**REVIEW
ARTICLE*****Helicobacter pylori* – a conundrum of genetic diversity**

David G. Marshall, William G. Dundon, Sarah M. Beesley
and Cyril J. Smyth

Author for correspondence: Cyril J. Smyth. Tel: +353 1 608 1195. Fax: +353 1 679 9294.
e-mail: csmyth@tcd.ie

Department of Microbiology, Moyné Institute of Preventive Medicine, Trinity College, University of Dublin, Dublin 2, Republic of Ireland

Keywords: *Helicobacter pylori*, genetic diversity, horizontal transfer, natural transformation, repetitive DNA

Background

In the past 15 years the recognition of *Helicobacter pylori* infection as a cause of gastroduodenal morbidity has probably had more impact on health care worldwide than any other emerging infection barring HIV. *H. pylori* is one of the commonest human gastrointestinal pathogens, infecting almost 50% of the human population (Taylor & Blaser, 1991). Since its rediscovery as curved bacteria in the stomach by histological examination of gastric biopsies (Warren & Marshall, 1983; Marshall, 1989; Rathbone & Heatley, 1989b; Buckley & O'Morain, 1998) and its subsequent first isolation by Marshall & Warren (1984), *H. pylori* (a microaerophilic, Gram-negative spiral bacterium) has been shown to be associated with chronic superficial gastritis (inflammation of the gastric mucosa), chronic active gastritis (gastric mucosal inflammation with polymorphonuclear cell infiltration), peptic ulcer disease (gastric and duodenal ulcers) and gastric cancer (NIH Consensus Development Panel, 1994). Infection of the stomach with *H. pylori* appears to occur in early infancy. The bacterium colonizes the gastric mucosa and causes acute gastritis that progresses into chronic active gastritis, a condition that appears to persist for life (Taylor & Parsonnet, 1995). During the years or decades that follow the initial infection, the chronic active gastritis may remain asymptomatic, or may progress into more severe forms of the disease, such as a peptic ulcer, or the almost complete loss of gastric epithelial cell function observed in atrophic gastritis (NIH Consensus Development Panel, 1994).

A recent survey to analyse trends in microbiological research revealed that approximately 7.5% of all publications between 1991 and 1997 were concerned with the genus *Helicobacter* and that the annual publication rate on *Helicobacter* had tripled over that period (Gálvez *et al.*, 1998). Indeed, several books devoted entirely to *H. pylori* have been published in the last decade (Blaser,

1989; Rathbone & Heatley, 1989a; Malfertheiner & Ditschuneit, 1990; Marshall *et al.*, 1991; Goodwin & Worsley, 1993; Ó'Moráin & O'Connor, 1994; Calam, 1996; Hunt & Tytgat, 1996; Lee & Mégraud, 1996; Moran & O'Morain, 1997; Farthing & Patchett, 1998), as well as comprehensive reviews in journals (e.g. Lee & Hazell, 1988; Dunn *et al.*, 1997), numerous topical articles (e.g. Suerbaum, 1995; Covacci *et al.*, 1997a; Kuipers & Blaser, 1998) and a journal launched in 1996, namely, *Helicobacter*.

One of the most intriguing aspects of *H. pylori* is its genetic diversity, which is at a level as yet unseen among other bacterial pathogens, the biological significance of which is enigmatic (Blaser, 1996; Logan & Berg, 1996; Covacci & Rappuoli, 1998). Digestive organs with acid-secreting epithelia, i.e. stomachs, arose over 300 million years ago in vertebrate evolution. Blaser (1996) has suggested that the diversity of *H. pylori* possibly reflects the diversity of humans, their ancestors, their environments and their diets, which has led to a state of balanced polymorphism in the *H. pylori* population in a genetically diverse human population during life-long infection of individuals. This review aims to describe aspects of the genetic diversity and genomic malleability of *H. pylori*.

The *H. pylori* genome sequence

A complete genome sequence of *H. pylori* was published by Tomb *et al.* (1997). The authors chose to sequence the genome of *H. pylori* strain 26695 which had been originally isolated from a patient with gastritis. This strain colonizes piglets, elicits an immune and inflammatory response and is also known to be toxigenic, plasmid-free and naturally competent (Tomb *et al.*, 1997). It contains a single circular chromosome of approximately 1.7 Mb with 1590 predicted ORFs, one third of which are of unknown function (for review see Berg *et al.*, 1997a). Analysis of this genome sequence

may reveal important information on the genome structure of this *H. pylori* strain and perhaps provide clues as to how the genetic variability within this species may occur. However, despite the wealth of information that has already been gleaned and may be further gathered from the sequence of the *H. pylori* genome, unfortunately the genome sequence of strain 26695 does not hold all the answers. To fully understand the genetic diversity of *H. pylori*, a second if not a third genome should be sequenced and released into the public domain. The genome of a second strain, J99, has been sequenced in large part, but is licensed to the pharmaceutical company Astra (Hancock *et al.*, 1998). This has allowed those with access to both genomic sequences to make some interesting observations on pairs of genes encoding outer-membrane proteins (Hancock *et al.*, 1998). Indeed, Hancock *et al.* (1998) have suggested from comparative analysis of the complete genome sequence of strain 26695 and the incomplete genome sequence of strain J99 that many of the *H. pylori* genes are highly conserved.

Since the release of the genome sequence of strain 26695 it has become apparent that the assignment of roles to putative ORFs in *H. pylori*, and indeed in other genomes, must be performed with caution as, firstly, most of the putative functions have not as yet been subjected to biochemical or mutational confirmation, thereby overlooking the possibility of the evolution of a new function for the gene product, and, secondly, ORF assignment can vary depending on the system used for analysis of the genome sequence (Berg *et al.*, 1997a; Saunders *et al.*, 1998). Berg *et al.* (1997a) cite, for example, the tricarboxylic acid cycle enzyme 2-oxoglutarate oxidoreductase, which is present in *H. pylori* but is not listed among the ORFs in the genome sequence. Furthermore, the analysis of the genome sequence by Saunders *et al.* (1998) used a different strategy to that of Tomb *et al.* (1997) and in doing so revealed clearly identifiable differences in ORF assignment.

Physical comparison of genomes

H. pylori

Macrodiversity within a species is described as gross genetic differences between strains, a phenomenon that can be readily seen in *H. pylori*. Using various molecular techniques to compare the genomic DNA of unique isolates, i.e. isolates recovered from gastric biopsy samples from different individuals, data supporting the macrodiversity of *H. pylori* have been obtained.

The extent of genomic DNA sequence similarity among *H. pylori* isolates has been investigated by Yoshimura *et al.* (1993) and Takami *et al.* (1994) using thermal denaturation analysis. Both groups reported similar findings and calculated the DNA sequence similarity between the majority of isolates to be above 80%. However, a number of isolates had sequence similarities of less than 80%, a finding that Takami *et al.* (1994)

postulated to be evidence for the existence of *H. pylori* subspecies.

The use of PFGE has allowed analysis of *H. pylori* genomes for macrorestriction patterns by digestion of DNA with restriction endonucleases that cut genomic DNA infrequently. The genome sizes in *H. pylori* have been assessed by this technique and were reported to vary from 1.0 Mb to 1.8 Mb (Takami *et al.*, 1993). In addition, the presence of plasmids in *H. pylori* isolates is not uniform, with carriage rates reported to vary between 35 to 59% of isolates (e.g. Simor *et al.*, 1990; Kleanthous *et al.*, 1991; Xiang *et al.*, 1995).

The degree of conservation of gene location between the genomes of isolates has also been determined using PFGE, which, in combination with Southern hybridization with gene probes, generates information about gene order within a strain, thereby allowing a map of the genetic loci to be deduced. In all, four *H. pylori* strains have been mapped by PFGE, namely, strain UA802 which was mapped by Taylor *et al.* (1992) and a further three strains, NCTC 11637, NCTC 11639 and UA861, which were mapped by Jiang *et al.* (1996). An ordered cosmid library of strain NCTC 11638 was made by Bukanov & Berg (1994), which, in combination with PFGE, yielded a high-resolution physical genetic map of this strain. The gene orders of all five isolates were compared by Jiang *et al.* (1996), revealing that there is little conservation of gene location among the strains, with some genes being found in the same 25% of the genome but with little maintenance of their order. This inter-strain variability in gene location is in marked contrast, for example, to the inter-genus conservation of gene order observed between *Escherichia coli* and *Salmonella typhimurium* (Krawiec & Riley, 1990). The genome of *Neisseria gonorrhoeae* also displays considerable variability but, unlike *H. pylori*, this diversity is not manifest in variability in genome arrangement (Dempsey *et al.*, 1995).

Macrodiversity among other *Helicobacter* genomes

Comparison of the level of genomic diversity of *H. pylori* with that of other *Helicobacter* spp. by PFGE revealed that *Helicobacter hepaticus* has a level of genomic variability which is intermediate between that of *H. pylori* and *Helicobacter mustelae* (Saunders *et al.*, 1997). *H. mustelae* appears not to exhibit the heterogeneity present in *H. pylori*, although minor differences in PFGE pattern are present (Taylor *et al.*, 1994). Why the profound genomic diversity of *H. pylori* is not present across the entire *Helicobacter* genus is an interesting question.

Genotypic analysis of *H. pylori* populations

DNA fingerprinting of *H. pylori*

A number of genotyping methods have been applied to *H. pylori* for epidemiological analysis. Random amplification of polymorphic DNA analysis (RAPD)

(Akopyanz *et al.*, 1992), oligofingerprinting (Marshall *et al.*, 1996) and repetitive element-PCR (Rep-PCR) (Go *et al.*, 1995) are the most discriminatory fingerprinting methods applied to *H. pylori*. All can easily distinguish unrelated isolates by the fingerprint profiles they generate. Other DNA-fingerprinting methods, namely, *Hind*III-RFLP (Langenberg *et al.*, 1986), *Hae*III-ribotyping (Owen *et al.*, 1992), PCR-RFLP analysis of *fla* (Forbes *et al.*, 1995) and *ure* genes (Foxall *et al.*, 1992; Fujimoto *et al.*, 1994; Shortridge *et al.*, 1997), have also been frequently applied to the differentiation of *H. pylori* strains, but fail to discriminate all epidemiologically unrelated isolates tested. From the DNA-fingerprinting studies applied to *H. pylori*, it is clear that isolates display a level of heterogeneity in fingerprint profiles unseen in other bacteria, with isolates recovered from different individuals generally having different fingerprint profiles by most methods.

Evidence for multiple strain colonization

The use of the above techniques on isolates from multiple biopsies recovered from colonized individuals has revealed data on the proportion of individuals colonized by more than a single strain. Colonization of individuals with more than a single strain appears to be rare among Europeans and North Americans. For example, Taylor *et al.* (1995), using RAPD, found more than one colonizing strain in only three out of 15 patients tested, Marshall (1996), using oligofingerprinting, found two strains in only one out of 19 patients screened, and Shortridge *et al.* (1997), using PCR-RFLP of *ureC* genes, found multiple strains in only two out of 81 patients examined. In comparison, in developing countries, where colonization with *H. pylori* is more common, up to 34% of individuals tested have been found to be colonized with multiple strains (Berg *et al.*, 1997b).

Substrain diversity revealed by genotypic analysis – evidence for microevolution

When the most discriminatory DNA-fingerprinting methods are used to examine isolates recovered from different biopsy sites within individuals, i.e. on epidemiologically related isolates, very similar but not identical fingerprint profiles are often observed, suggesting that rearrangements are occurring in the genomes of strains during infection. The first evidence to support the occurrence of genomic rearrangements in *H. pylori* came from *Hind*III-RFLP analysis of isolates recovered from multiple biopsies obtained from individuals, where the profiles generated for approximately 33% of isolates displayed subtypic variations (Oudbier *et al.*, 1990). Furthermore, using oligofingerprinting, up to 60% of individuals tested were found to be colonized with subtypic isolates (Marshall, 1996). Moreover, when 10 biopsies were recovered from each patient, Jorgensen *et al.* (1996) found that 16 out of 17 individuals tested displayed subtypic *H. pylori* populations. Several studies have looked at the fingerprint profiles of multiple single

colony isolates from the same biopsy specimen using RFLP or RAPD analysis (Oudbier *et al.*, 1990; Marshall *et al.*, 1995; Jorgensen *et al.*, 1996). These studies have all found a homogeneous population to be present at a single site except for Jorgensen *et al.* (1996), who suggested that substrains may be present in a small proportion of biopsies. *Hae*III-ribotyping analysis of isolates recovered from members of the same family has also revealed subtypic isolate profiles in different individuals, suggesting that a common clone may have initially infected the individuals (Nwokolo *et al.*, 1992).

The minor band differences obtained by the most discriminatory techniques in fingerprint profiles of epidemiologically related *H. pylori* isolates probably represent rearrangements occurring *in vivo* in the genomic organization within single clones of *H. pylori* (Oudbier *et al.*, 1990; Jorgensen *et al.*, 1996; Marshall, 1996). Such changes in genetic loci between clones over a relatively short period of time are referred to as microevolution (Achtman, 1994; Hobbs *et al.*, 1994). These events probably occur in most, if not all, *H. pylori* strains colonizing individuals and are, therefore, likely to be a major contributory factor to the heterogeneous nature of the genomic organization of this species. It appears likely that different strains undergo microevolution in isolation in different individuals which would undoubtedly contribute to the extensive genetic diversity present between isolates.

The discovery of transposable elements in *H. pylori* (IS605; see section on *cag*-pathogenicity island below) is of interest in this regard (Censini *et al.*, 1996; Akopyants *et al.*, 1998; Höök-Nikanne *et al.*, 1998). Movement of insertion sequences (ISs) in a genome occurs with about the same frequency as other mutational events (Berg & Berg, 1996). Höök-Nikanne *et al.* (1998) have suggested that tests for the copy number and chromosomal distribution of these IS elements could be useful in tracing the microevolution of a strain during long-term carriage in individual hosts and in tracking the transmission of a strain among epidemiologically linked individuals.

Population structure of *H. pylori*

Natural variability in *H. pylori*

Population genetics is the study of the natural variability of bacterial populations and has led to the formulation of theories to account for this variability. By using multilocus enzyme electrophoresis (MLEE) data, the frequency of recombination in natural populations can be estimated by calculating the index of association (I_A) value between loci. *H. pylori* has an I_A value that does not differ significantly from zero and, therefore, has a panmictic or freely recombining population structure (Go *et al.*, 1996), with genetic diversity in the sample tested exceeding the level of diversity recorded in all other bacterial species studied by MLEE. Non-clonal bacterial species, like *H. pylori*, are unlikely to show the genetic stability of clones. There is, therefore, little

rationale in attempting to compare the genetic relatedness of epidemiologically unrelated isolates of non-clonal species by phylogenetic trees produced from DNA-fingerprinting profiles (Maynard Smith, 1995). Accordingly, if meaningful conclusions are to be drawn from such data, studies of isolate populations should probably be restricted to epidemiologically related isolates.

Quasi-species

The possibility that within *H. pylori* there may exist subspecies or what has also been termed a species-complex or quasi-species has been put forward because of the relatively low level of clonality observed in this species (Hazell *et al.*, 1997; Covacci & Rappuoli, 1998). While such hypotheses are valid, a study by Eckloff *et al.* (1994), which examined 16S rDNA sequences, classified *H. pylori* as a unique species. They found that the sequence of 16S rDNA, the defining molecule of a species, among five *H. pylori* isolates showed 0.2–0.5% divergence. In comparison the 16S rDNA sequences of other *Helicobacter* species revealed divergences between species which ranged from 2.7 to 8.0%.

Genetic diversity in virulence-associated loci

Virulence-associated loci

The association of *H. pylori* infection with the pathogenesis of a spectrum of diseases has been recognized since its isolation but remains poorly understood. The disparate outcome of infection in different individuals suggests the existence of strains of *H. pylori* with enhanced virulence potential and the possibility of the presence in some strains of virulence-associated loci (Atherton, 1997). The contribution of host and environmental factors to the outcome of *H. pylori*-associated disease cannot be discounted, but is beyond the scope of this review (Riccardi & Rotter, 1994; Pérez-Pérez & Parsonnet, 1994; Kuipers & Blaser, 1998). Accordingly, the variance present among *H. pylori* isolates at the level of (i) carriage of specific genes by sub-groups of isolates and their reported associations with clinical outcome and (ii) the genetic diversity at the level of gene structure and gene expression of specific genes is highlighted below.

vacA

The *vacA* gene which encodes the vacuolating cytotoxin is present in all *H. pylori* strains but active toxin is produced by only 50% as determined by the ability of strains to induce vacuolation of epithelial cells *in vitro* (Cover & Blaser, 1992). It was thought up to very recently that lack of expression of the vacuolating cytotoxin in 50% of strains was due to sequence divergence in the *vacA* genes of cytotoxin-negative strains which had approximately 65% nucleotide sequence identity with the corresponding loci in cytotoxin-producing strains (Cover *et al.*, 1994). However, Forsyth *et al.* (1998) have shown that the *vacA* gene is transcribed

in all strains, but that transcription occurs at higher levels in cytotoxin-positive strains than in cytotoxin-negative strains. Therefore, they suggest that differences in *vacA* transcription levels among *H. pylori* strains may be a factor that contributes to different vacuolating cytotoxin phenotypes (Forsyth *et al.*, 1998).

In addition to this inter-strain heterogeneity in *VacA* expression, the *vacA* alleles also exhibit a type of genotypic variation known as mosaicism (Fig. 1a), with conserved regions and highly divergent regions occurring within these alleles (Atherton *et al.*, 1995). Three signal sequence types exist, namely, s1a, s1b and s2. Until recently, all mid-regions were classified as types m1 and m2. However, additional variants have been identified in Germany, designated m1a (Strobel *et al.*, 1998), in Taiwan, designated m1T and m1Tm2 (a chimaeric type) (Wang *et al.*, 1998), and in China, designated m1b and m1b–m2 (a hybrid type) (Pan *et al.*, 1998), suggesting that *vacA* typing schemes may need to be revised.

There is much evidence for a strong correlation between the *vacA* genotype of a strain and disease outcome. Strains of the signal sequence/mid-region combination type s1/m1 were identified by Atherton *et al.* (1995) as producing high levels of the vacuolating cytotoxin *in vitro* and were associated with increased gastric epithelial damage, enhanced gastric inflammation and duodenal ulceration (Atherton *et al.*, 1997). Furthermore, strains of signal sequence type s1 are associated with the presence of the *cagA* gene (see below) and numerous studies have highlighted the prevalence of type s1 strains in individuals with peptic ulcer disease (Rudi *et al.*, 1998; Strobel *et al.*, 1998; van Doorn *et al.*, 1998a). Interestingly, strains of signal sequence type s1a and mid-region type m1 were isolated with remarkable prevalence (almost 100%) in Japan (Ito *et al.*, 1997), although no correlation with clinical consequences was reported. The same high prevalence of type s1a strains was also found in Taiwan and China, but, in contrast to Japan, the dominant mid-region type was m2 (Pan *et al.*, 1998; Wang *et al.*, 1998). However, strains with the type m1b and m1b–m2 alleles produced significantly higher levels of the *VacA* cytotoxin than did those with type m2 alleles (Pan *et al.*, 1998) and strains with the type m1T allele occurred statistically more often in peptic ulcer disease patients than did strains with type m2 alleles (Wang *et al.*, 1998).

fla and *ure*

The flagella and urease of *H. pylori* are important factors involved in the initial colonization of the gastric mucosa (Josenhans & Suerbaum, 1997; Mobley *et al.*, 1997) and are present in all *H. pylori* isolates. The gene sequences which encode the flagellin, *flaA* and *flaB*, showed extensive sequence variability, as evidenced by RFLPs when PCR products were digested with restriction endonucleases (Ohta-Tada *et al.*, 1997). Forbes *et al.* (1995) presented evidence that the sequence diversity found in the *flaA* and *flaB* genes had resulted

(a) VacA protein family

Signal sequence types

s1a

s1b

s2

Mid-region types

m1

m1a

m2

(b) Outer-membrane protein family

(c) 3' regions of *cagA* genes

Fig. 1. (a) Schematic representation of the diversity in the VacA protein family. The s2/m2 combination of signal sequence and mid-region of *vacA* alleles produces no detectable cytotoxin, while the s2/m1 combination has not yet been encountered (adapted from Blaser, 1997). (b) Schematic representation of two *bab* allele products, BabA1 and BabA2, showing regions of amino acid sequence similarity (black) and heterogeneity (white). Inserted below is the nucleotide sequence of the *babA2* upstream and signal peptide region showing the translation initiation codon (in bold) and the 10 bp repeat motif which is absent in the *babA1* allele (adapted from Ilver *et al.*, 1998). (c) Schematic representation of the primary structure of the 3' regions of the *cagA* gene types A, B, C and D. R1 repeats, 15 bp; R2 repeats, 42 bp; R3 repeats, 147 bp. The deduced amino acid sequence of the R3 region of the CagA protein differed with respect to a 15-amino-acid-long motif between Japanese isolates and strains from other geographic regions (adapted from Yamaoka *et al.*, 1998). These schematic figures are not drawn to scale.

typing of strains (Foxall *et al.*, 1992; Fujimoto *et al.*, 1994; Shortridge *et al.*, 1997).

babA

The BabA adhesin of *H. pylori*, which binds to the fucosylated Lewis b (Le^b) histo-blood group antigen, is a putative virulence factor and is strongly associated with the presence of the *cag*-pathogenicity island (see below). The gene encoding BabA was recently cloned by Ilver *et al.* (1998). Three *bab* alleles have been identified (Fig. 1b), but it is not yet known whether *bab* alleles are present in all *H. pylori* strains. Two of these alleles, namely *babA2* and *babB*, encode functional proteins. However, the third allele *babA1* is identical to the *babA2* allele with the exception of a 10 bp deletion of a repeat motif in the signal peptide sequence, which results in ablation of the translational initiation codon. This suggests that hot spots for phenotypic (phase) variation exist within the *bab* gene family. The *bab* genes belong to a family of over 30 *H. pylori* genes which have been proposed to encode outer-membrane proteins showing extensive amino acid sequence identity in the N-terminal and C-terminal domains, suggesting the potential for recombinational events leading to mosaicism (Tomb *et al.*, 1997; Ilver *et al.*, 1998).

cagA

H. pylori strains can be arbitrarily divided into six groups, based on whether they possess the *cag* genetic locus and express the CagA (cytotoxin-associated antigen) protein and active VacA toxin (Xiang *et al.*, 1995). Isolates carrying the *cag* genetic locus are more commonly recovered from patients with duodenal ulcers and gastric tumours (Blaser *et al.*, 1995). However, studies from Japan, Korea and China have shown that more than 90% of *H. pylori* strains were *cagA*⁺ irrespective of clinical presentation (Miehlke *et al.*, 1996; Pan *et al.*, 1997; Shimoyama *et al.*, 1997; Maeda *et al.*, 1998).

The *cagA* gene is part of the *cagA*-pathogenicity island (see following section) and encodes a highly immunogenic outer-membrane protein of variable molecular mass, 120–140 kDa, and of unknown function (Covacci *et al.*, 1993; Tummuru *et al.*, 1993). Neither expression of the VacA toxin nor induction of interleukin 8 production was affected by inactivation of the *cagA* gene, suggesting that CagA may only be a marker for increased virulence (Tummuru *et al.*, 1994; Crabtree *et al.*, 1995). The *cagA* gene has a highly conserved 5' region, while variation in the size of the CagA protein has been correlated with the presence of a variable number of repeat sequences encoded by the 3' region of the *cagA* gene (Covacci *et al.*, 1993; Tummuru *et al.*, 1993).

Analysis of the 3' sequences of 155 *cagA*-gene-positive isolates identified four types of *cagA* structure, designated A, B, C and D, based on their primary sequences and structural organizations (Fig. 1c; Yamaoka *et al.*, 1998). Interestingly, part of the amino

from reassortment of sequences between strains in the bacterial population. Similarly, the urease structural genes *ureA*, *ureB* and *ureC* exhibit extensive restriction site diversity, which has been exploited for the molecular

acid sequence of the R3 region of the CagA protein in isolates from Japan, KIASAGKGVGGFSGA, was different from that of strains from other geographical areas, FPLKRHDKVDDLSKV (Covacci *et al.*, 1993; Tummuru *et al.*, 1993). Type C strains were associated with more severe degrees of gastric atrophy, higher levels of anti-CagA antibody and gastric cancer. The findings suggested that the *cagA* genotype may provide a useful marker for differences in virulence among *cagA*-positive isolates (Yamaoka *et al.*, 1998).

cag-pathogenicity island

In an attempt to more fully understand the relationship between the CagA protein and virulence, Censini *et al.* (1996) analysed the DNA regions flanking the *cagA* gene. This led to the discovery of the *cag*-pathogenicity island (*cag*-PAI), recently confirmed by the extensive study of Akopyants *et al.* (1998).

The *cag*-PAI has many of the features of classic pathogenicity islands (Censini *et al.*, 1996; Covacci *et al.*, 1997b; Hacker *et al.*, 1997), namely, (i) a G+C content of 35 mol% that differs from the mean of 39 mol% for the rest of the *H. pylori* genome, (ii) flanking direct repeats involved in the integration of the 40 kb locus into the genome, (iii) putative virulence genes and genes that encode a type IV secretion system and (iv) insertion sequences.

The nucleotide sequencing of the 40 kb *cag*-PAI revealed the presence of a novel insertion sequence, IS605, which consists of two genes, *tnpA* and *tnpB*, that, based on the amino acid sequence similarities of the translated gene products, are thought to encode transposases (Censini *et al.*, 1996) (Fig. 2). These genes are flanked by two nucleotide sequences of 33 bp and 41 bp, respectively, that possess dyadic symmetry and a common core sequence. IS605 was shown to be present in 56% of the 44 clinical isolates tested and was found in varying copy number within these strains, inserted both into the *cag*-

PAI and into sites on the rest of the genome. The maximal number of copies of IS605 was associated with *cag*-PAI⁺ strains, while *cag*-PAI⁻ strains did not appear to possess the IS605 element (Censini *et al.*, 1996). However, Akopyants *et al.* (1998) found that 9 out of 16 *cag*-PAI⁺ strains lacked IS605 and Höök-Nikanne *et al.* (1998) reported that the prevalence of IS605 was similar in 37 *cag*-positive (27%) and in 13 *cag*-negative (23%) strains. The presence of IS605 did not correlate with production of the VacA toxin, with the s-type and m-type of *vacA* alleles, or with clinical manifestations of *H. pylori* infection (Höök-Nikanne *et al.*, 1998). From one to nine copies of IS605 have been detected per genome (Censini *et al.*, 1996; Akopyants *et al.*, 1998; Höök-Nikanne *et al.*, 1998). Interestingly, five full-length copies of IS605 occur in the genome sequence of strain 26695, but none of them is associated with the *cag*-PAI (Tomb *et al.*, 1997).

The structure of the *cag*-PAI varies among strains (Fig. 3): (i) it can consist of a single uninterrupted unit; (ii) it can be split into two regions, *cagI* and *cagII*, by IS605; (iii) it can be split into *cagI* and *cagII* by a large piece of chromosomal DNA that is flanked by IS605; and (iv)–(vi) it can consist of partially deleted segments of *cagI*, *cagII* or both (Censini *et al.*, 1996). This variation in the structure of the *cag*-PAI is most likely due to the insertion of IS605 followed by homologous recombination, which may explain the existence of *cag*-PAI⁺ and *cag*-PAI⁻ strains and their intermediates (Censini *et al.*, 1996; Covacci *et al.*, 1997a).

The identification of the *cag*-PAI and its strong association with a functional *vacA* gene have led to some hypotheses on the evolution of *cag*-PAI⁺ *H. pylori*. It has been suggested that the inheritance of the *cag*-PAI resulted in the generation of a functional vacuolating toxin from a pre-existing molecule by a gain-of-function mechanism, thereby allowing for a more virulent clone, which could then out-compete less virulent strains in the stomach of the host (Covacci *et al.*, 1997a). It has been reported that in some cases both *cagA*-positive and *cagA*-negative strains can be isolated from the same individual (van der Ende *et al.*, 1996). These findings have previously been presented as evidence for multiple infections. However, it is now thought that a dynamic equilibrium may exist between the gain and loss of the *cag*-PAI (Covacci *et al.*, 1997a; Covacci & Rappuoli, 1998).

Strain-variable genetic loci

In addition to the *cag*-PAI several novel genetic loci, which are only carried by a proportion of strains, have been identified. A recently identified region containing an ORF related to the virulence-associated protein D (VapD) from *Dichelobacter nodosus* was situated 3.5 kb downstream from the *vacA* gene in strain 60190 (Cao & Cover, 1997). In *H. pylori*, the *vapD* gene has been detected in approximately 61% of isolates. The presence of the *vapD* gene is not associated with a specific *vacA* allele type (Cao & Cover, 1997).

Fig. 2. Schematic representation of the IS605 element. (a) General structure showing the direction of transcription of the *tnpA* and *tnpB* genes. L and R, left (41 bp) and right (33 bp) ends of the element. (b) Nucleotide sequence of the left (L) and right (R) ends of the IS605 element. The arrows indicate the dyadic structure of the ends. A conserved hexanucleotide between both ends is underlined. (Adapted from Censini *et al.*, 1996.)

Fig. 3. *cag*-PAI structures: (i) single uninterrupted unit; (ii) *cag*-PAI split into two segments by the insertion element IS605; (iii) *cag*-PAI split into two segments by chromosomal DNA flanked by IS605; (iv)–(vi) partially deleted segments of *cagI* and *cagII* or both. Δ , Partial deletions; L and R, left and right end of the IS605 element, respectively. The figure is not to scale. (Adapted from Censini *et al.*, 1996.)

The search for genes that are possibly induced by contact with epithelial cells led to the identification by Peek *et al.* (1996) of a locus known as *ice* (induced by contact with epithelium). There are two distinct *ice* sub-groups, namely, *iceA1* and *iceA2*. Only *iceA1* expression increased following adherence to epithelium. The *iceA1* gene product showed 65% identity with a restriction endonuclease, *NlaIII* of *Neisseria lactamica*, suggesting a putative function (Peek *et al.*, 1996; Xu *et al.*, 1997). Strains of the *iceA1* sub-group have been shown to be strongly associated with duodenal ulcer disease, while *iceA2* strains were present at higher frequency in individuals with gastritis (Peek *et al.*, 1996; van Doorn *et al.*, 1998b). The *iceA* allelic type was independent of the *cagA* and *vacA* status (van Doorn *et al.*, 1998b).

A genetic locus of unknown function designated *trl* (for tRNA-associated locus, formerly known as *del*), which was isolated using arbitrarily primed PCR of RNA, was only present in 48% of clinical isolates tested (Dundon *et al.*, 1998). No correlation between the presence or absence of *trl* and the *cagA*/*CagA* or *vacA*/*VacA* status or the clinical origin of strains has been revealed (Dundon *et al.*, 1998).

Phase-variable genes

A number of homopolymeric tracts and dinucleotide repeat regions have been identified from analysis of the *H. pylori* genome sequence, many of which are within the promoter regions or ORFs of putative genes. Slipped strand mispairing of these repeat sequences 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 (Levinson & Gutman, 1987). The presence of the homopolymeric tracts and dinucleotide repeat regions may, therefore, facilitate the phase variation of the *H. pylori* ORFs they are associated with. Similar repeat sequences are presumed to be sites for slipped strand mispairing in *Haemophilus influenzae* (Moxon *et al.*, 1994; Hood *et al.*, 1996).

The *H. pylori* genome study by Tomb *et al.* (1997) identified a total of 30 genes with either dinucleotide repeats or homopolymeric tracts associated with them

(Table 1). Eight of these genes, five of which encode outer-membrane proteins, have dinucleotide repeats at their 5' end, suggesting that a mechanism for phase variation may exist in *H. pylori* (Saunders *et al.*, 1998). There is also a suggestion from Tomb *et al.* (1997) that phenotypic variation may occur at the transcriptional level for 27 genes that have homopolymeric tracts in their promoter regions. Analysis of the genome sequence by Saunders *et al.* (1998), however, revealed nine such additional putative phase-variable genes (Table 1). These genes, along with the 27 described previously by Tomb *et al.* (1997), can be divided into five groups: (i) five are putatively involved in LPS biosynthesis; (ii) 14 encode putative cell-surface-associated proteins; (iii) six are part of the DNA restriction/modification system; (iv) six are of unknown function, showing no sequence similarity with genes or gene products in the public domain nucleotide databases; and (v) five encode miscellaneous proteins. However, Saunders *et al.* (1998) recommend caution when ascribing phase-variable functions based on the presence of homopolymeric tracts alone. They highlight that the absence of a number of the genes involved in the SOS response results in mismatch repair in *H. pylori* and may explain the prevalence of long homopolymeric tracts in the genome.

The O-antigen of the LPS of 85% of *H. pylori* isolates contains the tissue/blood group structure Lewis x (Le^x) or Lewis y (Le^y) or both, similar to those commonly occurring in human gastric mucosa (Aspinall *et al.*, 1994; Aspinall & Monteiro, 1996). This molecular mimicry between *H. pylori* LPS and host Lewis antigens is thought to have a role in autoimmunity in *H. pylori*-linked disease (Negrini *et al.*, 1993; Appelmelk *et al.*, 1996; Aspinall & Moran, 1997). Evidence for phase variation was recently reported for the expression of LPS in *H. pylori* (Appelmelk *et al.*, 1998). In a type strain of *H. pylori* (NCTC 11637) expression of LPS was not stable. The LPS in this given strain, which expresses Le^x , displayed a high frequency of phase variation, resulting in the presence of several LPS variants in one bacterial cell population. Variants that expressed a derivative of Le^x lacking fucose, monomeric Le^y , and both Le^x and Le^y were detected. Part of the variation observed was ascribed to altered levels of expression of fucosyl-transferases that are encoded by genes known to possess oligonucleotide repeats, thereby suggesting possible

Table 1. Phase-variable candidate genes [adapted from Tomb *et al.* (1997) and Saunders *et al.* (1998)]

Repeat	Gene similarities	TIGR number	Reference*
LPS biosynthesis			
11 [AG], 11 A	RfaJ/glycosyltransferase	0208	1, 2
13 C	(1-3)- α -Fucosyltransferase	0379	1, 2
13 C	Lex2B	0619	1, 2
13 C	(1-3)- α -Fucosyltransferase	0651	1, 2
14 C	(1-2)- α -Fucosyltransferase	0093 and 0094	2
Cell-surface associated			
11 [CT], 14 A	Outer-membrane protein	0009	1
15 T	Outer-membrane protein	0025	1
15 C	Hypothetical protein; streptococcal M protein	0058	1, 2
12 C	Methyl-accepting chemotaxis protein	0103	2
14 T	Outer-membrane protein	0227	1
9 T	Haem-binding lipoprotein/transport protein	0298	2
6 [CT]	Outer-membrane protein	0638	1, 2
9 C	FliP	0684 and 0685	2
8 [CT], 16 T	Outer-membrane protein/adhesin	0722	1, 2
6 [CT], 14 T	Outer-membrane protein/adhesin	0725	1, 2
5 [GA]	Alginate O-acetylation protein	0855	2
16 T	Iron-regulated outer-membrane protein (<i>frpB</i>)	0876	1
11 [CT], 14 A	Outer-membrane protein/adhesin	0896	1, 2
13 T	Outer-membrane protein	0912	1
10 A	ABC transporter/secretion protein	1206	2
14 A	Outer-membrane protein	1342	1
9 [GA]	Conserved hypothetical protein; outer-membrane adherence protein associated protein	1417	1, 2
DNA restriction/modification			
5 [CT]	Cytosine-specific methyltransferase	0051	2
15 G	Type I restriction enzyme R protein	0464	1, 2
12 C and 15 C	Adenine-specific methyltransferase	1353 and 1354†	1, 2
10 C	R/M methyltransferase	1369 and 1370	2
14 C	Restriction enzyme β -subunit	1471	1, 2
12 G	Adenine-specific methyltransferase	1522	1, 2
Miscellaneous			
5 [AT]	Hypothetical	0211	2
12 G	Hypothetical	0217	1, 2
14 A	Inner-membrane protein	0228	1
9 G	Hypothetical	0335	2
15 T	<i>pyrG</i>	0349	1
15 A	Inner-membrane protein	0350	1
14 A	<i>cagA</i>	0547	1
8 T	Hypothetical	0586 and 0585	2
15 T	Hypothetical	0629	1
14 T	Hypothetical	0733	1
9 [AG]	Hypothetical	0744	1, 2
15 C	Hypothetical	1353	1
16 A	<i>fecA</i>	1400	1

* 1, Tomb *et al.* (1997); 2, Saunders *et al.* (1998).† Saunders *et al.* (1998) only.

involvement of slipped strand mispairing (Table 1; Ge *et al.*, 1997; Martin *et al.*, 1997; Tomb *et al.*, 1997; Appelmek *et al.*, 1998). However, it should be noted that the *fucT* gene sequenced by Martin *et al.* (1997) lacks the oligo-C repeat described by Ge *et al.* (1997).

Generation of genomic diversity in *H. pylori*

The generation of genetic diversity within a non-clonal species of bacteria is likely to result from either of two events. Diversity can result from exchange of DNA

within a particular bacterial cell or lineage (intra-strain rearrangements) or it may result from exchange of DNA between different strains or in some rare cases even different species (horizontal transfer).

Intra-strain generation of genomic diversity

A variety of intra-genomic rearrangements (larger rearrangements than frameshifts) may occur in bacteria, namely, deletions, tandem duplications, inversions and translocations within a genome. Many of these rearrangements may occur as a result of the presence of repetitive DNA sequences. The repeated DNA sequences may be indigenous to the species, or may be derived from DNA introduced through horizontal transfer (Bachellier *et al.*, 1996; Lloyd & Low, 1996). The genome sequence of *H. pylori* (Tomb *et al.*, 1997) reveals that *H. pylori* possesses a large number of repetitive DNA sequences. As well as several full-length and partial copies of the IS605 insertion sequence, there are four copies (two full-length and two partial) of a novel insertion element, IS606. Due to the number of copies and their presence throughout the chromosome, these elements may lead to recombinational events resulting in chromosomal rearrangements. The short tandem repeats identified in *H. pylori* may also be associated with genomic rearrangements. Such repeats, like simple repeat DNA of eukaryotes, may be hot spots for intra-genomic rearrangements, such as gene conversion and illegitimate recombination.

Covacci (1996) has proposed the existence of 'hinge regions' (perhaps involving repeat sequences) in the chromosome which facilitate recombination while leaving regulatory and coding sequences intact, thereby permitting the rapid development of the macrodiversity identified by discriminatory fingerprinting techniques. The bacterial nucleoid is compacted *in vivo* with RNA, RNA polymerase components and protein (Trun & Marko, 1998). The propensity of *H. pylori* DNA to rearrange perhaps suggests that the loop domain structure of its DNA may be more highly dynamic than that of other bacteria, thereby facilitating intra- and inter-molecular recombination.

Horizontal gene transfer

The phenomenon of horizontal transfer is likely to have played a significant role in the genetic make-up of *H. pylori*. Analysis of G+C content and sequence data obtained from *H. pylori* prior to 1997 suggested the presence of sequences related to eukaryotes, archaea and other prokaryotes. For example, analysis of the *H. pylori* major sigma factor (RpoD) revealed that the amino acid sequence of RpoD showed strongest similarity to those of Gram-positive bacteria, e.g. *Staphylococcus aureus*, *Listeria monocytogenes* and *Enterococcus faecalis* (Solnick *et al.*, 1997).

This has since been confirmed by the publishing of the entire genome sequence of *H. pylori* strain 26695 and strongly suggests that inter-species horizontal transfer

of genes has occurred during the evolutionary development of *H. pylori*. Many of the proteins described in this Gram-negative bacterium show greatest amino acid sequence similarity to proteins from the different major taxonomic groups. Indeed, Goldman & Kranz (1998) published evidence for the transfer of an entire biosynthetic pathway. The biogenesis of *c*-type cytochromes can be divided into three system types. Type I is found in most Gram-negative bacteria, while type II is characteristic of Gram-positive bacteria, cyanobacteria and chloroplasts. It is the type II system that is found in *H. pylori*. These findings suggest that transfer of genetic information to *H. pylori* has occurred from different phylogenetic groups.

A total of five regions within the *H. pylori* genome have significantly different %G+C contents compared to the mean content of 39 mol% for the rest of the genome (Table 2). Two of these regions contain at least one copy of the IS605 insertion sequence discussed earlier, flanked by a 5S rDNA sequence and a 521 bp repeat. The other regions include the *cag*-PAI, a region associated with β and β' RNA polymerase and elongation factor EF-G, and a region associated with two restriction/modification systems (Tomb *et al.*, 1997).

Mechanisms of horizontal gene transfer

In order for exchange of DNA to occur between strains, DNA must be transferred from one bacterium to another and be taken up by the host cell. There are three main mechanisms by which naturally occurring DNA transfer may occur in bacteria, namely, natural transformation, conjugation and transduction. Despite the electron microscopic evidence for the occurrence of phage in *H. pylori* (Heintschel von Heinegg *et al.*, 1993), transduction has not been demonstrated to occur. Recently Kuipers *et al.* (1998), in a study to define conditions under which *H. pylori* strains could exchange DNA, found evidence for a deoxyribonuclease-resistant conjugation-like mechanism involving cell-to-cell contact which contributed to DNA transfer between *H. pylori* cells. A similar process of deoxyribonuclease-resistant transformation not involving conjugal plasmids has been described for *N. gonorrhoeae* (Catlin, 1981). Natural transformation has also been extensively demonstrated for *H. pylori* and, accordingly, it would appear to be an important mechanism for horizontal gene transfer in *H. pylori* (Nedenskov-Sørensen *et al.*, 1990; Wang *et al.*, 1993; Hofreuter *et al.*, 1998).

Natural transformation and competence

In vitro studies have revealed that most strains of *H. pylori* are naturally competent for DNA uptake (Nedenskov-Sørensen *et al.*, 1990; Wang *et al.*, 1993). It is currently unknown how competence is regulated in *H. pylori*, although a gene cluster involved in the regulation of competence has recently been described (Hofreuter *et al.*, 1998). This cluster consists of four tandemly arranged genes with partially overlapping ORFs (*orf2*, *comB1*, *comB2* and *comB3*) constituting a single trans-

Table 2. Genomic features that predict diversity (adapted from Tomb *et al.*, 1997)

Feature	Associated genes
Distinct G + C regions	
Region 1 (33 mol % G + C)	IS605, 5S rDNA and 521 bp repeat; <i>virB4</i>
Region 2 (35 mol % G + C)	<i>cag</i> -PAI
Region 3 (33 mol % G + C)	IS605, 5S rDNA and 521 bp repeat
Region 4 (43 mol % G + C)	β and β' RNA polymerase, EF-G (<i>fusA</i>)
Region 5 (33 mol % G + C)	Two restriction/modification systems
Insertion sequences	
IS605	<i>tnpA</i> and <i>tnpB</i>
IS606	<i>tnpA</i> and <i>tnpB</i>

riptional unit. This locus was shown to be present by Southern hybridization analysis in all *H. pylori* strains tested but interestingly no homologues were detected in chromosomal DNA from *Helicobacter felis* or *H. mustelae* (Hofreuter *et al.*, 1998).

In several bacterial species the genes involved in natural competence are coordinately expressed under the influence of particular environmental conditions, suggesting that cells may actively respond to conditions in the habitat by adjusting their level of gene acquisition through natural transformation rather than by being passively subjected to environmentally forced gene entry (Lorenz & Wackernagel, 1994). The release of DNA from individual *H. pylori* cells for uptake by other *H. pylori* cells may be facilitated by autolysis (Phadnis *et al.*, 1996). In *Streptococcus pneumoniae* the regulation of competence for genetic transformation depends on a quorum-sensing system (Alloing *et al.*, 1996; Pestova *et al.*, 1996) where a concerted response is initiated by signalling within the bacterial population. Such a mechanism, although not identified by analysis of the genome sequence, may possibly be present in *H. pylori* and may also be autolysis-linked (Phadnis *et al.*, 1996). Hofreuter *et al.* (1998) postulated that the 37-amino-acid-long secretory peptide encoded by *orf2*, the first ORF of the *comB* operon, may have a function in the regulation of competence in *H. pylori*.

In *Haem. influenzae* and *Neisseria meningitidis* the uptake of DNA by transformation is facilitated by the presence of uptake-specific sequences, i.e. a DNA consensus sequence of 29 bp and an invariant 10 bp sequence, respectively (Dempsey *et al.*, 1995; Smith *et al.*, 1995). Numerous copies of these uptake-specific sequences are present in the genomes of these bacteria, e.g. in the *Haem. influenzae* genome 1465 uptake-specific sequences have been identified (Fleischmann *et al.*, 1995; Smith *et al.*, 1995). However, as yet no such uptake-specific sequences have been identified in *H. pylori*. Following uptake of foreign DNA, repeats or repetitive elements often provide the substrate DNA for homologous recombination to occur between DNA molecules (Bachellier *et al.*, 1996).

It is difficult to assess the importance of natural transformation in gene transfer within and between

natural populations due to the absence of adequate *in vitro* models. Heretofore, what data there are may be of questionable relevance to the *in vivo* situation. Analysis of the *H. pylori* genome sequence identified 11 restriction/modification systems which have been presumed to protect the bacterium from integration of foreign DNA. Berg *et al.* (1997a) questioned whether or not these systems are fully functional and whether or not they were acquired before or after the uptake of foreign genetic elements. However, more detailed analysis performed by Saunders *et al.* (1998) has provided evidence that all of these systems are unlikely to be functional. Furthermore, the integration of transformed DNA into the genomes of other naturally competent species, namely, *Bacillus subtilis*, *S. pneumoniae* and *Haemophilus* spp., appears to be relatively unaffected by the restriction systems of the recipient bacteria (Lorenz & Wackernagel, 1994). Sequence divergence is reported to be the most likely barrier to successful transformation with foreign DNA (Matic *et al.*, 1996).

Model for microevolution during infection

Initial colonization of the gastric mucosa by founder isolates of *H. pylori* leads to a persistent infection by the bacterium. These founder colonies may all undergo microevolution of their genome structure. The failure to identify greater than a single strain or subtype within individual biopsy specimens indicates that within discrete sites (foci) within the gastric mucosa, identical clones colonize. The bacteria colonizing different sites within the stomach probably come under slightly different selection pressures, e.g. the pH level within different regions of the stomach is known to vary considerably, and different rearrangements may possibly be selected for by the different environmental conditions encountered at the different sites. *H. pylori* possesses flagella which facilitate its movement in the viscous gastric mucus and may perhaps play a role in the spread of infection within the gastric mucosa. Perhaps biopsies which are taken from adjacent sites are sampling the same focus of infection, whereas biopsies from regions of the gastric mucosa further removed from each other sample different foci of infection and contain different subtypes.

H. pylori lives predominantly in the mucus layer and must cope with the continual washing generated by peristalsis. The infecting bacteria are likely to be under significant selective pressure. In a hostile environment, like the gastric mucosa with low pH and an active immune response, it may be an important adaptation for a bacterium to mutate or rearrange its DNA frequently to generate genetic diversity within the population (Taddei *et al.*, 1997). By continually generating cells with varied phenotypes, mixed populations are likely to exist, some members of which are better suited for particular environmental conditions (Dybvig, 1993). Many questions regarding the modes of genetic variation involved remain to be answered. Why *H. pylori*, which has phenotypic adaptations apparently specific for this habitat, e.g. flagella and urease, needs to generate such a level of genomic diversity and how it tolerates such levels of mutations and rearrangements within its genome are intriguing questions.

Pathogenesis is a complex and multifactorial process involving the coordinated action of a large assembly of virulence-associated loci. Genes associated with short repetitive sequences which encode putative surface-expressed proteins in *H. pylori* may represent examples of contingency loci, where phase variation of these proteins may result in adaptive advantage (Moxon *et al.*, 1994). Horizontal gene transfer of DNA to *H. pylori* is facilitated by the fact that it is naturally competent for DNA uptake (Nedenskov-Sørensen *et al.*, 1990). Mosaic gene structures, such as found in the *vacA* alleles of *H. pylori*, are indicative of horizontal gene exchange and associated recombination (Maynard Smith *et al.*, 1991). Only approximately 10% of individuals in the developed world appear to be colonized with more than one strain, so frequent exposure to transformation of DNA from other strains appears to be unlikely. Therefore, the possibility for intra-species gene exchange appears to be limited. However, in the developing world, where *H. pylori* infection is more common, multiple infections are more frequent (Berg *et al.*, 1997b).

Autolysis has been shown to occur in *H. pylori* (Phadnis *et al.*, 1996), although recent data appear to conflict with this (Vanet & Labigne, 1998). Autolysis is associated with the release of DNA from cells which is then available for transformation. In other competent bacterial species, the autolysis of the bacterial population is correlated with the induction of competence genes (Lorenz & Wackernagel, 1994). Perhaps autolysis is associated with competence in *H. pylori*. Such a system would conform to a modular lifestyle, whereby the success of the individual is not as important as that of the population (Andrews, 1995). In this model, bacteria undergoing autolysis could be conceived as DNA donors, maintaining the genetic diversity and high rates of genome rearrangements to help the fitness of the population as a whole. Further analysis of the transformation and conjugation-like mechanisms of *H. pylori* is required to elucidate their role in DNA transfer and any association with the autolysis phenomenon.

Concluding remarks

H. pylori is a truly fascinating organism with many unique characteristics. The key issue raised by this review is why is this pathogen so genotypically diverse. The level of genomic diversity among *H. pylori* isolates indicates that this organism is an excellent model for studying the mechanisms of natural gene transfer and recombination. As a chronic human pathogen, which most likely causes life-long infections and is associated with several diseases including gastric cancer, further analysis of the macrodiversity of *H. pylori* populations is required as well as comparative genome sequences to unravel the details of its unique chromosomal structure.

Acknowledgements

The work of the authors has been supported by the Forbairt Basic Research Programme (SMB), the Irish American Partnership-Forbairt Scholars Programme (WGD) and the Health Research Board of Ireland (DGM). The authors are grateful to Professor Colm A. Ó Moráin, Department of Gastroenterology, the Meath and Adelaide Hospitals, Dublin, for his valued contribution, research enthusiasm, and financial support and to Jane Twohig for technical support.

References

- Achtman, M. (1994). Clonal spread of serogroup A meningococci: a paradigm for the analysis of microevolution in bacteria. *Mol Microbiol* 11, 15–22.
- Akopyants, N. S., Clifton, S. W., Kersulyte, D. & 7 other authors (1998). Analyses of the *cag* pathogenicity island of *Helicobacter pylori*. *Mol Microbiol* 28, 37–53.
- Akopyanz, N., Bukanov, N. O., Westblom, T. U., Kresovich, S. & Berg, D. E. (1992). DNA diversity among clinical isolates of *Helicobacter pylori* detected by PCR-based RAPD fingerprinting. *Nucleic Acids Res* 20, 5137–5142.
- Alloing, G., Granadel, C., Morrison, D. A. & Claverys, J.-P. (1996). Competence pheromone, oligopeptide permease, and induction of competence in *Streptococcus pneumoniae*. *Mol Microbiol* 21, 471–478.
- Andrews, J. (1995). What if bacteria are modular organisms? *ASM News* 61, 627–632.
- Appelmek, B. J., Simoons-Smit, I., Negrini, R. & 17 other authors (1996). Potential role of molecular mimicry between *Helicobacter pylori* lipopolysaccharide and host Lewis blood group antigens in autoimmunity. *Infect Immun* 64, 2031–2040.
- Appelmek, B. J., Shiberu, B., Trinks, C. & 10 other authors (1998). Phase variation in *Helicobacter pylori* lipopolysaccharide. *Infect Immun* 66, 70–76.
- Aspinall, G. O. & Monteiro, M. A. (1996). Lipopolysaccharides of *Helicobacter pylori* strains P466 and MO19: structures of the O antigen and core oligosaccharide chains. *Biochemistry* 35, 2498–2504.
- Aspinall, G. O. & Moran, A. P. (1997). *Helicobacter pylori* lipopolysaccharide structure and mimicry of Lewis blood group antigens. In *Pathogenesis and Host Response in Helicobacter pylori Infections*, pp. 34–42. Edited by A. P. Moran & C. A. O'Morain. Bad Homburg: Normed Verlag.
- Aspinall, G. O., Monteiro, M. A., Pang, H., Walsh, E. J. & Moran, A. P. (1994). O antigen chains in the lipopolysaccharide of *Helicobacter pylori* NCTC 11637. *Carbohydr Lett* 1, 151–156.

- Atherton, J. C. (1997). The clinical relevance of strain types of *Helicobacter pylori*. *Gut* 40, 701–703.
- Atherton, J. C., Cao, P., Peek, R. M., Tummuru, M. K. R., Blaser, M. J. & Cover, T. L. (1995). Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. *J Biol Chem* 270, 17771–17777.
- Atherton, J., Peek, R. M., Jr, Tham, K. T., Cover, T. L. & Blaser, M. J. (1997). Clinical and pathological importance of heterogeneity in *vacA*, the vacuolating cytotoxin gene of *Helicobacter pylori*. *Gastroenterology* 112, 92–99.
- Bachellier, S., Gilson, E., Hofnung, M. & Hill, C. (1996). Repeated sequences. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, vol. 2, pp. 2588–2612. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Berg, C. M. & Berg, D. E. (1996). Transposable elements: tools for microbial genetics. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, vol. 2, pp. 2012–2040. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Berg, D. E., Hoffmann, P. S., Appelmelk, B. J. & Kusters, J. G. (1997a). The *Helicobacter pylori* genome sequence: genetic factors for long life in the gastric mucosa. *Trends Microbiol* 5, 468–474.
- Berg, D. E., Gilman, R. H., Lelwala-Guruge, J. & 9 other authors (1997b). *Helicobacter pylori* populations in Peruvian patients. *Clin Infect Dis* 25, 996–1002.
- Blaser, M. J. (1989). *Campylobacter pylori in Gastritis and Peptic Ulcer Disease*. New York: Igaku-Shoin Medical Publishers.
- Blaser, M. J. (1996). Genetic basis for heterogeneity of *Helicobacter pylori*. In *Helicobacter pylori: Basic Mechanisms to Clinical Cure* 1996, pp. 33–39. Edited by R. H. Hunt & G. N. J. Tytgat. Dordrecht: Kluwer.
- Blaser, M. J. (1997). Heterogeneity of *Helicobacter pylori*. *Eur J Gastroenterol Hepatol* 9 (suppl. 1), S3–S7.
- Blaser, M. J., Perez-Perez, G. I., Kleanthous, H., Cover, T. L., Peek, R. M., Chyou, P. H., Stemmermann, G. N. & Normura, A. (1995). Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. *Cancer Res* 55, 2111–2115.
- Buckley, M. J. M. & O'Morain, C. A. (1998). *Helicobacter* biology – discovery. *Br Med Bull* 54, 7–16.
- Bukanov, N. O. & Berg, D. E. (1994). Ordered cosmid library and high-resolution physical-genetic map of *Helicobacter pylori* strain NCTC11638. *Mol Microbiol* 11, 509–523.
- Calam, J. (1996). *Clinician's Guide to Helicobacter pylori*. London: Chapman & Hall Medical.
- Cao, P. & Cover, T. L. (1997). High-level genetic diversity in the *vapD* chromosomal region of *Helicobacter pylori*. *J Bacteriol* 179, 2852–2856.
- Catlin, B. W. (1981). Cell-to-cell transmission of chromosomal loci in *Neisseria gonorrhoeae*. In *Genetic Exchange: a Celebration and a New Generation, Proceedings of the 25th Wind River Conference*, pp. 310–325. Edited by U. Strieps. New York: Marcel Dekker.
- Censini, S., Lange, C., Xiang, Z., Crabtree, J. E., Ghiara, P., Borodovsky, M., Rappuoli, R. & Covacci, A. (1996). *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proc Natl Acad Sci USA* 93, 14648–14653.
- Covacci, A. (1996). *Mobilis in mobile*: unexpected flexibility and quantum leaps in the *Helicobacter pylori* genome. In *Helicobacter pylori: Basic Mechanisms to Clinical Cure* 1996, pp. 40–49. Edited by R. H. Hunt & G. N. J. Tytgat. Dordrecht: Kluwer.
- Covacci, A. & Rappuoli, R. (1998). *Helicobacter pylori*: molecular evolution of a bacterial quasi-species. *Curr Opin Microbiol* 1, 96–102.
- Covacci, A., Censini, S., Bugnoli, M. & 8 other authors (1993). Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. *Proc Natl Acad Sci USA* 90, 5791–5795.
- Covacci, A., Falkow, S., Berg, D. E. & Rappuoli, R. (1997a). Did the inheritance of a pathogenicity island modify the virulence of *Helicobacter pylori*? *Trends Microbiol* 5, 205–208.
- Covacci, A., Falkow, S., Censini, S. & 8 other authors (1997b). The *cag* pathogenicity island of *Helicobacter pylori*: origin, molecular biology and relevance to virulence. In *Pathogenesis and Host Response in Helicobacter pylori Infections*, pp. 88–100. Edited by A. P. Moran & C. A. O'Morain. Bad Homburg: Normed Verlag.
- Cover, T. L. & Blaser, M. J. (1992). Purification and characterization of the vacuolating toxin from *Helicobacter pylori*. *J Biol Chem* 267, 10570–10575.
- Cover, T. L., Tummuru, M. K., Cao, P., Thompson, S. A. & Blaser, M. J. (1994). Divergence of genetic sequences for the vacuolating cytotoxin among *Helicobacter pylori* strains. *J Biol Chem* 269, 10566–10577.
- Crabtree, J. E., Xiang, Z., Lindley, I. J., Tompkins, D. S., Rappuoli, R. & Covacci, A. (1995). Induction of interleukin-8 secretion from gastric epithelial cells by a *cagA* negative isogenic mutant of *Helicobacter pylori*. *J Clin Pathol* 48, 967–969.
- Dempsey, J. A. F., Wallace, A. C. & Cannon, J. G. (1995). The physical map of the chromosome of a serotype A strain of *Neisseria meningitidis* shows complex rearrangements relative to the chromosome of the two mapped strains of the closely related species *N. gonorrhoeae*. *J Bacteriol* 177, 6390–6400.
- van Doorn, L. J., Figueiredo, C., Rossau, R., Jannes, G., van Asbroeck, M., Sousa, J. C., Carneiro, F. & Quint, W. G. V. (1998a). Typing of *Helicobacter pylori vacA* gene and detection of *cagA* gene by PCR and reverse hybridization. *J Clin Microbiol* 36, 1271–1276.
- van Doorn, L. J., Figueiredo, C., Sanna, R., Plaisier, A., Schneeberger, P., de Boer, W. & Quint, W. G. V. (1998b). Clinical relevance of the *cagA*, *vacA*, and *iceA* status of *Helicobacter pylori*. *Gastroenterology* 115, 58–66.
- Dundon, W. G., Marshall, D. G., O'Morain, C. A. & Smyth, C. J. (1998). The identification and characterization of a differentially expressed locus in *Helicobacter pylori* using arbitrarily primed PCR of RNA. *Gastroenterology* 114, abstract G3970, A968.
- Dunn, B. E., Cohen, H. & Blaser, M. J. (1997). *Helicobacter pylori*. *Clin Microbiol Rev* 10, 720–741.
- Dybvig, K. (1993). DNA rearrangements and phenotypic switching in prokaryotes. *Mol Microbiol* 10, 465–471.
- Eckloff, B. W., Podzorski, R. P., Kline, B. C. & Cockerill, F. R., III (1994). A comparison of 16S ribosomal DNA sequences from five isolates of *Helicobacter pylori*. *Int J Syst Bacteriol* 44, 320–323.
- van der Ende, A., Rauws, E. A., Feller, M., Mulder, C. J., Tytgat, G. N. & Dankert, J. (1996). Heterogeneous *Helicobacter pylori* isolates from members of a family with a history of peptic ulcer disease. *Gastroenterology* 111, 638–647.
- Farthing, M. J. G. & Patchett, S. E. (1998). *Helicobacter Infection*. London: The Royal Society of Medicine.
- Fleischmann, R. D., Adams, M. D., White, O. & 37 other authors (1995). Whole-genome random sequencing and assembly of *Haemophilus Rd*. *Science* 269, 496–512.

- Forbes, K. J., Fang, Z. & Pennington, T. H. (1995). Allelic variation of the *Helicobacter pylori* flagellin genes *flaA* and *flaB*: its consequences for strain typing schemes and population structures. *Epidemiol Infect* 114, 257–266.
- Forsyth, M. H., Atherton, J. C., Blaser, M. J. & Cover, T. L. (1998). Heterogeneity in levels of vacuolating cytotoxin gene (*vacA*) transcription among *Helicobacter pylori* strains. *Infect Immun* 66, 3088–3094.
- Foxall, P. A., Hu, L.-T. & Mobley, H. L. T. (1992). Use of polymerase chain reaction-amplified *Helicobacter pylori* urease structural genes for differentiation of isolates. *J Clin Microbiol* 30, 739–741.
- Fujimoto, S., Marshall, B. & Blaser, M. J. (1994). PCR-based restriction fragment length polymorphism typing of *Helicobacter pylori*. *J Clin Microbiol* 32, 331–334.
- Gálvez, A., Maqueda, M., Martínez-Bueno, M. & Valdivia, E. (1998). Publication rates reveal trends in microbiological research. *ASM News* 64, 269–275.
- Ge, M. F., Chan, N. W. C., Palcic, M. & Taylor, D. E. (1997). Cloning and heterologous expression of an α 1,3-fucosyltransferase gene from the gastric pathogen *Helicobacter pylori*. *J Biol Chem* 272, 21357–21363.
- Go, M. F., Chan, Y. K., Yersalovic, J., Koeuth, T., Graham, D. Y. & Lupski, J. R. (1995). Cluster analysis of *Helicobacter pylori* genomic DNA fingerprints suggests gastroduodenal disease-specific associations. *Scand J Gastroenterol* 30, 640–646.
- Go, M. F., Kapur, V., Graham, D. Y. & Musser, J. M. (1996). Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. *J Bacteriol* 178, 3934–3938.
- Goldman, B. & Kranz, R. (1998). Evolution and horizontal transfer of an entire biosynthetic pathway for cytochrome *c* biogenesis: *Helicobacter*, *Deinococcus*, *Archae* and more. *Mol Microbiol* 27, 871–874.
- Goodwin, C. S. & Worsley, B. W. (1993). *Helicobacter pylori: Biology and Clinical Practice*. Boca Raton, FL: CRC Press.
- Hacker, J., Blum-Oehler, G., Mühldorfer, I. & Tschäpe, H. (1997). Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol* 23, 1089–1097.
- Hancock, R. E. W., Alm, R., Bina, J. & Trust, T. (1998). *Helicobacter pylori*: a surprisingly conserved bacterium. *Nat Biotechnol* 16, 216–218.
- Hazell, S. L., Andrews, R. H., Mitchell, H. M. & Daskalopoulos, G. (1997). Genetic relationship among isolates of *Helicobacter pylori*: evidence for the existence of a *Helicobacter pylori* species complex. *FEMS Microbiol Lett* 150, 27–32.
- Heintschel von Heinegg, E., Nalik, H. P. & Schmid, E. N. (1993). Characterisation of *Helicobacter pylori* phage (HP1). *J Med Microbiol* 38, 245–249.
- Hobbs, M. M., Seiler, A., Achtman, M. & Cannon, J. G. (1994). Microevolution within a clonal population of pathogenic bacteria: recombination, gene duplication and horizontal genetic exchange in the *opa* gene family of *Neisseria meningitidis*. *Mol Microbiol* 12, 171–180.
- Hofreuter, D., Odenbreit, S., Henke, G. & Hass, R. (1998). Natural competence for DNA transformation in *Helicobacter pylori*: identification and genetic characterization of the *comB* locus. *Mol Microbiol* 28, 1027–1038.
- Hood, D. W., Deadman, M. E., Jennings, M. P., Bisercic, M., Fleischmann, R. D., Venter, J. C. & Moxon, E. R. (1996). DNA repeats identify novel virulence genes in *Haemophilus influenzae*. *Proc Natl Acad Sci USA* 93, 11121–11125.
- Höök-Nikanne, J., Berg, D. E., Peek, R. M., Jr, Kersulyte, D., Tummuru, M. K. R. & Blaser, M. J. (1998). DNA sequence conservation and diversity in transposable element IS606 of *Helicobacter pylori*. *Helicobacter* 3, 79–85.
- Hunt, R. H. & Tytgat, G. N. J. (1996). *Helicobacter pylori: Basic Mechanisms to Clinical Cure* 1996. Dordrecht: Kluwer.
- Ilver, D., Arnquist, A., Ögren, J. & 7 other authors (1998). *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* 279, 373–377.
- Ito, Y., Azuma, T., Ito, S. & 7 other authors (1997). Analysis and typing of the *vacA* gene from *cagA*-positive strains of *Helicobacter pylori* isolated in Japan. *J Clin Microbiol* 35, 1710–1714.
- Jiang, Q., Hiratsuka, K. & Taylor, D. E. (1996). Variability of gene order in different *Helicobacter pylori* strains contributes to genome diversity. *Mol Microbiol* 20, 833–842.
- Jorgensen, M., Daskalopoulos, G., Warburton, C., Mitchell, H. M. & Hazell, S. L. (1996). Multiple strain colonization and metronidazole resistance in *Helicobacter pylori*-infected patients: identification from sequential and multiple biopsy specimens. *J Infect Dis* 174, 631–635.
- Josenshans, C. & Suerbaum, S. (1997). Flagella and motility of *Helicobacter pylori*. In *Pathogenesis and Host Response in Helicobacter pylori Infections*, pp. 6–15. Edited by A. P. Moran & C. A. O'Morain. Bad Homburg: Normed Verlag.
- Kleanthous, H., Clayton, C. L. & Tabaqchali, S. (1991). Characterisation of a plasmid from *Helicobacter pylori* encoding a replication protein common to plasmids in Gram-positive bacteria. *Mol Microbiol* 5, 2377–2389.
- Krawiec, S. & Riley, M. (1990). Organization of the bacterial chromosome. *Microbiol Rev* 54, 502–539.
- Kuipers, E. J. & Blaser, B. J. (1998). *Helicobacter pylori* and gastroduodenal disorders. In *Emerging Pathogens*, pp. 191–206. Edited by W. M. Scheld, D. Armstrong & J. M. Hughes. Washington, DC: American Society for Microbiology.
- Kuipers, E. J., Israel, D. A., Kusters, J. G. & Blaser, B. J. (1998). Evidence for a conjugation-like mechanism of DNA transfer in *Helicobacter pylori*. *J Bacteriol* 180, 2901–2905.
- Langenberg, W., Rauws, E. A. J., Widjojokusumo, A., Tytgat, G. N. J. & Zanen, H. C. (1986). Identification of *Campylobacter pyloridis* isolates by restriction endonuclease DNA analysis. *J Clin Microbiol* 24, 414–417.
- Lee, A. & Hazell, S. L. (1988). *Campylobacter pylori* in health and disease: an ecological perspective. *Microb Ecol Health Dis* 1, 1–16.
- Lee, A. & Mégraud, F. (1996). *Helicobacter pylori: Techniques for Clinical Diagnosis and Basic Research*. London: W. B. Saunders.
- Levinson, G. & Gutman, G. A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol* 4, 203–221.
- Lloyd, R. G. & Low, K. B. (1996). Homologous recombination. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, vol. 2, pp. 2236–2255. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Logan, R. P. H. & Berg, D. E. (1996). Genetic diversity of *Helicobacter pylori*. *Lancet* 348, 1462–1463.
- Lorenz, M. G. & Wackernagel, W. (1994). Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol Rev* 58, 563–602.
- Maeda, S., Ogura, K., Yoshida, H., Funai, F., Ikenoue, T., Kato, N., Shiratori, Y. & Omata, M. (1998). Major virulence factors, *VacA* and *CagA*, are commonly positive in *Helicobacter pylori* isolates in Japan. *Gut* 42, 338–343.

- Malfertheiner, P. & Ditschuneit, H. (1990). *Helicobacter pylori*, Gastritis and Peptic Ulcer. Berlin: Springer.
- Marshall, B. J. (1989). History of the discovery of *C. pylori*. In *Campylobacter pylori in Gastritis and Peptic Ulcer Disease*, pp. 7–23. Edited by M. J. Blaser. New York: Igaku-Skoin Medical Publishers.
- Marshall, B. J. & Warren, J. R. (1984). Unidentified curved bacilli in the stomachs of patients with gastritis and peptic ulceration. *Lancet* i, 1311–1315.
- Marshall, B. J., McCallum, R. W. & Guerrant, R. L. (1991). *Helicobacter pylori in Peptic Ulceration and Gastritis*. Oxford: Blackwell Scientific Publications.
- Marshall, D. G. (1996). *Population analysis of Helicobacter pylori*. PhD thesis, Trinity College, University of Dublin.
- Marshall, D. G., Chua, A., Keeling, P. W. N., Sullivan, D. J., Coleman, D. C. & Smyth, C. J. (1995). Molecular analysis of *Helicobacter pylori* populations in antral biopsies from individual patients using randomly amplified polymorphic DNA (RAPD) fingerprinting. *FEMS Immunol Med Microbiol* 10, 317–324.
- Marshall, D. G., Coleman, D. C., Sullivan, D. J., Xia, H., O'Morain, C. A. & Smyth, C. J. (1996). Genomic DNA fingerprinting of clinical isolates of *Helicobacter pylori* using oligonucleotide probes containing repetitive sequences. *J Appl Bacteriol* 81, 509–517.
- Martin, S. L., Edbrooke, M. R., Hodgman, T. C., van den Eijnden, D. H. & Bird, M. I. (1997). Lewis x biosynthesis in *Helicobacter pylori*: molecular cloning of an α -(1,3)-fucosyltransferase gene. *J Biol Chem* 272, 21349–21356.
- Matic, I., Taddei, F. & Radman, M. (1996). Genetic barriers among bacteria. *Trends Microbiol* 4, 69–73.
- Maynard Smith, J. (1995). Do bacteria have population genetics? In *Population Genetics of Bacteria*, pp. 1–12. Edited by S. Baumberg, J. P. W. Young, E. M. H. Wellington & J. R. Saunders. Cambridge: Cambridge University Press.
- Maynard Smith, J., Dowson, C. G. & Spratt, B. G. (1991). Localized sex in bacteria. *Nature* 349, 29–31.
- Miehlke, S., Kibler, K., Kim, J. G., Figura, N., Small, S. M., Graham, D. Y. & Go, M. F. (1996). Allelic variation in the *cagA* gene of *Helicobacter pylori* obtained from Korea compared to the United States. *Am J Gastroenterol* 91, 1322–1325.
- Moble, H. L. T., Fulkerson, J., Jr, Hendricks, J. K. & Garner, R. M. (1997). Expression of catalytically active urease by *Helicobacter pylori*. In *Pathogenesis and Host Response in Helicobacter pylori Infections*, pp. 58–66. Edited by A. P. Moran & C. A. O'Morain. Bad Homburg: Normed Verlag.
- Moran, A. P. & O'Morain, C. A. (1997). *Pathogenesis and Host Response in Helicobacter pylori Infections*. Bad Homburg: Normed Verlag.
- Moxon, E. R., Rainey, P. B., Nowak, M. A. & Lenski, R. E. (1994). Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr Biol* 4, 24–33.
- Nedenskov-Sørensen, P., Bukholm, G. & Bøvre, K. (1990). Natural competence for genetic transformation in *Campylobacter pylori*. *J Infect Dis* 161, 365–366.
- Negrini, R., Savio, M., Graffeo, M., Rolfi, F. & Ghielmi, S. (1993). Auto-antibodies and gastric *Helicobacter pylori* infection: does autoimmunity affect progression to atrophic gastritis? *Eur J Gastroenterol Hepatol* 5 (suppl. 2), S27–S29.
- NIH Consensus Development Panel on *Helicobacter pylori* in Peptic Ulcer Disease (1994). *Helicobacter pylori* in peptic ulcer disease. *J Am Med Assoc* 272, 65–69.
- Nwokolo, C. U., Bickley, J., Attard, A. R., Owen, R. J., Costas, M. & Fraser, I. A. (1992). Evidence of clonal variants of *Helicobacter pylori* in three generations of a duodenal ulcer disease family. *Gut* 33, 1323–1327.
- Ohta-Tada, U., Takagi, A., Koga, Y., Kamiya, S. & Miwa, T. (1997). Flagellin gene diversity among *Helicobacter pylori* strains and IL-8 secretion from gastric epithelial cells. *Scand J Gastroenterol* 32, 455–459.
- Ó'Moráin, C. A. & O'Connor, H. (1994). *Helicobacter pylori: Implications and Practice*. Bad Homburg: Normed Verlag.
- Oudbier, J. H., Langenberg, W., Rauws, E. A. J. & Bruin-Mosch, C. (1990). Genotypic variation of *Campylobacter pylori* from gastric mucosa. *J Clin Microbiol* 28, 559–565.
- Owen, R. J., Hunton, C., Bickley, J., Moreno, M. & Linton, D. (1992). Ribosomal RNA gene restriction patterns of *Helicobacter pylori*: analysis and appraisal of *HaeIII* digests as a molecular typing system. *Epidemiol Infect* 109, 35–47.
- Pan, Z.-J., van der Hulst, R. W. M., Feller, M., Xiao, S.-D., Tytgat, G. N. J., Dankert, J. & van der Ende, A. (1997). Equally high prevalences of infection with *cagA*-positive *Helicobacter pylori* in Chinese patients with peptic ulcer disease and those with chronic gastritis-associated dyspepsia. *J Clin Microbiol* 35, 1344–1347.
- Pan, Z.-J., Berg, D. E., van der Hulst, R. W. M., Su, W.-W., Raudonikienė, A., Xiao, S.-D., Dankert, J., Tytgat, G. N. J. & van der Ende, A. (1998). Presence of vacuolating cytotoxin production and distribution of distinct *vacA* alleles in *Helicobacter pylori* from China. *J Infect Dis* 178, 220–226.
- Peek, R. M., Jr, Thompson, S. A., Atherton, J. C., Blaser, M. J. & Miller, G. G. (1996). Expression of *iceA*, a novel ulcer-associated *H. pylori* gene, is induced by contact with gastric epithelial cells and is associated with enhanced mucosal IL-8. *Gut* 39 (suppl. 2), A71.
- Pérez-Pérez, G. I. & Parsonnet, J. (1994). Epidemiology of *Helicobacter pylori* infections. In *Helicobacter pylori: Implications and Practice*, pp. 52–57. Edited by C. A. O'Moráin & H. O'Connor. Bad Homburg: Normed Verlag.
- Pestova, E. V., Håvarstein, L. S. & Morrison, D. A. (1996). Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. *Mol Microbiol* 21, 853–862.
- Phadnis, S. H., Parlow, M. H., Levy, L., Ilver, D., Caulkins, C. M., Connors, J. B. & Dunn, B. E. (1996). Surface localization of *Helicobacter pylori* urease and a heat shock protein homolog requires bacterial autolysis. *Infect Immun* 64, 905–912.
- Rathbone, B. J. & Heatley, R. V. (1989a). *Campylobacter pylori and Gastrointestinal Disease*. Oxford: Blackwell Scientific Publications.
- Rathbone, B. J. & Heatley, R. V. (1989b). The historical associations between bacteria and peptic-ulcer disease. In *Campylobacter pylori and Gastrointestinal Disease*, pp. 1–4. Oxford: Blackwell Scientific Publications.
- Riccardi, V. M. & Rotter, J. I. (1994). Familial *Helicobacter pylori* infection. Societal factors, human genetics, and bacterial genetics. *Ann Intern Med* 120, 1043–1045.
- Rudi, J., Kolb, C., Maiwald, M., Kuck, D., Sieg, A., Galle, P. R. & Stremmel, W. (1998). Diversity of *Helicobacter pylori vacA* and *cagA* genes and relationship to *VacA* and *CagA* expression, cytotoxin production, and associated diseases. *J Clin Microbiol* 36, 944–948.
- Saunders, K. E., McGovern, K. J. & Fox, J. G. (1997). Use of pulse-field gel electrophoresis to determine genomic diversity in strains

- of *Helicobacter hepaticus* from geographically distant locations. *J Clin Microbiol* 35, 2859–2863.
- Saunders, N. J., Peden, J. F., Hood, D. W. & Moxon, E. R. (1998). Simple sequence repeats in the *Helicobacter pylori* genome. *Mol Microbiol* 27, 1091–1098.
- Shimoyama, T., Fukuda, S., Tanaka, M., Mikami, T., Saito, Y. & Munakata, A. (1997). High prevalence of CagA-positive *Helicobacter pylori* strains in Japanese asymptomatic patients and gastric cancer patients. *Scand J Gastroenterol* 32, 465–468.
- Shortridge, V. D., Stone, G. G., Flamm, R. K., Beyer, J., Versalovic, J., Graham, D. W. & Tanaka, S. K. (1997). Molecular typing of *Helicobacter pylori* isolates from a multicenter US clinical trial by *ureC* restriction fragment length polymorphism. *J Clin Microbiol* 35, 471–473.
- Simor, A. E., Shames, B., Drumm, B., Sherman, P., Low, D. E. & Penner, J. L. (1990). Typing of *Campylobacter pylori* by bacterial DNA restriction endonuclease analysis and determination of plasmid profile. *J Clin Microbiol* 28, 83–86.
- Smith, H. O., Tomb, J.-F., Dougherty, B. A., Fleischmann, R. D. & Venter, J. C. (1995). Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. *Science* 269, 538–540.
- Solnick, J. V., Hansen, L. M. & Syvanen, M. (1997). The major sigma factor (RpoD) from *Helicobacter pylori* and other Gram-negative bacteria shows an enhanced rate of divergence. *J Bacteriol* 179, 6196–6200.
- Strobel, S., Bereswill, S., Balig, P., Allgaier, P., Sonntag, H.-G. & Kist, M. (1998). Identification and analysis of a new *vacA* genotype variant of *Helicobacter pylori* in different patient groups in Germany. *J Clin Microbiol* 36, 1285–1289.
- Suerbaum, S. (1995). The complex flagella of gastric *Helicobacter* species. *Trends Microbiol* 3, 168–170.
- Taddei, F., Matic, I., Godelle, B. & Radman, M. (1997). To be a mutator, or how pathogenic and commensal bacteria can evolve rapidly. *Trends Microbiol* 5, 427–429.
- Takami, S., Hayashi, T., Tonokatsu, Y., Shimoyama, T. & Tamura, T. (1993). Chromosomal heterogeneity of *Helicobacter pylori* by pulsed-field gel electrophoresis. *Int J Med Microbiol Virol Parasitol Infect Dis* 280, 120–127.
- Takami, S., Hayashi, T., Akashi, H., Shimoyama, T. & Tamura, T. (1994). Genetic heterogeneity of *Helicobacter pylori* by pulse-field electrophoresis and re-evaluation of DNA homology. *Eur J Gastroenterol Hepatol* 6 (suppl. 1), S53–S56.
- Taylor, D. E. & Blaser, M. J. (1991). The epidemiology of *Helicobacter pylori* infection. *Epidemiol Rev* 13, 42–59.
- Taylor, D. E., Eaton, M., Chang, N. & Salama, S. M. (1992). Construction of a *Helicobacter pylori* genome map and demonstration of diversity at the genome level. *J Bacteriol* 174, 6800–6806.
- Taylor, D. E., Chang, N., Taylor, N. S. & Fox, J. G. (1994). Genome conservation in *Helicobacter mustelae* as determined by pulse-field gel electrophoresis. *FEMS Microbiol Lett* 118, 31–36.
- Taylor, D. N. & Parsonnet, J. (1995). Epidemiology and natural history of *H. pylori* infections. In *Infections of the Gastrointestinal Tract*, pp. 551–564. Edited by M. J. Blaser, P. F. Smith, J. Ravdin, H. Greenberg & R. L. Guerrant. New York: Raven Press.
- Taylor, N. S., Fox, J. G., Akopyants, N. S. & 8 other authors (1995). Long-term colonization with single and multiple strains of *Helicobacter pylori* assessed by DNA fingerprinting. *J Clin Microbiol* 33, 918–923.
- Tomb, J.-F., White, O., Kerlavage, A. R. & 39 other authors (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388, 539–547.
- Trun, N. J. & Marko, J. F. (1998). Architecture of a bacterial chromosome. *ASM News* 64, 276–283.
- Tummuru, M. K., Cover, T. L. & Blaser, M. J. (1993). Cloning and expression of a high molecular-mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production. *Infect Immun* 61, 1799–1809.
- Tummuru, M. K., Cover, T. L. & Blaser, M. J. (1994). Mutation of the cytotoxin-associated *cagA* gene does not affect the vacuolating cytotoxin activity of *Helicobacter pylori*. *Infect Immun* 62, 2609–2613.
- Vanet, A. & Labigne, A. (1998). Evidence for specific secretion rather than autolysis in the release of some *Helicobacter pylori* proteins. *Infect Immun* 66, 1023–1027.
- Wang, H.-J., Kuo, C.-H., Yeh, A. A. M., Chang, P. C. L. & Wang, W.-C. (1998). Vacuolating toxin production in clinical isolates of *Helicobacter pylori* with different *vacA* genotypes. *J Infect Dis* 178, 207–212.
- Wang, Y., Roos, K. P. & Taylor, D. E. (1993). Transformation of *Helicobacter pylori* by chromosomal metronidazole resistance and by a plasmid with a selectable chloramphenicol resistance marker. *J Gen Microbiol* 139, 2485–2493.
- Warren, J. R. & Marshall, B. J. (1983). Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet* i, 1273–1275.
- Xiang, Z., Censini, S., Bayeli, P. F., Telford, J. L., Figura, N., Rappuoli, R. & Covacci, A. (1995). Analysis and expression of CagA and VacA virulence factors in 43 strains of *Helicobacter pylori* reveals that clinical isolates can be divided into two major types and that CagA is not necessary for expression of the vacuolating cytotoxin. *Infect Immun* 63, 94–98.
- Xu, Q., Peek, R. M., Jr, Miller, G. G. & Blaser, M. J. (1997). The *Helicobacter pylori* genome is modified at CATG by the product of *hpyLM*. *J Bacteriol* 179, 6807–6815.
- Yamaoka, Y., Kodama, T., Kashima, K., Graham, D. Y. & 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. *J Clin Microbiol* 36, 2258–2263.
- Yoshimura, H. H., Evans, D. G. & Graham, D. Y. (1993). DNA–DNA hybridization demonstrates apparent genetic differences between *Helicobacter pylori* from patients with duodenal ulcer and asymptomatic gastritis. *Dig Dis Sci* 38, 1128–1131.