Comparative genomics of *Helicobacter pylori*

Quan-Jiang Dong, Qing Wang, Ying-Nin Xin, Ni Li, Shi-Ying Xuan

Quan-Jiang Dong, Qing Wang, Ying-Nin Xin, Ni Li, Shi-Ying Xuan, Department of Gastroenterology, Qindao Municipal Hospital, Qindao 266071, Shandong Province, China

Author contributions: Dong QJ and Xuan SY performed the majority of the work; Wang Q, Xin YN and Li N were all involved in the preparation of the manuscript.

Correspondence to: Shi-Ying Xuan, Professor, Department of Gastroenterology, Qindao Municipal Hospital, Qindao 266071, Shandong Province, China. jiangacer@126.com

Telephone: +86-532-88905629
Received: March 16, 2009
Revised: June 13, 2009
Accepted: June 20, 2009
Published online: August 28, 2009

Abstract

Genomic sequences have been determined for a number of strains of *Helicobacter pylori* (*H pylori*) and related bacteria. With the development of microarray analysis and the wide use of subtractive hybridization techniques, comparative studies have been carried out with respect to the interstrain differences between *H pylori* and inter-species differences in the genome of related bacteria. It was found that the core genome of *H pylori* constitutes 1111 genes that are determinants of the species properties. A great pool of auxiliary genes are mainly from the categories of *cag* pathogenicity islands, outer membrane proteins, restriction-modification system and hypothetical proteins of unknown function. Persistence of *H pylori* in the human stomach leads to the diversification of the genome. Comparative genomics suggest that a host jump has occurred from humans to felines. Candidate genes specific for the development of the gastric diseases were identified. With the aid of proteomics, population genetics and other molecular methods, future comparative genomic studies would dramatically promote our understanding of the evolution, pathogenesis and microbiology of *H pylori*.

© 2009 The WJG Press and Baishideng. All rights reserved.

Key words: *Helicobacter pylori*; Genomics; Pathogenesis; Cancer

Peer reviewer: Da-Jun Deng, Professor, Department of Cancer Etiology, Peking University School of Oncology, 1 Da-Hong-Luo-Chang Street, Western District, Beijing 100034, China

Dong QJ, Wang Q, Xin YN, Li N, Xuan SY. Comparative genomics of *Helicobacter pylori*. World J Gastroenterol 2009; 15(32): 3984-3991. Available from: URL: http://www.wjgnet.com/1007-9327/15/3984.asp DOI: http://dx.doi.org/10.3748/wjg.15.3984

INTRODUCTION

The gastric pathogen, *Helicobacter pylori* (*H pylori*), is a member of the epsilon-bacteria. This microaerophilic, Gram-negative bacterium colonizes the human stomach[1]. It is estimated that over half of the human population are infected by *H pylori*[2]. The infection causes mucosal inflammation, atrophy, ulceration and cancer[3,4]. Five strains of *H pylori* and a number of related bacteria have been sequenced. Genomics, evolutionary studies and population genetics have advanced our understanding of this bacterium.

GENOMIC FEATURES

In 1997, *H pylori* strain 26695 was firstly sequenced[5]. It was isolated from an English patient with chronic gastritis. The chromosome of strain 26695 is circular and composed of 1667 867 base pairs. The average GC content is approximately 39%. In the initial annotation, it has 1590 open reading frames that are possibly protein-coding[6], in addition to the RNA coding genes (2 copies of 16S rRNA and 23S rRNA genes, 36 tRNA genes). Later analysis of the genome sequence suggested a smaller number of ORFs in strain 26695[7]. The ongoing studies have found genes that were neglected in the initial analysis. A general secretion machinery is widely present in bacteria, which functions in secretion of outer membrane proteins from the inner membrane and delivery of proteins to extracellular environments[3]. The initial annotation revealed a partial general secretion machinery because it lacked SecE in 26695[6]. A further analysis of the genome sequences with GeneMark, Glimmer and BlastX found a small open reading frame (2 copies of 16S rRNA and 23S rRNA genes, 36 tRNA genes). Later analysis of the genome sequence suggested a smaller number of ORFs in strain 26695[7]. The initial annotation revealed a partial general secretion machinery because it lacked SecE in 26695[6]. A further analysis of the genome sequences with GeneMark, Glimmer and BlastX found a small open reading frame between nusG and rpmG (HP1203-HP1204)[8]. It has a high homology and structural similarity to the SecE protein in related bacteria. Therefore, strain 26695 has a complete general secretion machinery, which is consistent with the fact that the bacterium is capable of protein secretion. In addition, small RNA genes are universally present in bacteria[9]. The tRNA gene (tRNA-4) has been found in *H pylori*, which encodes a functional RNA molecule and a small peptide that is involved in...
quality control of translation. In addition, _H pylori_ also possesses a sRNA gene encoding the RNA component of Rnase P and the 4.5S RNA gene which is involved in secretion.

In 1999, strain J99 was sequenced which was isolated from an American patient with a duodenal ulcer. Compared to strain 26695, it has a slightly smaller circular chromosome (1,643,831). The overall genomic organization, gene order and predicted proteomes of the sequenced strains are very similar. The predicted open reading frames are less in strain J99, amounting to 1,495. There are 1,406 genes shared by both strains, but 86 open reading frames are absent from strain 26695. Both strains contain a complete _cag_ pathogenicity island that codes for a type IV secretion system which delivers the CagA cytotoxin protein into gastric epithelial cells. Comparison of the two genomes reveals the occurrence of translocation and inversion events. A 83 kb inversion contains most of the strain specific genes. This region was named a plasticity zone since it has a much lower GC content (35%) than the rest of the genome.

In 2006, a chronic atrophic gastritis _H pylori_ strain, HPAG1, was sequenced. It was isolated from an 80 year-old female patient who was enrolled in a Swedish case-control study of gastric cancer. Similar to the sequenced strains 26695 and J99, HPAG1 is a type 1 strain that contains _cagA_ and a virulente allele of _vacA_. The genome of HPAG1 (1,596,366 bp) is the smallest in the three sequenced strains. A total of 1,536 open reading frames were predicted. Of these, 43 genes are only present in HPAG1. Analysis revealed that 29 genes that are found in both J99 and 26695 were missing from HPAG1. If genes in a strain are absent from other strains, they are called strain-specific genes. The comparison of three sequenced _H pylori_ strains shows that the majority of strain-specific genes are functionally unknown. Another group of strain-specific genes is composed of genes of the R-M system (restriction-modification). They encode proteins involved in DNA restriction or modification. Other strain-specific genes include those encoding outer membrane proteins and _cag_ proteins.

_**DETERMINATION OF THE CORE GENOME OF H PYLORI**_

Inter-strain diversity, represented by variations in number and contents of genes, chromosomal rearrangements and allelic diversities, is not unique to _H pylori_. This has been noted in a number of other bacterial species. For _H pylori_, each strain contains many strain-specific genes. It has been proposed that a particular bacterial species contains a core set of genes and the auxiliary genes. The core genome contains genes that are present in all or nearly all of the strains. It determines the properties that are characteristic of the species. The auxiliary genes are present in some of the strains. They are determinants of the biological properties unique to some of the strains. Salama et al. firstly explored the core set of genes in _H pylori_. A total of 15 strains of _H pylori_ mainly isolated from Western countries were examined using a microarray method. It was found that 1,281 genes were common to all the examined strains, therefore constituting the core genome of _H pylori_. Considering the limited number of strains examined and the fact that _H pylori_ is highly prevalent in human, could these genes represent the actual core set of the _H pylori_ species? Additionally, these strains were only isolated from Western individuals. In fact, molecular typing of global strains has found that the modern population of _H pylori_ divides into five major

It is predicted that strain G27 has 58 genes that are not found in 26695, J99, or HPAG1. The _H pylori_ S hiak genome has also been sequenced by Washington University Medical School. It is 1.61 Million bp long and contains approximately 1,609 predicted genes. The sequences are available on the university website (http://hpylori.ucsc.edu).

The finding of strain-specific genes from the comparison of the sequenced strains is in agreement with the earlier studies which demonstrated the high diversity of the _H pylori_ genome. No identical strains of _H pylori_ have been found in their genetic types unless they are isolated from a family. _H pylori_ has great mutation and recombination capacities. Analysis of the genomic sequences failed to identify a complete mismatch repair system controlling the confidentiality of replication, despite the presence of a homology of Muts. This results in a high mutation rate of _H pylori_. Examination of 29 clinical isolates revealed that approximately 1/4 of them had mutator phenotype with higher mutation frequencies than Enterobacteriaceae mismatch-repair defective mutants. In another study, examination of paired strains isolated from a patient at different times suggested a mutation rate of 4.1 × 10⁻⁶, which is comparable to that of _E. coli_ mutator. _H pylori_ is naturally competent for transformation. Nonrandomly distributed repetitive sequences are found in the genome, which leads to frequent recombination events. The recombination rate (recombination events starting at any particular nucleotide) is estimated to be 6.9 × 10⁻₁². High levels of recombination and mutation could explain the observed genomic diversity in _H pylori_.

It was found that 1281 genes were common to all the examined strains, therefore constituting the core genome of _H pylori_. Considering the limited number of strains examined and the fact that _H pylori_ is highly prevalent in human, could these genes represent the actual core set of the _H pylori_ species? Additionally, these strains were only isolated from Western individuals. In fact, molecular typing of global strains has found that the modern population of _H pylori_ divides into five major
groups, hpEurope, Hpafrica1, hpafrica2, hpEastAsia and hpAsia2[14,41]. They are possibly derived from different ancestral groups. Gressmann et al[41] further examined 56 globally representative strains of H pylori. The number of the genes in the core set of H pylori was diminished to 1150. The author concluded through a calculation that the core genome of H pylori only consists of 1111 genes.

The auxiliary set of genes in H pylori amounts to 22%-27% of the genome[14,41]. In agreement with the findings from the whole genome sequencing of H pylori strains, the auxiliary genes consist of those coding for functionally unknown proteins, cag protein, outer membrane proteins and proteins of DNA metabolism[1].

Candidate genes specific for development of gastric diseases.

The long term clinical outcomes of the H pylori infection are diverse. The infected gastric mucosa may develop inflammation, atrophy, intestinal metaplasia, ulceration, cancer and MALT lymphomas[3,4,6]. Genes in the auxiliary set are specific only for some strains. Do they play roles in the determination of the final outcome of an infected individual? H pylori broth culture filtrates cause the formation of intracellular vacuoles in mammalian cells[49]. The protein which has the vacuolation activity was purified and named VacA (Vacuolating cytotoxin). Despite of the universal presence of the vacA gene in H pylori, some strains do not cause the vacuolation of epithelial cells. This is attributed to the mosaic structure of vacA[26]. A signal region in the N-terminal and a mid region of vacA are polymorphic. The signal region affects the vacuolation activity of the cytotoxin: a 12 amino acid extension on the s2 form blocks the activity, although not all s1 forms have the cytotoxic activity[43]. The mid region is a determinant of the cell specificity of VacA[46]. There are three vacA genotypes, s1/m1, s1/m2 and s2/m2 in H pylori. The association of s1/m1 strains with severe diseases has been observed in some studies. A recent study found an intermediate region (i-region) of vacA between the signal region and the mid region that also contributes to the levels of vacuolating activity[46]. The genotype i1 was more frequently found in gastric cancer-associated H pylori than the i2[49]. Strains possessing vacA i1 are strongly associated with peptic ulcer[60]. Another protein has been found to be co-present in almost all of the strains possessing the vacuolation activity[47]. The protein was named as cytotoxin-associated gene A (cagA) protein. The gene is present in the majority of strains. Those possessing the vacuolation activity and the CagA expression are called type 1 strains, or virulent strains[48]. The presence of cagA is generally the marker for a large DNA region called vacA island named Clone P32 with homology to GepA in Dichelobacter nodosus[39,40]. Proteins produced by the vacA island make up a type IV secretion system which delivers CagA into the epithelial cells[49,50]. The type IV secretion system locates across the inner and outer membrane and forms a pilus-like structure at the surface[31,32]. The CagL protein is a specialized adhesin that is targeted to the pilus surface[33]. Through an arginine-glycine-aspartate motif, it binds to and activates integrin α5β1 receptor on gastric epithelial cells. This interaction triggers CagA delivery into target cells[54] and activation of Src of gastric epithelial cells[55]. Translocated CagA remains associated with the host membrane and becomes tyrosine phosphorylated on carboxyl-terminal repeat motifs (Glu-Pro-Ile-Tyr-Ala, or EPIYA motifs) by members of the Src family of protein tyrosine kinases such as c-Src, Fyn, Lyn, and Yes[56]. Phosphorylated CagA interacts with SHP-2[59], which thereafter activates a number of phosphorylases inducing alteration of signaling pathways. This alters the spreading, migration, adhesion, polarity and cytoskeletal structures of epithelial cells[60-63] . A large European study, demonstrated that cagA positive strains are significantly associated with the development of gastric cancer[64]. The cag island is thus an important determinant of the clinical outcomes of the H pylori infection. Most H pylori strains, and almost all in certain geographical locations, however, are virulent (that is they expressing CagA and VacA). Are there any other genomic differences associated with the clinical outcomes?

Comparison of the genomic contents of different strains has found genes that are potentially disease-specific. Peptic ulcer disease frequently occurs in humans with severe, or even lethal complications. The disease may also affect children. Oleastro et al[65] reported the study of the genomic comparison of a H pylori strain isolated from a child presenting with duodenal ulcer and a strain from a sex and age matched child with gastritis. It was found that genes jhp0562 and jhp0870 are more frequently seen in children with peptic ulcer than in those with gastritis. Both genes are located in the plasticity zone. Jhp0562 encodes a putative LPS glycosyltransferase involved in LPS biosynthesis[66], whereas jhp0870 codes for an outer membrane protein. LPS and outer membrane proteins play roles in the induction of an inflammatory response from the gastric mucosa[66,67]. Whether jhp0562 and jhp0870 contribute to the development of ulceration in children deserves further study. Other genomic comparison studies have found that the cag island and a 670 bp-long DNA fragment that is partially homologous to the hymidylate kinase gene are potentially associated with peptic ulcer diseases[68]. Gastric mucosa infected by H pylori develops inflammation, and gradually become atrophic. Mucosal atrophy is an important stage in stomach carcinogenesis. A thorough examination of the genome of 6 strains from atrophic gastritis found a set of 121 “ChAG-associated” (ChAG, chronic atrophic gastritis) genes[14,69]. They are universally present in these 6 strains but absent from 56 globally derived strains of H pylori[40]. Their putative roles in the development of atrophy and promotion of carcinogenesis are yet to be studied. Intestinal metaplasia of gastric mucosa is a precancerous lesion. H pylori in the patient with intestinal metaplasia is likely associated with progression into gastric cancer. Comparison of a cancer strain and a duodenal ulcer strain of H pylori found a novel sequence named Clone P32 with homology to GepA in Dichelobacter nodosus[39]. Examination of strains from diverse gastric diseases demonstrated that Clone P32 is inversely associated with intestinal metaplasia. Gastric B cell lymphoma of mucosa associated lymphoid tissue is highly associated with the H pylori infection[1]. Eradication of the bacterium leads to the alleviation of the disease. Comparing
the genome of a strain from gastric B cell lymphoma with that from gastritis revealed that jhp0950 encoding a *H pylori* specific protein of unknown function was potentially associated with the development of the disease.[71] It was present in about 3 quarters of strains from gastric lymphoma, but only present in about half of strains from gastritis, duodenal ulcer or gastric adenocarcinoma. If other virulent factors were taken into account, the odds of having gastric MZBL among patients harbouring JHP950, IceA1 (coding for a restriction enzyme), and sabA (coding for a major adhesin) “on” strains were 10 times higher than for the others.[72] Although these genes are specific for strains from a specified disease, it is uncertain whether they are pathogenic for a particular disease. Actually, different gastric diseases greatly differ in a variety of environmental factors that have potential impacts on the biological behaviors and genetics of the bacterium. For example, secretion of gastric acid is varied in different diseases.[73] Therefore, further studies are required to say that a gene is specific for the pathogenesis of a particular disease.

**INTRA-HOST EVOLUTION OF THE H* PYLORI* GENOME**

It is believed that the *H pylori* infection is usually acquired in childhood.[74] The bacterium is transmitted from parents to their children with a bias of mother to children transmission.[75,76] Once the infection is established, the bacterium persists in the host for decades unless eradicated with antibiotics. Transmission of bacteria to a new host is a major barrier for bacterial spreading. It may affect the bacterial genome contents. Four healthy adults were experimentally infected with *H pylori*. Examination of isolates form 15 d or 90 d postinfection demonstrated that their genomic contents were identical to the challenging strain.[77] A similar result was found in a study of experimental infected mice.[78] These findings suggest that for *H pylori*, transmission does not cause any alteration of the gene components of the genome, or, in other words, the establishment of the *H pylori* infection does not require the involvement of novel genes. Further evidence supporting this conclusion comes from a study of the analysis of the strains from a mother and her three children.[79] Microarray analysis demonstrated that the genomic contents of isolates from the mother were identical to those from the children. *H pylori* persists in the human stomach for decades, probably from childhood. It may experience a variety of ecological alterations, which may in turn have large impacts on the genome of the organism. The output of gastric acid alters with aging and with infection by *H pylori*. Alterations of the constituents and the quantity of the gastric mucus underneath which the bacterium resides are observed during the course of the *H pylori* infection. The gastric epithelial cells may undergo metaplasia and changes in surface proteins, which affect adhesion and the supply of nutrients. The gastric mucosa may produce immune and inflammatory products against the bacterium. The co-infection with other microbes is also frequent seen in the stomach. These alterations may affect genomic contents of *H pylori*. Kraft et al.[80] examined paired strains of *H pylori* with respect to their genomic contents using the microarray method. Paired strains were isolated from the same patients at an interval from 3 to 36 mo. Of 21 pairs of strains examined, 4 pairs showed differences in their genomic contents, suggesting the occurrence of evolutionary events. These included a complete deletion and a partial loss of the sag pathogenicity island, a replacement of an open reading frame of unknown function with the restriction-modification system HpyAIV, an acquisition of 14 genes in the plasticity zone, a duplication of the cenE genes (HP1561/HP1562) and a truncation of tandem arranged ackA and pta genes resulting in the formation of pseudogenes. A study has compared 2 pairs of strains obtained from the same patients at an interval of 4 years[81]. The patients had progressed from atrophic gastritis to cancer. Six genes were absent, including 3 genes involved in DNA repair, an outer membrane protein and two hypothetical proteins. Nine genes were gained, including a ligase, a metalloprotease, a tRNA formyltransferase, a putative ribonuclease, a restriction enzyme and four hypothetical genes. The comparison suggests that with the progression of the atrophy to cancer, the bacterium may have a propensity for losing its diversifying capacity. Findings from these studies demonstrate that *H pylori* may acquire or lose genes during the intra-host colonization[82]. The genes involved fall into the same categories as the strain specific genes. This was further supported by the results from the comparison of the sequenced strain J99 with isolates obtained 6 years later.[83] These comparative studies of the *H pylori* genome draw a picture of the genomic changes during the cycle of invasion, colonization and transmission to a new host. It appears that invasion into a new host has little effect on the gene composition of the genome. This indicates that the current genome of *H pylori* has sufficient capacities for permitting bacterial invasion into a human host or even into a host of different species under experimental conditions. Once the infection is established, the bacterium has to cope with the dynamic changes of the ecology during the long-term coexistence with the host. Genomic diversifications, or gain and/or loss of genes, occur in response to these changes. The diversifications involve genes that are mainly those strain specific genes observed from comparative studies of unrelated strains of *H pylori*. Intra-host evolution of *H pylori*, thus, results in the creation of a pool of genes that are generally needed by some strains. This pool of genes can be considered as the auxiliary set of genes of *H pylori*.

**COMPARATIVE STUDIES OF H* PYLORI* AND ITS RELATED BACTERIA**

Since the isolation of *H pylori*, a number of closely related bacteria have been identified, constituting a
new bacterial genus named Helicobacter acinonychis.[82,83] Bacteria within this genus have been shown to colonise the gastrointestinal tract of mammals. Many of these Helicobacter species are involved in the pathogenesis of gastrointestinal diseases.[82,83] Phylogenetic analysis has shown that the helicobacters can be separated into two clusters[84]. Gastric species that colonise the stomach of mammals form a cluster. Species that inhabit the intestine and biliary tracts cluster together to form the enterohelipater cluster. In addition to H pylori, the genome sequences have been determined for several other helicobacters, including H. mustelae from ferret, H. acinonychis from large felines (cheetahs, lions and tigers)[85], H. hepaticus from mice which causes hepatomain[86], and Wolinella succinogenes from cattle[87].

General genomic features of these helicobacters are listed in Table 1, which also includes information for several species of campylobacters[88-91]. Of these related bacteria, the size and GC content of H. acinonychis are most similar to those of H pylori[88]. Comparison of 612 orthologues that are present in both H. acinonychis and H pylori found that they differ at only few of their amino acids. The Blast scores against H pylori of most coding sequences in H. acinonychis are very high. These findings lead to the conclusion that a host jump has occurred from human to feline[89]. This event probably occurred 100,000 years ago. More studies are required to confirm this conclusion considering that universally accepted criteria to identify a host jump event are currently unavailable. The study also found a set of fragmented genes and newly acquired genes in H. acinonychis. These genes include a set of genes encoding outer membrane proteins and a cluster of genes encoding proteins for sialylation of bacterial surface carbohydrates. It has been suggested that these genes are probably beneficial for the bacterium to evade host immune defenses[89].

Information from comparative genomics has greatly enhanced our understanding of the microbiology, physiology, evolution and pathogenesis of H pylori. Candidate genes specific for the development of the gastric disease, particularly gastric cancer have been identified. Considering the striking diversities in the H pylori genome which are intensified by intra-host evolution, further studies exploring these genes must take account of them.

REFERENCES

1. Blaser MJ, Atherton JC. Helicobacter pylori persistence: biology and disease. J Clin Invest 2004; 113: 321-333
2. Lehoux P, Yilmaz O. Epidemiology of Helicobacter pylori infection. Helicobacter 2007; 12 Suppl 1: 1-3
3. Blaser MJ. Helicobacter pylori and gastric diseases. BMJ 1998; 316: 1507-1510
4. Sipponen P, Hyvarinen H, Seppala K, Blaser MJ. Review article: Pathogenesis of the transformation from gastritis to malignancy. Aliment Pharmacol Ther 1998; 12 Suppl 1: 61-71
5. Tomb JF, White O, Kurokawa K, Clayton RA, Sutton G, Fleischmann RD, Ketchum KA, Lyle S, Dougherty BA, Nelson K, Quakenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glocalized A, McKenney K, Fitzgerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. The complete genome sequence of the gastric pathogen Helicobacter pylori. Nature 1997; 388: 539-547
6. Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig C, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel DE, Vovis GF, Trust TJ. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. J Bacteriol 1999; 181: 329-334
7. Biener KL, Silhavy T. The genetics of protein secretion in E. coli. Trends Genet 1990; 6: 329-334
8. Medigue C, Wong BC, Lin MC, Bocs S, Danchin A. The see gene of Helicobacter pylori. J Bacteriol 2002; 184: 2837-2840
9. Wassarman KM, Repola F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15: 1637-1651
10. Dong Q, Zhang L, Goh KL, Forman D, O'Rourke J, Harris TE, Voss G, Fitch TJ, Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature 1997; 397: 176-180
11. Biekker KL, Silhavy TJ. The genetics of protein secretion in E. coli. Trends Genet 1990; 6: 329-334
12. Medigue C, Wong BC, Lin MC, Bocs S, Danchin A. The see gene of Helicobacter pylori. J Bacteriol 2002; 184: 2837-2840
13. Wassarman KM, Repola F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15: 1637-1651
14. Dong Q, Zhang L, Goh KL, Forman D, O'Rourke J, Harris TE, Voss G, Fitch TJ, Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature 1997; 397: 176-180
15. Biekker KL, Silhavy TJ. The genetics of protein secretion in E. coli. Trends Genet 1990; 6: 329-334
16. Medigue C, Wong BC, Lin MC, Bocs S, Danchin A. The see gene of Helicobacter pylori. J Bacteriol 2002; 184: 2837-2840
17. Wassarman KM, Repola F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15: 1637-1651
18. Dong Q, Zhang L, Goh KL, Forman D, O'Rourke J, Harris TE, Voss G, Fitch TJ, Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature 1997; 397: 176-180
19. Biekker KL, Silhavy TJ. The genetics of protein secretion in E. coli. Trends Genet 1990; 6: 329-334
20. Medigue C, Wong BC, Lin MC, Bocs S, Danchin A. The see gene of Helicobacter pylori. J Bacteriol 2002; 184: 2837-2840
21. Wassarman KM, Repola F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15: 1637-1651
22. Dong Q, Zhang L, Goh KL, Forman D, O'Rourke J, Harris TE, Voss G, Fitch TJ, Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature 1997; 397: 176-180
23. Biekker KL, Silhavy TJ. The genetics of protein secretion in E. coli. Trends Genet 1990; 6: 329-334
24. Medigue C, Wong BC, Lin MC, Bocs S, Danchin A. The see gene of Helicobacter pylori. J Bacteriol 2002; 184: 2837-2840
25. Wassarman KM, Repola F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15: 1637-1651
26. Dong Q, Zhang L, Goh KL, Forman D, O'Rourke J, Harris TE, Voss G, Fitch TJ, Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature 1997; 397: 176-180
27. Biekker KL, Silhavy TJ. The genetics of protein secretion in E. coli. Trends Genet 1990; 6: 329-334
28. Medigue C, Wong BC, Lin MC, Bocs S, Danchin A. The see gene of Helicobacter pylori. J Bacteriol 2002; 184: 2837-2840
29. Wassarman KM, Repola F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15: 1637-1651
30. Dong Q, Zhang L, Goh KL, Forman D, O'Rourke J, Harris TE, Voss G, Fitch TJ, Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature 1997; 397: 176-180
31. Biekker KL, Silhavy TJ. The genetics of protein secretion in E. coli. Trends Genet 1990; 6: 329-334
32. Medigue C, Wong BC, Lin MC, Bocs S, Danchin A. The see gene of Helicobacter pylori. J Bacteriol 2002; 184: 2837-2840
33. Wassarman KM, Repola F, Rosenow C, Storz G, Gottesman S. Identification of novel small RNAs using comparative genomics and microarrays. Genes Dev 2001; 15: 1637-1651
Dong QJ et al. H pylori genomics 3989

13 Kutter S, Buhdorf R, Haas J, Schneider-Bracht W, Haas R, Fischer W. Protein subassemblies of the Helicobacter pylori Cag type IV secretion system revealed by localization and infection studies. J Bacteriol 2008; 190: 2161-2171

14 Oh JD, Kling-Backhed H, Giannakis M, Xu J, Fulton RS, Fulton LA, Cordum HS, Wang C, Elliott G, Edwards J, Mardis ER, Engstrand LG, Gordon JL. The complete genome sequence of a chronic atrophic gastritis Helicobacter pylori strain: evolution during disease progression. Proc Natl Acad Sci USA 2006; 103: 9999-10004

15 Enroth H, Kraae W, Engstrand L, Nyren O, Rohan T. Helicobacter pylori strain types and risk of gastric cancer: a case-control study. Cancer Epidemiol Biomarkers Prev 2000; 9: 981-985

16 Baltrus DA, Amieva MR, Covacci A, Lowe TM, Merrell DS, Ottemann KM, Stein M, Salama NR, Guillemin K. The complete genome sequence of Helicobacter pylori strain G27. J Bacteriol 2009; 191: 447-448

17 Covacci A, Censini S, Bugnoli M, Petracca R, Burroni D, Macchia G, Massone A, Papini E, Xiang Z, Figura N. Molecular characterization of the 128-kDa immunodominant antigen of Helicobacter pylori associated with cytotoxicity and duodenal ulcer. Proc Natl Acad Sci USA 1993; 90: 5791-5795

18 Censini S, Lange C, Xiang Z, Crabtree JE, Ghia B, Burovsky M, Rappuoli R. CagA, a pathogenicity island of Helicobacter pylori, encodes type I-specific and disease-typical virulence factors. Proc Natl Acad Sci USA 1996; 93: 14648-14653

19 El-ETR SH, Mueller A, Tompkins LS, Falkow S, Merrell DS. Phosphorylation-independent effects of CagA during interaction between Helicobacter pylori and T84 polarized monolayers. J Infect Dis 2004; 190: 1516-1523

20 Guillemin K, Salama NR, Tompkins LS, Falkow S. CagA, a pathogenicity island-specific responses of gastric epithelial cells to Helicobacter pylori infection. Proc Natl Acad Sci USA 2002; 99: 15136-15141

21 Amieva MR, Vogelmann R, Covacci A, Tompkins LS, Nelson WJ, Falkow S. Disruption of the epithelial apical-junctional complex by Helicobacter pylori CagA. Science 2003; 300: 1430-1434

22 Akopyanz N, Bukanov NO, Westblom T, Kresovich S, Berg DE. DNA diversity among clinical isolates of Helicobacter pylori detected by PCR-based RAPD fingerprinting. Nucleic Acids Res 2002; 20: 5137-5142

23 Han FC, Ng HC, Ho B. Stability of randomly amplified polymorphic DNA fingerprinting in genotyping clinical isolates of Helicobacter pylori. World J Gastroenterol 2003; 9: 2021-2024

24 Burucu C, Lhomme V, Fauchere JL. Performance criteria of DNA fingerprinting methods for typing of Helicobacter pylori isolates: experimental results and meta-analysis. J Clin Microbiol 1999; 37: 4071-4080

25 van der Ende A, Rauws EA, Feller M, Mulder CJ, Tytgat GN, Dankert J. Heterogeneous Helicobacter pylori isolates from members of a family with a history of peptic ulcer disease. Gastroenterology 1996; 111: 638-647

26 Roma-Giannikou E, Karameris A, Balatsos B, Panayiotou J, Manika Z, Van-Vliet C, Roskas T, Skandalis N, Kattamis C. Intrafamilial spread of Helicobacter pylori: a genetic analysis. Helicobacter 2003; 8: 15-20

27 Raymond J, Thiberge JM, Kalach N, Bergeret M, Dupont C, Labigne A, Dauga C. Using macrarrays to study routes of infection of Helicobacter pylori in three families. PLoS One 2008; 3: e2259

28 Kang J, Huang S, Blaser MJ. Structural and functional divergence of MuttS2 from bacterial MutS1 and eukaryotic MSH4-MSH5 homologs. J Bacteriol 2005; 187: 3528-3537

29 Pinto AW, Mattheu A, Marsan S, Veauze X, Ielpi L, Labigne A, Radiello JA. Suppression of homologous and homeologous recombination by the bacterial MuttS2 protein. Mol Cell 2005; 17: 113-120

30 Bjorkholm B, Sjolund M, Falk PG, Berg OG, Engstrand L, Andersson DI. Mutation frequency and biological cost of antibiotic resistance in Helicobacter pylori. Proc Natl Acad Sci USA 2001; 98: 14607-14612

31 Falush D, Kraft C, Taylor NS, Corea P, Fox JG, Achtman M, Suerbaum S. Recombination and mutation during long-term gastric colonization by Helicobacter pylori: estimates of clock rates, recombination size, and minimal age. Proc Natl Acad Sci USA 2001; 98: 15056-15061

32 Karmholz A, Hoefler C, Odenbreit S, Fischer W, Hofreuter D, Haas R. Functional and topological characterization of novel components of the comB DNA transformation competence system in Helicobacter pylori. J Bacteriol 2006; 188: 882-893

33 Aras RA, Kang J, Tschumi AI, Harasaki Y, Blaser MJ. Extensive repetitive DNA facilitates prokaryotic genome plasticity. Proc Natl Acad Sci USA 2003; 100: 13579-13584

34 Suerbaum S, Achtman M. Evolution of Helicobacter pylori: the role of recombination. Trends Microbiol 1999; 7: 182-184

35 Suerbaum S, Smith JM, Bapumia K, Morelli G, Smith NH, Kunstmann E, Dyrek I, Achtman M. Free recombination within Helicobacter pylori. Proc Natl Acad Sci USA 1998; 95: 12619-12624

36 Pupo GM, Lan R, Reeves PR, Bavestock PR. Population genetics of Escherichia coli in a natural population of native Australian rats. Environ Microbiol 2000; 2: 594-610

37 Lan R, Reeves PR. Intraspecies variation in bacterial genomes: the need for a species genome concept. Trends Microbiol 2000; 8: 396-401

38 Salama N, Guillemin K, McDaniel TK, Sherlock G, Tompkins LS, Falkow S. A whole-genome microarray reveals genetic diversity among Helicobacter pylori strains. Proc Natl Acad Sci USA 2000; 97: 14668-14673

39 Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Megraud F, Otto K, Reichard U, Katozaitis E, Wang X, Achtman M, Suerbaum S. Traces of human migrations in Helicobacter pylori populations. Science 2003; 299: 1582-1585

40 Linz B, Balloux F, Moodley Y, Manica A, Liu H, Roumagnoc P, Falush D, Stamer C, Prugnolle F, van der Merwe SW, Yamaoka Y, Graham DY, Perez-Trallero E, Wadstrom T, Suerbaum S, Achtman M. An African origin for the intimate association between humans and Helicobacter pylori. Nature 2007; 445: 915-918

41 Grellmann H, Linz B, Gharai R, Pleissner K, Schlaphach R, Yamaoka Y, Kraft C, Suerbaum S, Meyer TF, Achtman M and Gain of loss of multiple genes during the evolution of Helicobacter pylori. PLoS Genet 2005; 1: e43

42 Atherton JC, Cao P, Peek RM Jr, Tummuru MK, Blaser MJ, Cover TL. Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori. Association of specific vacA types with cytotoxin production and peptic ulceration. J Biol Chem 1995; 270: 17771-17777

43 Leelley DP, Atherton JC. Natural diversity in the N terminus of the mature vacuolating cytotoxin of Helicobacter pylori determines cytotoxin activity. J Bacteriol 2000; 182: 3278-3280

44 Ji X, Fernandez T, Burroni D, Pagliaccia C, Atherton JC, Reyrat JM, Rappuoli R, Telford JL. Cell specificity of Helicobacter pylori cytotoxin is determined by a short region in the polymorphic midregion. Infect Immun 2000; 68: 3754-3757

45 Rhead JL, Letley DP, Mohammadi M, Hussein N, Moghaghehi MA, Estahban Hosseini M, Atherton JC. A new Helicobacter pylori vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology 2007; 133: 926-936

46 Basso D, Zambon CF, Letley DP, Stranges A, Marchet A, Rhead JL, Schiavon S, Guariso G, Ceroni M, Nitti D, Rugge M, Plebani M, Atherton JC. Clinical relevance of Helicobacter pylori cagA and vacA gene polymorphisms. Gastroenterology 2008; 135: 91-99

47 Tummuru MK, Cover TL, Blaser MJ. Cloning and expression of a high-molecular-mass major antigen of Helicobacter
chronic Helicobacter pylori infection. J Bacteriol 2006; 188: 249-254
80 Gressmann H, Linz B, Ghai R, Pleissner KP, Schlapbach R, Yamaoka Y, Kraft C, Suerbaum S, Meyer TF, Achtman M. Gain and loss of multiple genes during the evolution of Helicobacter pylori. PLoS Genet 2005; 1: e43
81 Israel DA, Salama N, Krishna U, Rieger UM, Atherton JC, Falkow S, Peek RM Jr. Helicobacter pylori genetic diversity within the gastric niche of a single human host. Proc Natl Acad Sci USA 2001; 98: 14625-14630
82 Fox JG. The non-H pylori helicobacters: their expanding role in gastrointestinal and systemic diseases. Gut 2002; 50: 273-283
83 O’Rourke JL, Grehan M, Lee A. Non-pylori Helicobacter species in humans. Gut 2001; 49: 601-606
84 Mikkonen TP, Karenlampi RI, Hanninen ML. Phylogenetic analysis of gastric and enterohepatic Helicobacter species based on partial HSP60 gene sequences. Int J Syst Evol Microbiol 2004; 54: 753-758
85 Eppinger M, Baar C, Linz B, Raddatz G, Lanz C, Keller H, Morelli G, Gressmann H, Achtman M, Schuster SC. Who ate whom? Adaptive Helicobacter genomic changes that accompanied a host jump from early humans to large felines. PLoS Genet 2006; 2: e120
86 Suerbaum S, Josenhans C, Sterzenbach T, Drescher B, Brandt P, Bell M, Droge M, Hartmann B, Fischer HP, Ge Z, Horster A, Holland R, Klein K, Konig J, Macko L, Mendz GL, Nyakatura G, Schauer DB, Shen Z, Weber J, Frosch M, Fox JG. The complete genome sequence of the carcinogenic bacterium Helicobacter hepaticus. Proc Natl Acad Sci USA 2003; 100: 7901-7906
87 Baar C, Eppinger M, Raddatz G, Simon J, Lanz C, Klimmek O, Nandakumar R, Gross R, Rosinus A, Keller H, Jagtap P, Linke B, Meyer F, Lederer H, Schuster SC. Complete genome sequence and analysis of Wolinella succinogenes. Proc Natl Acad Sci USA 2003; 100: 11690-11695
88 Eppinger M, Baar C, Raddatz G, Huson DH, Schuster SC. Comparative analysis of four Campylobacterales. Nat Rev Microbiol 2004; 2: 872-885
89 Fouts DE, Mongodin EF, Mandrell RE, Miller WG, Rasko DA, Ravel J, Brinkac LM, DelBoy RT, Parker CT, Daugherty SC, Dodson RJ, Durkin AS, Madupu R, Sullivan SA, Shetty JU, Ayodeji MA, Shvartsbeyn A, Schatz MC, Badger JH, Fraser CM, Nelson KE. Major structural differences and novel potential virulence mechanisms from the genomes of multiple campylobacter species. PLoS Biol 2005; 3: e15
90 Gundogdu O, Bentley SD, Holden MT, Parkhill J, Dorrell N, Wren BW. Re-annotation and re-analysis of the Campylobacter jejuni NCTC11168 genome sequence. BMC Genomics 2007; 8: 162
91 Parkhill J, Wren BW, Mungall K, Ketley JM, Churche C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream MA, Rutherford KM, van Vliet AH, Whitehead S, Barrett BG. The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. Nature 2000; 403: 665-668
92 Dailidiene D, Dailide G, Ogura K, Zhang M, Mukhopadhyay AK, Eaton KA, Cattoli G, Kusters JC, Berg DE. Helicobacter acinonychis: genetic and rodent infection studies of a Helicobacter pylori-like gastric pathogen of cheetahs and other big cats. J Bacteriol 2004; 186: 356-365

S-Editor Li LF L-Editor Lalor PF E-Editor Ma WH