Coevolution and Adaptation of *Helicobacter pylori* and the Case for ‘Functional Molecular Infection Epidemiology’

Niyaz Ahmed\(^a, b, c\)

\(^a\)Pathogen Biology Laboratory, Department of Biotechnology, University of Hyderabad, and \(^b\)Institute of Life Sciences, Hyderabad, India; \(^c\)Institute of Biological Sciences, Universiti Malaya, Kuala Lumpur, Malaysia

**Key Words**

Chronic human pathogens • *Helicobacter pylori* • Coevolution • Molecular epidemiology • Chronological evolution • Replicate genomics

**Abstract**

*Helicobacter pylori* is a major human pathogen and its transmission and epidemiology have been extensively studied; it has been found that *H. pylori*’s prevalence and infection outcome is characterized by marked differences between the developing and the developed worlds. Recent data on genomic analyses and comparative core genome haplotyping have revealed that *H. pylori* has coevolved with its human host. While several studies advocate the protective effects of *H. pylori* colonization, it is prudent to systematically unleash the role of the strong virulence apparatus present within most *H. pylori* strains and to determine how to disarm them (or protect the host from the effects) if the intent is to allow it to remain a friendly organism or to use it as a vaccine delivery tool. While genotyping and phenotyping based on a few genetic markers have not provided much insight into such issues, use of replicate/chronological genomics (of virulent versus innocuous strains) coupled with functional screens in animal models is expected to be able to explain the acquisition and evolution of virulence factors of *H. pylori* and their discreet associations with serious clinical outcomes such as gastric cancer.

**Helicobacter pylori Epidemiology in Developing and Developed Countries**

The plausible role of *H. pylori* in gastroduodenal pathologies including cancer was not confirmed until the end of the 20th century. Subsequently, its epidemiology has been the subject of extensive studies in both the developing and the developed world, and an ever-growing body of scientific literature suggests that not all humans are equally at risk of infection by this pathogen [1].

In the last several years, many studies have forecasted mass migrations due to the consequences of climate change, whereby millions of people would perhaps flee from rising sea levels, floods, disease outbreaks and drought, leading to serious consequences for both the migrant and receiving societies [2, 3]. Many enteric infectious outbreaks occur in the aftermath of natural calam-
ities, mainly due to contamination or shortage of drinking water, such as occur in the summer months in many parts of India, wherein populations are forced to drink from unconventional or unreliable sources that might be contaminated. In the event of migrations, *H. pylori* also comigrate with their human host. Population dynamics and disease potentials of these pathogens then would change with the change in history, geography and ecology of their hosts. *H. pylori* is thus one of the prominent candidates whose epidemiology and evolution within different stationary and migratory communities will be of significant interest when the change of climate and lifestyle and their direct or indirect impact on enteric infections emerge as important concerns in recent times. Chronic *H. pylori* infection has already been described to be a predisposing factor for other enteric infections such as cholera [4], which results in heavy mortality upon spread as epidemic.

Khalifa et al. [1] have recently reviewed the different epidemiological aspects of *H. pylori* infection with emphasis on factors related to human poverty. This author agrees with this notion because the epidemiology of *H. pylori* infection in developing countries is very different from that in developed countries; the difference is more discernable when we take into account the incidence in children. Additionally, there is evidence to suggest that the societal and economic factors, in combination with lifestyle, could be the chief determinants of age-related acquisition rate of *H. pylori*, and consequently, its prevalence. These data appear quite alarming when we take into account the elements of global climate change and birth rate, which are supposed to alter the demography, ecology and sociobiology of our planet, thus putting more birth rate, which are supposed to alter the demography, into account the elements of global climate change and ecology of their hosts. *H. pylori* is thus one of the prominent candidates whose epidemiology and evolution within different stationary and migratory communities will be of significant interest when the change of climate and lifestyle and their direct or indirect impact on enteric infections emerge as important concerns in recent times. Chronic *H. pylori* infection has already been described to be a predisposing factor for other enteric infections such as cholera [4], which results in heavy mortality upon spread as epidemic.

Due to an unequal population growth and uneven economic growth, coupled with the rise in life expectancy all over the world, one should expect a higher incidence of *H. pylori* infection in children, who will mostly inhabit the developing countries in the next decade [5]. This augmented incidence in children is likely to lead to a rise rather than a decline of the worldwide prevalence of *H. pylori*. As a consequence, the elderly population of the developing countries shall be at higher risk of gastric cancer [1].

*H. pylori* incidence is in sharp decline in most of the developed nations of the west due to the success of eradication therapy and high standards of personal and community hygiene. However, the eradication of *H. pylori* has come with a price – a rise in the incidence of gastroesophageal reflux disease, esophageal cancer and childhood diarrheas (reviewed elsewhere) [6]. Due to better eradication compliance and outcome in these countries, *H. pylori* infection is largely confined to immigrant populations and is therefore not seen as a major health concern.

In a semi-developed setting such as India, *H. pylori* colonization rates, although reaching almost 100%, do not correlate with serious outcomes of infection. About 10% of those infected and showing dyspeptic symptoms are diagnosed with duodenal ulcers while the gastric cancer incidence has been nearly 3–4 per 100,000 of the population, which is almost negligible when we compare it with its incidence in Japan and other Asian countries [7]. In India, there seems to be a difference in the distribution frequency of gastric cancer incidence. Clearly, there seems to be some protective advantage [7] which must be learned and reproduced elsewhere. If it is food, then the adaptation will be easy. However, if it is genetic, then it will be difficult to replicate the effect.

**Coevolution, Expansion and Acquisition of Virulence**

The population structure of bacteria that have been determined to be coevolved with their human hosts is almost always similar to the genetic distribution patterns of their host. DNA analysis of humans in recent times revealed that the populations that fall along the major land routes leading out of Africa have become genetically isolated [8], which means that the farther they are from Eastern Africa, the more genetically heterogeneous they are when compared to other human populations; this could be explained based on a phenomenon called 'serial founder effect' where only a fraction of an ancestral population expands further to give birth to a new population. Due to this, in a general sense, the overall genetic diversity of a founder population is restrained in a stepwise manner commensurate to the distance it has traversed from its original habitat.

Similar observations made for some of the European populations revealed gradual differences analogous to the migration of Neolithic farmers northwards [8]. Comprehensive genetic analysis of *H. pylori* found almost exactly the same genetic distribution patterns for this bacterial parasite [9]. Genetic and computer simulation analyses incorporating human and bacterial data extend commendable support to the hypothesis that *H. pylori* may have started its journey from Eastern Africa at about the same time as did early humans, that is, approximately 60,000 years ago [9]. This observation leads to perhaps
the most authentic claim that humans and this bacterium have been intimately associated with each other for at least the last 60,000 years. Given this, the puzzle that remains unsolved is whether this 60,000-year-old, co-evolved H. pylori was as successful a pathogen as exists today. In other words, it is not sufficiently evident if H. pylori originally was equipped with its virulence genes or if it acquired them later during its traverse through various subcontinents and during its mix-up or encounter with other environmental organisms (occurring as a result of gradual change in human history and ecology) [2, 10].

Adoption of agrarian practices by early humans comprising their domestication of animals and pets and growing of crops leads us to think about a scenario which probably explains the acquisition of virulence genes by H. pylori. It was recently hypothesized that early microbial communities associated with crop plants or rodent pests or other animals etc. coinhabiting the early human dwellings may have contributed some of the virulence gene cassettes through horizontal genetic exchanges [2]. There are indirect hints of this possibility due to the presence within the H. pylori chromosome of some of the remnants/homologues of plant pathogen genes (vir genes of Agrobacterium tumifaciens for example) and the genes of environmental bacteria (Aeromonas spp.), as reviewed previously [2]. Also, it is possible that the import of virulence genes in the manner described above may have conferred the needed survival advantage upon H. pylori with respect to achieving fitness in different human and animal hosts. Consequently, the pathogen may have spread selectively in a geographically compartmentalized manner. The genetic structure of today’s H. pylori is highly geographically oriented, both with respect to the core genome [9] and the variome [10]. Many of the putative virulence genes supposedly of foreign origin in H. pylori have been described to be prevalent in a strain-specific manner.

Virulence Factors, Their Plasticity, Role in Imparting Survival Advantage and Use in Diagnostics and Epidemiology

H. pylori infection leads to a spectrum of gastric and duodenal symptoms. Among these, chronic gastritis is a definitive risk factor for the development of gastric adenocarcinoma. Traditionally, H. pylori’s virulence attributes and pathological presentations have been examined in association with the genetic integrity or diversity of the two chief virulence factors, namely CagA and VacA. However, it was soon discovered that genetic statuses of CagA and VacA do not always correlate with particular outcomes of infection [11, 12]. This dilemma could at best be understood through the perception of a complex interplay of many different virulence factors not essentially driven by CagA and VacA alone [12].

In view of this, it appears that virulence of H. pylori is a complex phenomenon evolved by the pathogen as a strategy for survival and adaptation. However, it is not clear how the bacterium colonizes stomach niches for the entire life of its host without being eliminated by the innate defenses of the latter. It is therefore evident that there operates some biological interaction between the host and the pathogen. For example, H. pylori traditionally harnesses its cardinal virulence factors to down-regulate T-cell responses (through the VacA-mediated cell cycle arrest) and up-regulates mucosal proinflammatory pathways (by CagA). However, when CagA and VacA become functionally impaired, there should be some backup mechanism to maintain the above-mentioned inflammatory and apoptotic balance. One such backup mechanism has been proposed by us wherein HP986, a hitherto unknown H. pylori protein, simultaneously functions as a proinflammatory and proapoptotic agent [Alvi et al., unpublished].

H. pylori has evolved by aggregating different genomic islands in its chromosome, including genomic islands associated with the three major type IV secretion systems, of which a very important one, the cagPAI, comprises 29 genes arranged in tandem. Because of the geographically compartmentalized insertion-deletion-substitution patterns occurring within [13] or contiguous to it [14], the cagPAI profiles have also been used as a surrogate marker of human prehistory [10, 14], just like the core genome-based markers such as MLSTs. Based on the genetic instability of this region, several molecular-epidemiology studies have been performed that involved the genotyping of hundreds of strains and isolates belonging to several different gene pools across the continents [15–21]. For instance, the survival and persistence of H. pylori in the aftermath of human migration in India and Peru have recently been explored using core-genome analysis and cagPAI genotypes [13, 16]. H. pylori can acquire DNA from cocolonizing strains belonging to different genetic stocks [21]. However, it appears that the cagPAI and its contiguous region on the right have been exchanged by the ancestral Amerindian strains in South America [16], a continent where human societies were shaped by major geopolitical and sociocultural waves in
recent history with the arrival of European conquerors. It is presumed that the European *H. pylori* strains were equipped with an intact, functional cagPAI, and the ‘indigenous’ *H. pylori* in the Americas such as in Peru that might have carried either no cagA or only a vestigial form of it could have been outcompeted by the newly arrived European strains [16]. The relatedness of cagA-gene sequences from many different Indian isolates and their European counterparts was analyzed to confirm the assumption of gene flow in India through the western corridor with the arrival of Indo-Aryans and the expansion of Neolithic practices and languages from the Fertile Crescent [13]. The cagPAI of the Indian isolates was found to be a complete and intact one without any in-dels in it, and was perhaps acquired well before the arrival of *H. pylori* in India [13, 22]. Such findings support the hypothesis that the genomic regions acquired in prehistoric times could be used as surrogate markers of human history at the interface of microbiology and anthropology [22]. As mentioned, one of the dominant virulence determinants, the CagA antigen, encoded on the cagPAI, is a very widely accepted pathogenicity marker associated with invasive outcomes of infection [23–25]. Strains are often designated as potentially carcinogenic or ulcerogenic according to the amino acid sequence of the CagA protein; for example, using the type and abundance of the EPIYA motif (type C or D and their number of repeats) within the CagA. These signature motifs are critically implicated in phosphorylation of CagA, which ultimately can be linked to the severity of atrophic gastritis and gastric carcinoma associated with CagA-positive strains of *H. pylori* [23]. Given this, several diagnostic and phenotypic studies have been based on this phenomenon [26–28], and many molecular diagnostic methods harness genes or gene products that are encoded by the cagPAI, including CagA [26, 29, 30].

Microevolution and Adaptation during Colonization

‘Chronological evolution’ is a rapidly developing area of specialization in microbiology wherein genetic changes accumulated by hierarchically obtained bacterial isolates can be studied across varied time scales. *H. pylori*, being a chronic colonizer of gastric niches, offers an attractive model system for the study of adaptive changes occurring in the genomes of parasitic organisms. In cases of mixed *H. pylori* infections, recombinant strains evolve with different allelic compositions [12]. Genomic rearrangements directed at achieving fitness by adapting to a changing host environment precipitate spontaneous mutations as demonstrated via serial passages in gnotobiotic piglets [31] or the changes that occur through allelic recombination lead to differential inactivation of important virulence genes such as the case of evolution of vacA in the subclones of *H. pylori* colonizing a single patient [32]. Another important mechanism supportive of the adaptive evolution could be insertion-deletion and substitution polymorphisms of either individual gene loci such as cagA or vacA [11, 12] or the entire genomic islands [12, 33]. In addition, various genotyping methods applied to serial isolates obtained from the same patient revealed similar fingerprints with minor differences [11, 33–35], possibly due to independent genomic alterations despite the isolates being descendants of the same parental strain [11]. This phenomenon has been best described under the term ‘microevolution’ [11, 12, 34]. However, since DNA fingerprinting alone may not be very informative, sequence data are necessary to verify microevolution as much as it is required to confirm (in animal models) whether microevolution results in virulence optimization fine-tuned to a changing gastric environment. Apart from this, the nature and extent of genetic polymorphisms accumulated in *H. pylori*’s genome across wide timescales and during the colonization of different host niches are not fully known. While certain studies explored diversity of strains from individual patients through genetic fingerprinting [11], microarrays [35] and genomic island sequencing [12], whole-genome sequencing of every single clinical isolate is perhaps necessary to investigate the type, extent, frequency and timing of the emergence of small insertions, deletions and substitutions and their functional significance in terms of adaptive mechanisms.

Replicate Genomics: An Inevitable Pursuit in Studying the Evolution of *H. pylori*

A new branch of (bacterial) genomics that entails sequencing and analysis of multiple genomes of serial isolates of the same species or its constituent strains propagated under the same or different environmental conditions or obtained from different hosts or the different niches of the same host could be referred to as ‘replicate genomics’ or ‘replicative genomics’. Unlike other important microbial pathogens such as *Mycobacterium tuberculosis* (whose population structure is clonal and its pan-genome closed), *H. pylori* has the potential to become a target of perhaps several hundred
whole-genome sequencing projects in the near future [36]. Moreover, because it has an overwhelming genetic diversity and, by probability, its pan-genome is infinitely open as dozens of new, undefined genes and genetic elements, it will be identified with the addition of each new genome to the existing pan-genome.

Given the diminishing costs of \textit{H. pylori} genome sequencing on next-generation platforms (such as Illumina), the \textit{H. pylori} genome programs are likely to become pedestrian. It is therefore the appropriate time for even institutions with fewer resources to embark upon genome-inferred molecular epidemiology and transition from genotyping to genome sequencing.

\textit{H. pylori}'s core genome is almost conserved with close to 1,200 genes. However, its ‘variome’, dominated by large, highly fluid, plasticity zones, is of significant research interest because it contains novel genes and genetic elements whose composition and structure vary and evolve in association with the geographic descent of the strains and host ecology/physiology. Functional level consequences of such plastic genomic repertoires could be very significant in terms of adaptations and shaping of the survival mechanisms over time. Therefore, the wealth of comparative genomic data emanating from multiple whole-genome sequencing projects will make it possible to systematically decipher functional consequences of genomic diversity, especially in terms of adaptation, differential virulence potentials and novel mechanisms of virulence. Such data will also be helpful to understand how genotypic information relevant in molecular epidemiology and evolutionary genetics could be scaled up (through functional screens of whole-genome insertion, deletion and substitution patterns of single strains) to ‘functional molecular infection epidemiology’ (FMIE). FMIE is an emerging area of molecular medicine that concerns the study of genetic variations (such as the single-nucleotide polymorphisms) in a pathogen in juxtaposition with the host genetic variations in the context of unique disease phenotype(s). In other words, FMIE could be defined as a holistic concept in which attempts are made to explore aspects of interplay between the pathogens and their hosts through functional analysis of the descriptive host pathogen genomic variations with regard to traits such as adhesion, invasion, persistence and adaptation (on the bacterial side) as well as disease severity, progression and genetic susceptibility or resistance (on the host side) [37].

Another important reason to pursue replicate genomics of \textit{H. pylori} entails the need to study chronological evolution within a single host [38, 39]. The nature and extent of genetic lesions that the chronically inhabiting \textit{H. pylori} accumulates across wide timescales during the colonization of different host niches are not known; this needs in-depth analysis involving those strains which are obtained at different intervals and sampled from different sites of individual patients to know the occurrence of possible insertions, deletions and substitutions enriched over time. Apart from this, geographically distinct strains and their multiple representatives could be sequenced to explore local advantages in terms of host adaptation or disease outcome; for example, \textit{H. pylori} infection in the Indian population (despite a very high colonization rate of up to 90%) rarely leads to serious consequences such as gastric cancer in a significant majority of patients who test positive for \textit{H. pylori} infection. Bacterial coordinates of such ‘protection’, if any, can be studied with the help of bacterial genome sequence data obtained from a number of strains. Multiple whole-genome sequencing of \textit{H. pylori} has recently become an exciting regulatory tool in vaccine studies involving \textit{H. pylori}-mediated antigen delivery. The possibility of using short-term \textit{H. pylori} colonization of the stomach mucosa for antigen delivery represents an attractive strategy for the development of an oral vaccine [40].

Owing to its major advantages over other bacteria and especially its innocuousness when disarmed of its major toxins, \textit{H. pylori} has a remarkable profile that provides for the safety of the organism. Full-genome sequencing could serve as a primary tool for validation of genomic integrity (or rearrangements) during the period(s) of experimental colonization and evaluation of genetic risks due to horizontal gene transfer or other genetic rearrangements that may lead to increased virulence or reestablishment of virulence of the live bacterial vector.

### The Future

\textit{H. pylori} constantly evolves and recasts its genome through microevolutionary changes and via horizontal gene transfer events which could be detected over time through genome sequencing. Therefore, the data generated based on chronological and replicate genomics are likely to explain host-microbe interactions that could potentially occur with colonization by \textit{H. pylori}, wherein some could prove beneficial, leading the coevolved bacterium and its host to reach homeostasis akin to a ‘symbiotic’ relationship. Such descriptive host-pathogen associations, when systematically unraveled, would form the core of approaches heralding the FMIE. Under such an umbrella as FMIE, the pathogenic or commensal attri-

Coevolution and Adaptation of \textit{H. pylori}

---

Med Princ Pract 2011;20:497–503

501
butes of a coevolved pathogen as *H. pylori* would be established with rigor and certainty. Perhaps, it might be relevant to carefully analyze and integrate into the FMIE projects the data that support harmful or beneficial effects of *H. pylori* colonization (see below).

The documented high rates of colonization by *H. pylori* worldwide, and particularly in resource-poor countries of the third world, contrast with the much lower incidence of *H. pylori*-associated diseases in these countries. It has also been suggested that *H. pylori* infection could actually be protective of other chronic pathogens such as *M. tuberculosis* [41], due to the fact that *H. pylori* infection induces bystander effects, modifying the risk of active tuberculosis in humans and nonhuman primates. These arguments and the data related to the association of *H. pylori*’s eradication with enhanced incidence of gastroesophageal reflux disease and esophageal cancers strengthen the idea that *H. pylori* has protective effects on its human host [6, 7].

On the other hand, there have been many attempts to associate *H. pylori*-induced gastric diseases with specific virulence factors. However, no specific associations of these virulence factors have been identified that prospectively link up with symptomatic disease in humans or clinical outcomes with epidemiological consistency. The combination of replicate genomics and phenotypic studies in animal models would be able to resolve the frustrating lack of correlation of putative virulence coordinates with pathogenicity or symptomatic infection.

**Conclusion**

Given the conflicting sides of *H. pylori*’s association as discussed above, scientists have tried to give this pathogen the benefit of the doubt and advocated its carriage in asymptomatic cases. It could be very pertinent, especially in developing countries, where eradication does not hold dramatically beneficial effects due to rampant drug resistance nor does it appear to be economically viable due to frequent recolonization. Some recommend its use as a vaccine vehicle when the organism is disarmed of its toxins. Finally, projecting *H. pylori* as a friendly bug that is otherwise condemned for its role in causing approximately 400,000 deaths per year from gastric adenocarcinoma runs the risk of reduced priority for funding *H. pylori* control and eradication schemes. The funders might reorient priority for research support in countries such as India where the overt outcomes of *H. pylori* colonization are only marginal, thus making *H. pylori* the less important pathogen.

**Acknowledgment**

*H. pylori*-related research in our laboratories is funded through a matching grant by the University of Hyderabad/UGC (India) under the aegis of a German Research Foundation (DFG)-sponsored international research training group entitled ‘Internationales Graduiertenkolleg-Functional Molecular Infection Epidemiology-GRK1673 (Berlin-Hyderabad).’

**References**


