115.

Neidhardt F.C. (1996) *Escherichia Coli* and *Salmonella*. Cellular and Molecular Biology. Washington D.C.: ASM Press.

Newton M., Bryan R., Burnham W.R. and Kamm M.A. (1997) Evaluation of *Helicobacter pylori* in reflux oesophagitis and Barrett's oesophagus. *Gut* 40: 9-13.

Nguyen V.Q., Caprioli R.M. and Cover T.L. (2001) Carboxy-terminal proteolytic processing of *Helicobacter pylori* vacuolating toxin. *Infection and Immunity* 69: 543-546.

Ni Bhriain N. and Dorman C.J. (1993) Isolation and characterisation of a *topA* mutant of *Shigella flexneri*. *Molecular Microbiology* 7: 351-358.

Nobusato A., Uchiyama I. and Kobayashi I. (2000) Diversity of restriction-modification gene homologues in *Helicobacter pylori*. *Gene* 259: 89-98.

Nogueira C., Figueiredo C., Carneiro F., Gomes T.A., Barreira R., Figueira P., Salgado C., Belo L., Peixoto A., Bravo J.C., Bravo L.E., Realpe J.L., Plaisier A.P., Quint W.G.V., Ruiz B., Correa P. and Van Doorn, L.J. (2001) *Helicobacter pylori* genotypes may determine gastric histopathology. *American Journal of Pathology* 158: 647-654.

Occhialini A., Marais A., Urdaci M., Sierra R., Munoz N., Covacci A. and Megraud F. (2001) Composition and gene expression of the *cag* pathogenicity island in *Helicobacter pylori* strains isolated from gastric carcinoma and gastritis patients in Costa Rica. *Infection and Immunity* 69: 1902-1908.

Occhialini A., Marais A., Alm R., Garcia F., Sierra R and Megraud, F. (2000) Distribution of open reading frames of plasticity region of strain J99 in *Helicobacter pylori* strains isolated from gastric carcinoma and gastritis patients in Costa Rica. *Infection and Immunity* 68: 6240-6249.

Ochman H. (2001) Lateral and oblique gene transfer. *Current Opinion in Genetics and Development* 11: 616-619.

Ochman H. and Moran N.A. (2001) Genes lost and gene found: evolution of bacterial pathogenesis and symbiosis. *Science* 292: 1096-1101.

Ochman H., Lawerenc J.G. and Groisman E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* 405: 289-304.

Odenbreit S., Puls J., Sedlmaier B., Gerland E., Fischer W. and Haas R. (2000) Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. *Science* 287: 1497-1500.

Ohkusa T., Fujiki K., Takashimizu I., Kumagai J., Tanizawa T., Eishi Y., Yokoyama T. and Watanabe M. (2001) Improvement in atrophic gastritis and intestinal metaplasia in patients in whom *Helicobacter pylori* was eradicated. *Annals of Internal Medicine* 134: 380-386.

Owen R.J. (1998) *Helicobacter*-Species classification and identification. *British Medical Bulletin* 54: 17-30.

Pan Z.J., Berg D.E., Hulst V.D., Su W.W., Raudonikiene A., Xiao S.D., Dankert J., Tytgat G.N., and van der Ende A. (1998) Prevalence of vacuolating cytotoxin production and distribution of distinct vacA alleles in *Helicobacter pylori* from China. *Journal of Infectious Diseases* 178: 220-226.

Papini E., Zoratti M. and Cover T.L. (2001) In search of the *Helicobacter pylori* VacA mechanism of action. *Toxicon* 39: 1757-1767.

Papini E., Satin B., Norais N., de Bernard M., Telford J.L., Rappuoli R. and Montecucco C. (1998) Selective increase of the permeability of polarized epithelial cell monolayers by *Helicobacter pylori* vacuolating toxin. *The Journal of Clinical Investigation* 102: 813-820.

Parsonnet J., Friedman G.D., Vandersteen D.P., Chang Y., Vogelstein J.H., Orentreich N. and Sibley R.K. (1991) *Helocbacter pylori* infection and the risk of gastric carcinoma. *New England Journal of Medicine* 325: 1127-31.

Peek R.M., Blaser M.J., Mays D.J., Forsyth M.H., Cover T.L., Song S.Y., Krishna U. and Pietenpol J.A. (1999) *Helicobacter pylori* strain-specific genotypes and modulation of the gastric epithelial cell cycle. *Cancer Research* 59: 6124-6131.

Peek R.M., Thompson S.A., Donahue J.P., Tham K.T., Atherton J.C., Blaser M.J. and Miller G.G. (1995) Adherence to gastric epithelial cells induces expression of a *Helicobacter pylori* gene, *iceA*, that is associated with clinical outcome. *Proceedings of the Association of American Physicians* 110: 531-544.

Perez-Perez GI Witkin S.S., Decker M.D. and Blaser M.J. (1991) Seroprevalence of *Helicobacter pylori* infection in couples. *Journal of Clinical Microbiology* 29: 642-644.

Piccolomini R., Di Bonaventura G., Festi D., Catamo G., Laterza F. and Neri M. (1997) Optimal combination of media for primary isolation of *Helicobacter pylori* from gastric biopsy specimens. *Journal of Clinical Microbiology* 35: 1541-1544.

Polish L.B., Douglas J.J., Davidson A.J., Perez-Perez G.I. and Blaser M.J. (1991) Characterization of risk factors for *Helicobacter pylori* infection among men attending a sexually transmitted disease clinic: lack of evidence for sexual transmission. *Journal of Clinical Microbiology* 29: 2139-2143.

Polonczyk P K.S., Karczewska E, Bielanski W, Kaczmarczyk W, Stachowska A. (1996) Oral cavity as permanent reservoir of *Helicobacter pylori* and potential source of reinfection. *Journal of Physiology and Pharmacology* 47: 121-129.

Puls J., Fischer W. and Haas R. (2002) Activation of *Helicobacter pylori* CagA by tyrosine phosphorylation is essential for dephosphorylation of host cell proteins in gastric epithelial cells. *Molecular Microbiology* 43: 961-969.

Qasim A. and O'Morain C.A. (2002) Treatment of *Helicobacter pylori* infection and factors influencing eradication. *Alimentary Pharmacology & Therapeutics* 16 Suppl 1: 24-30.

Qi H., Menzel, R. and Tse-Dinh, Y-C. (1999) Increased thermosensitivity associated with topoisomerase I deletion and promoter mutations in *Escherichia coli*. *FEMS Microbiology Letters* 178: 141-146.

Raudonikiene A., Zakharova N., Su W.W., Jeong J.Y., Bryden L., Hoffman P.S., Berg D.E. and Severinov K. (1999) *Helicobacter pylori* with separate beta- and beta'-subunits of RNA polymerase is viable and can colonize conventional mice. *Molecular Microbiology* 32: 131-138.

Rotimi O., Cairns A., Gray S., Moayyedi P. and Dixon M.F. (2000) Histological identification of *Helicobacter pylori*: comparison of staining methods. *Journal of Clinical Pathology* 53: 756-759.

Rudi J., Kolb C., Maiwald M., Kuck D., Sieg A., Galle P.R., and Stremmel W. (1998) Diversity of *Helicobacter pylori* *vacA* and *cagA* genes and relationship to VacA and CagA protein expression, cytotoxin production, and associated diseases. *Journal of Clinical Microbiology* 36: 944-948.

Rudi J., Toppe H. and Marx N. (1997) Risk of infection with *Helicobacter pylori* and hepatitis A virus in different groups of hospital workers. The American Journal of Gastroenterology 92: 258-262.

Russmann H., Kempf V.A., Koletzko S., Heesemann J. and Autenrieth I.B. (2001) Comparison of fluorescent in situ hybridization and conventional culturing for detection of *Helicobacter pylori* in gastric biopsy specimens. Journal of Clinical Microbiology 39: 304-308.

Ryan K.A., van Doorn L.J., Moran A.P., Glennon M., Smith T. and Maher M. (2001) Evaluation of clarithromycin resistance and *cagA* and *vacA* genotyping of *Helicobacter pylori* strains from the west of Ireland using line probe assays. Journal of Clinical Microbiology 39: 1978-1980.

Ryan K.A., Moran A.P., Hynes S.O., Smith T., Hyde D., O'Morain C.A. and Maher M. (2000) Genotyping of *cagA* and *vacA*, Lewis antigen status, and analysis of the poly-(C) tract in the alpha(1,3)-fucosyltransferase gene of Irish *Helicobacter pylori* isolates. FEMS Immunology and Medical Microbiology 28: 113-120.

Sahay P. and Axton A.T.R. (1996) Reservoirs of *Helicobacter pylori* and modes of transmission. *Helicobacter* 1: 175-182.

Sahay P., West A.P., Hawkey P.M. and Axon A.T.R. (1995) Isolation of *Helicobacter pylori* from faeces. Journal of Infection 30: 262-263.

Salama N., Guillemin K., McDaniel T.K., Sherlock G., Tompkins L. and Falkow S. (2000) A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. Proceedings of the National Academy of Sciences of the United States of America 97: 14668-14673.

Sambrook J., Fritsh E.F. and Maniatis T. (1989) Molecular cloning: A laboratory manual, 2nd edition. Cold Spring Harbour, New York: Cold Spring Harbor Press.

Satin B., Del-Giudice G., Della Bianca V., Dusi S., Laudanna C., Tonello F., Kelleher D., Rappuoli R., Montecucco C. and Rossi F. (2000) The neutrophil-activating protein (HP-NAP) of *Helicobacter pylori* is a protective antigen and a major virulence factor. The Journal of Experimental Medicine 191: 1467-1476.

Saunders N.J., Peden J.F., Hood D.W. and Moxon E.R. (1998) Simple sequence repeats in the *Helicobacter pylori* genome. Molecular Microbiology 27: 1091-1098.

Saunders K.E., McGovern K.J. and Fox J.G. (1997) Use of pulsed-field gel electrophoresis to determine genomic diversity in strains of *Helicobacter hepaticus* from geographically distant locations. Journal of Clinical Microbiology 35: 2859-2863.

Schwizer W., Thumshirn M., Dent J., Guldenschuh I., Menne D., Cathomas G. and Fried M. (2001) *Helicobacter pylori* and symptomatic relapse of gastro-oesophageal reflux disease: a randomised controlled trial. Lancet 357: 1738-1742.

Segal E.D., Lange C., Covacci A., Tompkins L.S. and Falkow S. (1997) Induction of host signal transduction pathways by *Helicobacter pylori*. *Proceedings of the National Academy of Sciences of the United States of America* 94: 7595-7599.

Sepulveda A., Peterson L.E., Shelton J., Gutierrez O. and Graham D.Y. (2002) Histological patterns of gastritis in *H. pylori*-infected individuals with a family history of gastric cancer. *The American Journal of Gastroenterology* 97: 1365-1370.

Seshadri R., Hendrix, L.R. and Samuel, J.E. (1999) Differential expression of translational elements by life cycle variants of *Coxiella burnetii*. *Infection and Immunity* 67: 6026-6033.

Seto K., Hayashi-Kuwabara Y., Yoneta T., Suda H. and Tamaki H. (1998) Vacuolation induced by cytotoxin from *Helicobacter pylori* is mediated by the EGF receptor in HeLa cells. *FEBS Letters* 431: 347-350.

Sevin E., Lamarque D., Delchier J.C., Soussy C.J. and Tankovic J. (1998) Co-detection of *Helicobacter pylori* and of its resistance to clarithromycin by PCR. *FEMS Microbiology Letters* 165: 369-372.

Shimizu T., Akamatsu T., Ota H. and Katsuyama T. (1996) Immunohistochemical detection of *Helicobacter pylori* in the surface mucus gel layer and its clinical-pathological significance. *Helicobacter* 1: 197-206.

Shirai M., Fujinaga R., Akada J.K. and Nakazawa T. (1999) Activation of *Helicobacter pylori ureA* promoter by a hybrid *Escherichia coli*-*H. pylori rpoD* gene in *E. coli*. *Gene* 239: 351-359.

Simor A.E., Cooter N.B. and Low D.E. (1990) Comparison of four stains and a urease test for rapid detection of *Helicobacter pylori* in gastric biopsies. *European Journal of Clinical Microbiology & Infectious Diseases* 9: 350-352.

Siu L.K., Leung W.K., Cheng A.F., Sung J.Y., Ling T.K., Ling J.M., Ng E.K., Lau J.Y. and Chung S.C. (1998) Evaluation of a selective transport medium for gastric biopsy specimens to be cultured for *Helicobacter pylori*. *Journal of Clinical Microbiology* 36: 3048-3050.

Skovgaard M., Jensen L.J., Brunak S., Ussery D. and Krogh A. (2001) On the total number of genes and their length distribution in complete microbial genomes. *Trends in Genetics* 17: 425-428.

Slater E., Owen R.J., Williams M. and Pounder R.E. (1999) Conservation of the *cag* pathogenicity island of *Helicobacter pylori*: associations with vacuolating cytotoxin allele and IS605 diversity. *Gastroenterology* 117: 1308-1315.

Slev P. and Potts W. (2002) Disease consequences of pathogen adaptation. *Current Opinion in Immunology* 14: 609-612.

Slonczewski J.L., McGee D.J., Phillips J., Kirkpatrick C. and Mobley H.L. (2000) pH-dependent protein profiles of *Helicobacter pylori* analyzed by two-dimensional gels. *Helicobacter* 5: 240-247.

Smeets L.C., Bijlsma J.J., Kuipers E.J., Vandenbroucke-Grauls C.M. and Kusters J.G. (2000) The *dprA* gene is required for natural transformation of *Helicobacter pylori*. *FEMS Immunology and Medical Microbiology* 27: 99-102.

Solnick J.V., Hansen L.M. and Syvanen M. (1997) The major sigma factor (RpoD) from *Helicobacter pylori* and other gram-negative bacteria shows an enhanced rate of divergence. *Journal of Bacteriology* 179: 6196-6200.

Sperandio V., Mellies, J. L., Nguyen, W., Shin, S., and Kaper, J. B. (1999) Quorum sensing controls expression of the type III secretion gene transcription and protein secretion in enterohemorrhagic and enteropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 96: 15196-15201.

Spohn G. and Scarlato V. (1999a) Motility of *Helicobacter pylori* is coordinately regulated by the transcriptional activator FlgR, an NtrC homolog. *Journal of Bacteriology* 181: 593-599.

Spohn G. and Scarlato V. (1999b) The autoregulatory HspR repressor protein governs chaperone gene transcription in *Helicobacter pylori*. *Molecular Microbiology* 34: 663-674.

Spohn G., Beier D., Rappuoli R. and Scarlato V. (1997) Transcriptional analysis of the divergent *cagAB* genes encoded by the pathogenicity island of *Helicobacter pylori*. *Molecular Microbiology* 26: 361-372.

Stein M., Bagnoli F., Halenbeck R., Rappuoli R., Fantl W.J. and Covacci A. (2002) c-Src/Lyn kinases activate *Helicobacter pylori* CagA through tyrosine phosphorylation of the EPIYA motifs. *Molecular Microbiology* 43: 971-980.

Stein M., Rappuoli R. and Covacci A. (2000) Tyrosine phosphorylation of the *Helicobacter pylori* CagA antigen after *cag*-driven host cell translocation. *Proceedings of the National Academy of Sciences of the United States of America* 97: 1263-1268.

Stingl K., Altendorf K. and Bakker E.P. (2002a) The compartment buffered by the urease of *Helicobacter pylori*: cytoplasm or periplasm? *Trends in Microbiology* 10: 218-219.

Stingl K., Altendorf K. and Bakker E.P. (2002b) Acid survival of *Helicobacter pylori*: how does urease activity trigger cytoplasmic pH homeostasis? *Trends in Microbiology* 10: 70-74.

Stingl K., Uhlemann E.M., Schmid R., Altendorf K. and Bakker E.P. (2002c) Energetics of *Helicobacter pylori* and its implications for the mechanism of urease-dependent acid tolerance at pH 1. *Journal of Bacteriology* 184: 3053-3060.

Stingl K., Uhlemann-Em E.M., Deckers-Hebestreit G., Schmid R., Bakker E.P. and Altendorf K. (2001) Prolonged survival and cytoplasmic pH homeostasis of *Helicobacter pylori* at pH 1. *Infection and Immunity* 69: 1178-1180.

Stock A.M., Robinson V.L. and Goudreau P.N. (2000) Two-component signal transduction. *Annual Review of Biochemistry* 69: 183-215.

Stone G.G., Shortridge D., Flamm R.K., Beyer J., Ghoneim A.T. and Tanaka S.K. (1997) PCR-RFLP typing of *ureC* from *Helicobacter pylori* isolated from gastric biopsies during a European multi-country clinical trial. *The Journal of Antimicrobial Chemotherapy* 40: 251-256.

Strobel S., Bereswill S., Balig P., Allgaier P., Sonntag H.G. and Kist M. (1998) Identification and analysis of a new *vacA* genotype variant of *Helicobacter pylori* in different patient groups in Germany. *Journal of Clinical Microbiology* 36: 1285-1289.

Suerbaum S. and Achtman M. (1999) Evolution of *Helicobacter pylori*: the role of recombination. *Trends in Microbiology* 7: 182-184.

Suerbaum S., Brauer Steppkes T., Labigne A., Cameron B. and Drlica K. (1998a) Topoisomerase I of *Helicobacter pylori*: juxtaposition with a flagellin gene (*flaB*) and functional requirement of a fourth zinc finger motif. *Gene* 210: 151-161.

Suerbaum S., Smith J.M., Bapumia K., Morelli G., Smith N.H., Kunstmann E., Dyrek I. and Achtman M. (1998b) Free recombination within *Helicobacter pylori*.

Proceedings of the National Academy of Sciences of the United States of America 95: 12619-12624.

Suerbaum S. (1995) The complex flagella of gastric *Helicobacter* species. Trends in Microbiology 3: 168-170; 170-161.

Suerbaum S., Josenhans C. and Labigne A. (1993) Cloning and genetic characterization of the *Helicobacter pylori* and *Helicobacter mustelae* flaB flagellin genes and construction of *H. pylori* flaA- and flaB-negative mutants by electroporation-mediated allelic exchange. Journal of Bacteriology 175: 3278-3288.

Szcebara F., Dhaenens L., Armand S. and Husson M.O. (1999) Regulation of the transcription of genes encoding different virulence factors in *Helicobacter pylori* by free iron. FEMS Microbiology Letters 175: 165-170.

Takata T., Aras R., Tavakoli D., Ando T., Olivares A.Z. and Blaser M.J. (2002) Phenotypic and genotypic variation in methylases involved in type II restriction-modification systems in *Helicobacter pylori*. Nucleic Acids Research 30: 2444-2452.

Taylor D.N., and Blaser M. J. (1991) The epidemiology of *Helicobacter pylori* infection. Epidemiological Reviews 13: 42-59.

Tazawa K. and Tsutsumi Y. (1998) Effect of prolonged staining with hematoxylin on detecting *Helicobacter pylori* in hematoxylin-eosin-stained gastric mucosa. International Pathology 48: 448-452.

Tee W., Fairley S., Smallwood R. and Dwyer B. (1991) Comparative evaluation of three selective media and a nonselective medium for the culture of *Helicobacter pylori* from gastric biopsies. *Journal of Clinical Microbiology* 29: 2587-2589.

Thomas J.E. Gibson G., Darboe M.K., Dale A. and Gibson G.R. (1992) Isolation of *H. pylori* from human faeces. *Lancet* 340: 1194-1195.

Tomb J.F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G., Fleischmann R.D., Ketchum K.A., Klenk H.P., Gill S., Dougherty B.A., Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S., Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D. and Venter J.C. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388: 539-547.

Tonello F., Dundon W.G., Satin B., Molinari M., Tognon G., Grandi G., Del-Giudice G., Rappuoli R. and Montecucco C. (1999) The *Helicobacter pylori* neutrophil-activating protein is an iron-binding protein with dodecameric structure. *Molecular Microbiology* 34: 238-246.

Trebesius K., Adler K., Vieth M., Stolte M. and Haas R. (2001) Specific detection and prevalence of *Helicobacter heilmannii*-like organisms in the human gastric mucosa by fluorescent *in situ* hybridisation and partial 16S ribosomal DNA sequencing. *Journal of Clinical Microbiology* 39: 1510-1516.

Tsang K.W. and Lam S.K. (1999) *Helicobacter pylori* and extra-digestive diseases. Journal of Gastroenterology and Hepatology 14: 844-850.

Tsuda M., Karita M., Mizote T., Morshed M.G., Okita K. and Nakazawa T. (1994) Essential role of *Helicobacter pylori* urease in gastric colonization: definite proof using a urease-negative mutant constructed by gene replacement. European Journal of Gastroenterology & Hepatology 6 Suppl 1: S49-52.

Tsuda M., Karita M., Morshed M.G., Okita K. and Nakazawa T. (1994) A urease-negative mutant of *Helicobacter pylori* constructed by allelic exchange mutagenesis lacks the ability to colonize the nude mouse stomach. Infection and Immunity 62: 3586-3589.

Tummuru M., Cover, T.L. and Blaser, M.J. (1993) Cloning and expression of a high-molecular-mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production. Infection and Immunity. 61: 1799-1809.

Uemura N., Okamoto S., Yamamoto S., Matsumura N., Yamaguchi S., Yamakido M., Taniyama K., Sasaki N. and Schlemper R.J. (2001) *Helicobacter pylori* infection and the development of gastric cancer. The New England Journal of Medicine 345: 784-789.

Uemura N., Mukai T., Okamoto S., Yamaguchi S., Mashiba H., Taniyama K., Sasaki N., Haruma K., Sumii K. and Kajiyama G. (1997) Effect of *Helicobacter pylori*

eradication on subsequent development of cancer after endoscopic resection of early gastric cancer. *Cancer Epidemiology, Biomarkers & Prevention* : 6: 639-642.

Vaira D., Gatta L., Ricci C. and Miglioli M. (2002) Diagnosis of *Helicobacter pylori* infection. *Alimentary Pharmacology & Therapeutics* 16 (Suppl. 1): 16-23.

Vaira D., Holton, J., Ricci, C., Menegatti, M., Gatta, L., Berardi, S., Tampieri, A. and Miglioli M. (2001) The transmission of *Helicobacter pylori* from stomach to stomach. *Alimentary Pharmacology & Therapeutics* 15: 33-40.

Vaira D., Malfertheiner P. and Megraud F. (1999) Diagnosis of *Helicobacter pylori* infection with a new non-invasive antigen-based assay. *Lancet* 354: 30-33.

Van der Hulst R.W., Verheul S.B., Weel J.F., Gerrits Y., ten Kate F.J., Dankert J. and Tytgat G.N. (1996) Effect of specimen collection techniques, transport media, and incubation of cultures on the detection rate of *Helicobacter pylori*. *European Journal of Clinical Microbiology & Infectious Diseases* 15: 211-215.

Van Doorn L.J., Figueiredo C., Megraud F., Pena S., Midolo P., Queiroz D.M., Carneiro F., Vanderborght B., Pegado M.D., Sanna R., De Boer W., Schneeberger P.M., Correa P., Ng E.K., Atherton J., Blaser M.J. and Quint W.G. (1999) Geographic distribution of *vacA* allelic types of *Helicobacter pylori*. *Gastroenterology* 116: 823-830.

Van Doorn L.J., Figueiredo C., Sanna R., Pena S., Midolo P., Ng E.K., Atherton J.C., Blaser M.J. and Quint W.G. (1998a) Expanding allelic diversity of *Helicobacter pylori* vacA. *Journal of Clinical Microbiology* 36: 2597-2603.

Van Doorn L.J., Figueiredo C., Sanna R., Plaisier A., Schneeberger P., de-Boer W. and Quint W. (1998b) Clinical relevance of the *cagA*, *vacA*, and *iceA* status of *Helicobacter pylori*. *Gastroenterology* 115: 58-66.

van Vliet A.H., Poppelaars S.W., Davies B.J., Stoof J., Bereswill S., Kist M., Penn C.W., Kuipers E.J. and Kusters J.G. (2002) NikR mediates nickel-responsive transcriptional induction of urease expression in *Helicobacter pylori*. *Infection and Immunity* 70: 2846-2852.

Vandamme P., Falsen E., Rossau R., Hoste B., Segers P., Tytgat R. and De Ley J. (1991) Revision of *Campylobacter*, *Helicobacter*, and *Wolinella* taxonomy: Recommendation of generic descriptions and proposal of *Arcobacter* gen. nov. *International Journal of Systematic Bacteriology* 41: 88-103.

Varanasi R.V., Fantry G.T. and Wilson K.T. (1998) Decreased prevalence of *Helicobacter pylori* infection in gastroesophageal reflux disease. *Helicobacter* 3: 188-194.

Veenendaal R.A., Lichtendahl-Bernards A.T., Pena A.S., Endtz H.P., van-Boven C.P. and Lamers C.B. (1993) Effect of transport medium and transportation time on culture

of *Helicobacter pylori* from gastric biopsy specimens. Journal of Clinical Pathology 46: 561-563.

Veldhuyzen Van Zanten S., Lauritsen K., Delchier J.C., Labenz J., De Argila C.M., Lind T., Treichel H.C., Stubberod A., Cockeram A., Hasselgren G., Gothe L., Wrangstadh M. and Sinclair P. (2000) One-week triple therapy with esomeprazole provides effective eradication of *Helicobacter pylori* in duodenal ulcer disease. Alimentary Pharmacology & Therapeutics 14: 1605-1611.

Versalovic J., Osato M.S., Spakovsky K., Dore M.P., Reddy R., Stone G.G., Shortridge D., Flamm R.K., Tanaka S.K. and Graham D.Y. (1997) Point mutations in the 23S rRNA gene of *Helicobacter pylori* associated with different levels of clarithromycin resistance. The Journal of Antimicrobial Chemotherapy 40: 283-286.

Vuthoori S., Bowers C.W., McCracken A., Dombroski A.J. and Hinton D.M. (2001) Domain 1.1 of the sigma(70) subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. Journal of Molecular Biology 309: 561-572.

Wang G., Ge Z., Rasko D.A. and Taylor D.E. (2000) Lewis antigens in *Helicobacter pylori*: biosynthesis and phase variation. Molecular Microbiology 36: 1187-1196.

Wang G., Humayun M.Z. and Taylor D.E. (1999) Mutation as an origin of genetic variability in *Helicobacter pylori*. Trends in Microbiology 7: 488-493.

Wang X., Sturegard E., Rupar R., Nilsson H.O., Aleljung P.A., Carlen B., Willen R. and Wadstrom T. (1997) Infection of *BALB/c* A mice by spiral and coccoid forms of *Helicobacter pylori*. *Journal of Medical Microbiology* 46: 657-663.

Warren J.R. (1983) Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet* i: 1273.

Watanabe T., Tada M., Nagai H., Sasaki S. and Nakao M. (1998) *Helicobacter pylori* infection induces gastric cancer in mongolian gerbils. *Gastroenterology* 115: 642-648.

Webb G.F. and Blaser M.J. (2002) Dynamics of bacterial phenotype selection in a colonized host. *Proceedings of the National Academy of Sciences of the United States of America* 99: 3135-3140.

Weeks D.L. and Sachs G. (2001) Sites of pH regulation of the urea channel of *Helicobacter pylori*. *Molecular Microbiology* 40: 1249-1259.

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

Weiss R. (2002) Virulence and pathogenesis. *Trends in Microbiology* 10: 314-317.

Weston A.P., Badr A.S., Topalovski M., Cherian R., Dixon A. and Hassanein R.S. (2000) Prospective evaluation of the prevalence of gastric *Helicobacter pylori* infection

in patients with GERD, Barrett's esophagus, Barrett's dysplasia, and Barrett's adenocarcinoma. *The American Journal of Gastroenterology* 95: 387-394.

Wirth H.P., Yang M., Karita M. and Blaser M.J. (1996) Expression of the human cell surface glycoconjugates Lewis x and Lewis y by *Helicobacter pylori* isolates is related to *cagA* status. *Infection and Immunity* 64: 4598-4605.

Wong B.C., Yin Y., Berg D.E., Xia H.H., Zhang J.Z., Wang W.H., Wong W.M., Huang X.R., Tang V.S. and Lam S.K. (2001) Distribution of distinct *vacA*, *cagA* and *iceA* alleles in *Helicobacter pylori* in Hong Kong. *Helicobacter* 6: 317-324.

Worku M.L., Sidebotham R.L., Walker M.M., Keshavarz T. and Karim Q.N. (1999) The relationship between *Helicobacter pylori* motility, morphology and phase of growth: implications for gastric colonization and pathology. *Microbiology* 145: 2803-2811.

Wosten M. (1998) Eubacterial sigma-factors. *FEMS Microbiology Reviews* 22: 127-150.

Xia H.X., Keane, C.T. and O'Morain, C.A. (1994) Culture of *Helicobacter pylori* under aerobic conditions on solid media. *European Journal of Clinical Microbiology and Infectious Diseases* 13: 406-409.

Xia H.X., English L., Keane C.T. and O'Morain C.A. (1993) Enhanced cultivation of *Helicobacter pylori* in liquid media. *Journal of Clinical Pathology* 46: 750-753.

Yahiro K., Niidome T., Hatakeyama T., Aoyagi H., Kurazono H., Padilla P.I., Wada A. and Hirayama T. (1997) *Helicobacter pylori* vacuolating cytotoxin binds to the 140-kDa protein in human gastric cancer cell lines, AZ-521 and AGS. *Biochemical and Biophysical Research Communications* 238: 629-632.

Yamaoka Y., El-Zimaity H.M., Gutierrez O., Figura N., Kim J.G., Kodama T., Kashima K., Graham D.Y. and Kim J.K. (1999) Relationship between the *cagA* 3' repeat region of *Helicobacter pylori*, gastric histology, and susceptibility to low pH. *Gastroenterology* 117: 342-349.

Yamaoka Y., Kodama T., Gutierrez O., Kim J.G., Kashima K. and Graham D.Y. (1999) Relationship between *Helicobacter pylori* *iceA*, *cagA*, and *vacA* status and clinical outcome: studies in four different countries. *Journal of Clinical Microbiology* 37: 2274-2279.

Yamaoka Y., Kodama T., Kashima K., Graham D.Y. and Sepulveda A.R. (1998) Variants of the 3' region of the *cagA* gene in *Helicobacter pylori* isolates from patients with different *H. pylori*-associated diseases. *Journal of Clinical Microbiology* 36: 2258-2263.

Yoshizawa Y. and Yamamoto Y. (1989) Characterisation of a nalidixic acid- resistant mutant of *Escherichia coli* as a strict aerobe. *Microbiology and Immunity* 33: 449-457.

Young C.C. and Furano A.V. (1981) Regulation of the synthesis of *E. coli* elongation factor Tu. Cell 24: 695-706.

Young K.A., Allaker R.P. and Hardie J.M. (2001) Morphological analysis of *Helicobacter pylori* from gastric biopsies and dental plaque by scanning electron microscopy. Oral Microbiology and Immunology 16: 178-181.

Zakharova N., Paster B.J., Wesley I., Dewhirst F.E., Berg D.E. and Severinov K.V. (1999) Fused and overlapping *rpoB* and *rpoC* genes in *Helicobacters*, *Campylobacters*, and related bacteria. Journal of Bacteriology 181: 3857-3859.

Zakharova N., Hoffman P.S., Berg D.E. and Severinov K. (1998) The largest subunits of RNA polymerase from gastric *Helicobacters* are tethered. The Journal of Biological Chemistry 273: 19371-19374.

APPENDIX

List of *H. pylori* strains used in this study

Strains	P32	F5	F10	F17	P18	<i>cagA</i>	<i>cag7</i>	<i>cag2</i>
98587	-	-	-	-	-	+	+	+
3655	-	+	+	-	-	+	+	+
3519*	+	+	-	-	+	+	+	+
3612	-	-	+	-	+	+	+	+
3582	+	+	+	-	+	+	+	+
3568	+	-	-	-	+	+	+	+
3526	-	-	+	-	+	+	+	+
8402	/	/	/	/	/	+	+	+
8131	/	/	/	/	/	-	+	+
71104	/	/	/	/	/	+	+	+
8928	/	/	/	/	/	+	+	+
5734	-	+	/	/	/	+	+	+
X363	-	-	/	/	/	+	+	+
5835	-	+	/	/	/	+	+	+
X306	-	+	/	/	/	+	+	+
775	-	+	/	/	/	+	+	+
7169	-	+	/	/	/	+	+	+
5803	+	+	/	/	/	+	+	+

List of *H. pylori* strains used in this study (continued)

Strains	P32	F5	F10	F17	P18	<i>cagA</i>	<i>cag7</i>	<i>cag2</i>
4573	-	+	/	/	/	+	+	+
41123	+	+	/	/	/	+	+	+
5713	+	+	/	/	/	+	+	+
5714	-	+	/	/	/	+	+	+
7189	-	+	/	/	/	+	+	+
1981	/	/	/	/	/	+	+	+
3802	/	/	/	/	/	+	+	+
31116	/	/	/	/	/	+	+	+
492	+	-	/	/	/	+	+	+
4728	-	-	/	/	/	+	+	+
2111	/	/	/	/	/	+	+	+
31125	/	/	/	/	/	+	+	+
2311	/	/	/	/	/	+	+	+
4694	-	+	/	/	/	+	+	+
4774	/	/	/	/	/	+	+	+
3567	/	/	/	/	/	+	+	+
3581	+	+	/	/	/	+	+	+
3620	/	/	/	/	/	+	+	+
4811	/	/	/	/	/	+	+	+
5351	/	/	/	/	/	+	+	+

List of *H. pylori* strains used in this study (continued)

Strains	P32	F5	F10	F17	P18	<i>cagA</i>	<i>cag7</i>	<i>cag2</i>
8295	/	/	/	/	/	+	+	+
461	/	/	/	/	/	+	+	+
X663	/	/	/	/	/	+	-	+
X625*	/	/	/	/	/	+	+	+
8136	+	-	/	/	/	+	+	+
X402	/	/	/	/	/	+	+	+
8290	/	/	/	/	/	+	+	+
5533	-	+	/	/	/	+	-	+
548*	-	+	/	/	/	+	+	+
5780	/	/	/	/	/	+	+	+
41174	/	/	/	/	/	-	+	+
4739	/	/	/	/	/	-	+	+
5561	/	/	/	/	/	-	+	+
4804	-	-	/	/	/	-	+	+
5671	-	-	/	/	/	+	+	+
5631	-	+	/	/	/	+	+	+
5556	/	/	/	/	/	+	+	+
5755	+	-	/	/	/	+	+	+
581	/	/	/	/	/	+	+	+
41091	-	+	/	/	/	+	+	+

List of *H. pylori* strains used in this study (continued)

Strains	P32	F5	F10	F17	P18	<i>cagA</i>	<i>cag7</i>	<i>cag2</i>
4786	+	+	/	/	/	+	+	+
4233	-	-	/	/	/	+	+	+
4730	+	-	/	/	/	+	+	+
41008	/	/	/	/	/	+	+	+
4956	-	+	/	/	/	+	+	+
8442	/	/	/	/	/	+	+	+
5786	/	/	/	/	/	+	+	+
8225*	-	+	/	/	/	+	+	+
8268	-	-	/	/	/	+	+	+
8283	-	+	/	/	/	+	+	+
8323	-	+	/	/	/	+	+	+
8325*	+	+	/	/	/	-	-	-
159	+	+	/	/	/	-	-	-
487	+	+	/	/	/	-	-	-
X143	-	+	/	/	/	-	-	-
X665	+	+	/	/	/	-	-	-
77 [#]	-	-	+	-	+	+	-	+
5612 [#]	-	+	+	-	+	+	+	+
8248 [#]	-	-	+	+	+	+	+	+
5536 [#]	-	-	+	-	+	+	+	+

List of *H. pylori* strains used in this study (continued)

Strains	P32	F5	F10	F17	P18	<i>cagA</i>	<i>cag7</i>	<i>cag2</i>
5283 [#]	-	+	-	-	+	+	+	+
7231 [#]	-	-	+	-	+	+	+	+
3517 [#]	-	+	+	-	-	+	+	+
8260 [#]	-	-	-	-	+	+	+	+
<i>CI</i>	+	+	+	+	+	+	+	+
41289	-	-	-	-	+	+	+	+
7109	+	+	-	-	+	+	+	+
863	-	-	+	-	+	+	-	+

+, present; - absent; /, not determined; * PCR failed to amplify *cagA* or *cag7*, but was detected using Southern blot; #, gastritis with intestinal metaplasia; strains in regular font, isolated from duodenal ulcer patients; strains in italic font, isolated from gastritis patients; strains in bold font, isolated from gastric cancer patients