Minireview

Discovering human history from stomach bacteria
Todd R Disotell

Address: Department of Anthropology, New York University, 25 Waverly Place, New York, NY 10003, USA. E-mail: todd.disotell@nyu.edu

Abstract

Recent analyses of human pathogens have revealed that their evolutionary histories are congruent with the hypothesized pattern of ancient and modern human population migrations. Phylogenetic trees of strains of the bacterium *Helicobacter pylori* and the polyoma JC virus taken from geographically diverse groups of human beings correlate closely with relationships of the populations in which they are found.

Charles Darwin recognized that the distribution and form of parasites was evolutionarily significant. He noted, for instance, that "... the Pediculi [lice] collected in different countries from the different races of man ... differ, not only in colour, but in the structure of their claws and limbs. In every case in which many specimens were obtained the differences were constant" [1]. More recently, several research groups [2-7] have found interesting correlations between the evolutionary relationships among various bacterial and viral strains hosted by humans and the pattern of migrations of modern humans throughout the world.

A particularly interesting case is that of *Helicobacter pylori*, a Gram-negative bacterium associated with gastritis, peptic ulcers, and gastric cancer that may infect up to half of all humans [8]. The discovery that a bacterial infection could lead to what were considered chronic diseases [8] was a striking example of the fact that infectious diseases have not yet been conquered. The continuing acquired immune deficiency syndrome (AIDS) epidemic, outbreaks of Ebola in Central Africa, and the current spread of West Nile Virus in the United States and of severe acute respiratory syndrome (SARS) from Asia provide evidence of the pervasiveness and health consequences of infectious agents even in the age of vaccination and antimicrobial and antiviral therapies. Many infectious diseases are thought to have arisen concurrently with the development of agriculture and the rise of urban living. If, instead, many pathogens' relationships with humans are much older, it would not be surprising to find

deeper evolutionary associations between humans and their microbial and viral invaders.

The evolutionary history of *H. pylori* may provide an example of the coevolution of a bacterium and its only known host. The *H. pylori* genome is relatively small at 1.67 megabases, with a minimal complement of metabolic genes [9]. Variation between *H. pylori* isolates from different people or even from one person is large, leading to unique fingerprints for nearly every isolate so far typed. The coding genes are not very diverse, however: most of the variation occurs in the third base position within codons or through inversions or translocations, leaving the encoded amino-acid sequences relatively similar [2]. This amino-acid sequence conservation is fortunate for vaccine research, as one vaccine is likely to be effective on many strains. More interesting is the fact that *H. pylori* has an extremely high rate of recombination, higher in fact than that of any other organism characterized to date [3,10].

Normally, such a high rate of recombination would make inferring an organism's evolutionary history very difficult, as information about the origin of each mutation would be lost as it spreads throughout a population. But coupled with the mode of transmission of *H. pylori*, this extremely high rate of recombination may in fact make evolutionary inferences easier. Several studies strongly suggest that *H. pylori* is usually transmitted within families, generally from mother to child [11,12]. Thus, the transmission of *H. pylori* in some ways mimics that of maternally transmitted mitochondrial

DNA [13]. Because mitochondrial DNA is transmitted solely from one parent (the mother) and does not recombine, it has proved to be an ideal genetic system for inferring human evolutionary history (see below) [14,15]. If *H. pylori* is indeed predominantly maternally transmitted, new strains will generally not infect a person during their lifetime; together with the high rate of recombination, this would mean that the mutations that accumulate within the population of bacteria in an individual's stomach will be relatively homogeneous. This should result in a swarm of strains that are very closely related to each other, containing many of the mutations that have occurred in individual bacteria. Swarms found in different people will thus be more different from each other than if there was less recombination.

Most infectious diseases spread quickly throughout the world and strains from different regions are relatively similar, but initial sampling of *H. pylori* from people from different regions of the world revealed fairly strong geographic partitioning into European and Asian *H. pylori* types [2-4]. Recently, Falush and colleagues [5] have examined this partitioning in greater detail. After sequencing eight genes - a total of 3,850 nucleotides - in 370 strains derived from 27 human populations, they found 1,418 polymorphic nucleotide positions. They then applied a new analytic tool, STRUCTURE [16], that was developed to infer human genetic structure from multilocus genotype data. The program uses Bayesian methods to identify subgroups with distinctive allele frequencies, and clusters of the subgroups, even in the presence of recombination [16]. When this technique was applied to the *H. pylori* sequences, four main clusters were found - two from Africa and one each from Europe and Asia (Figure 1a) [5].

Each cluster found by Falush and colleagues [5] could be divided into subgroups; for instance, the 'Africa 1' cluster could be subdivided further into West and South African subclusters, and the East Asia cluster could be split into East Asian, Amerind, and Maori subclusters. The geographic partitioning within the 200 European strains was particularly complicated, presumably because numerous groups have swept back and forth across Europe over the past several millennia. European strains also occasionally appeared in the Americas, Australia, and among South Africans, presumably reflecting colonial conquest.

The phylogenetic relationships of these clusters (Figure 1b) and their subdivisions [5] show a pattern similar to that
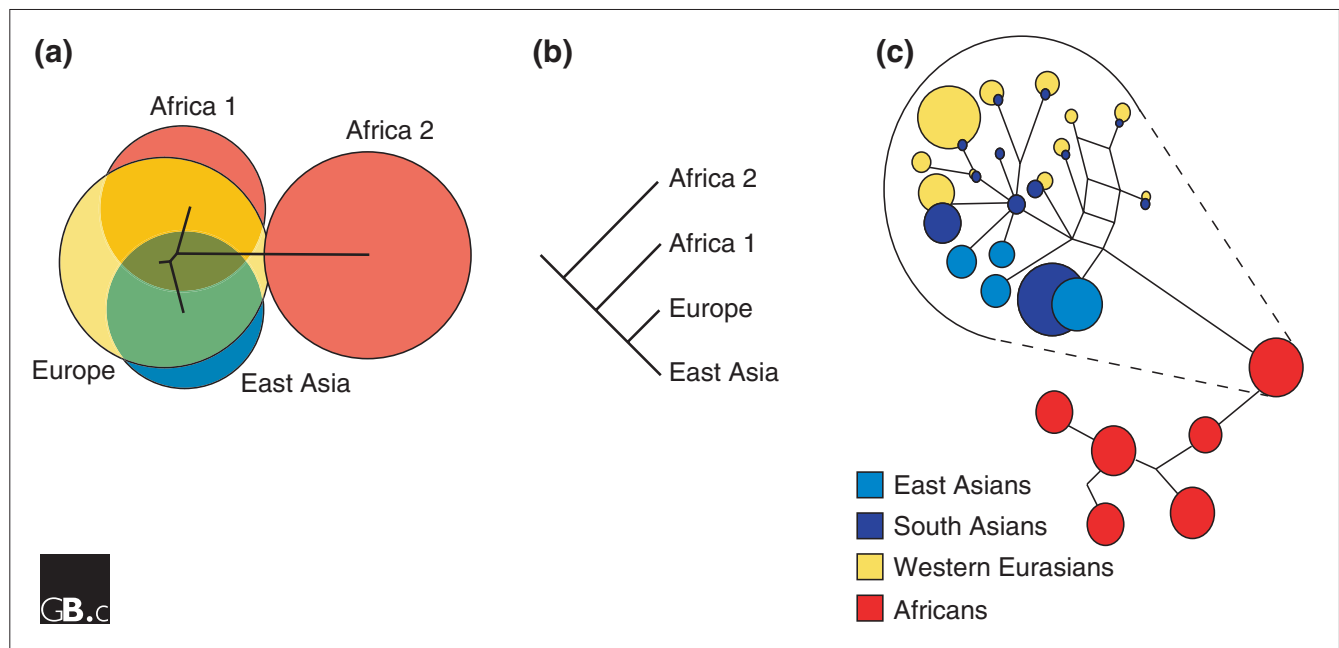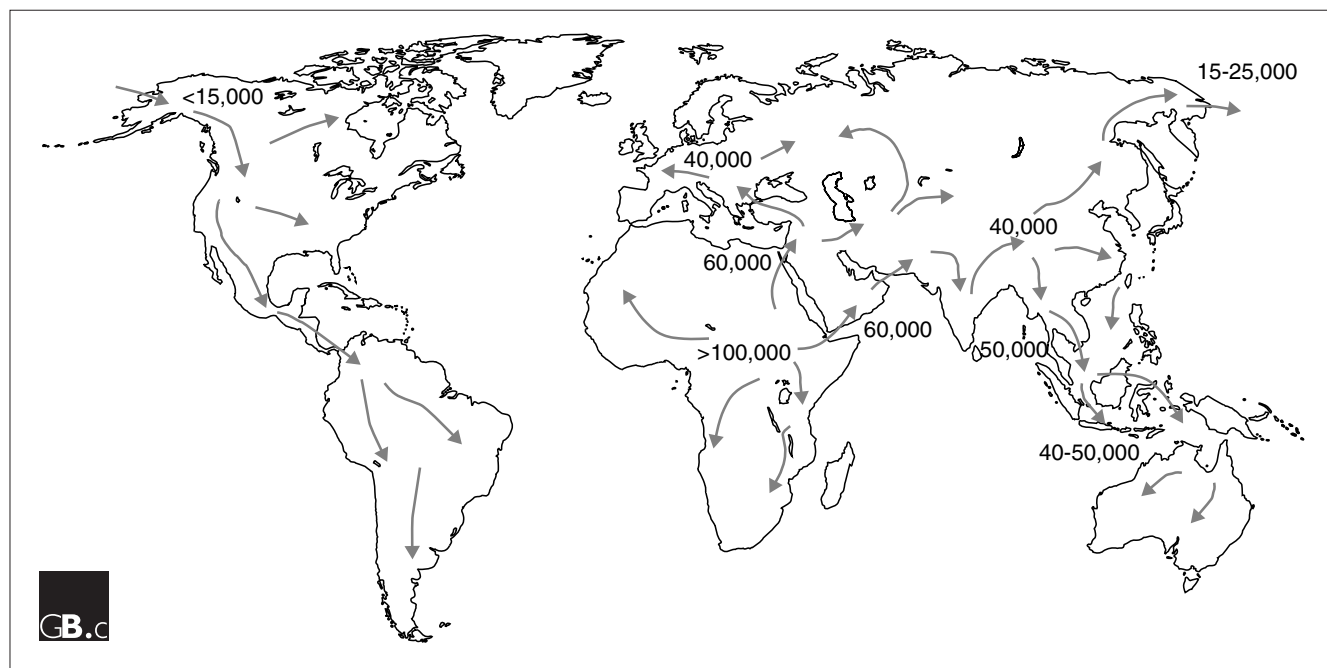

Figure 1
The relationships between human populations, as calculated from *H. pylori* found in stomachs and from mitochondrial DNA data. **(a)** Relationships between modern subpopulations of *H. pylori* [5]. Each subpopulation is represented by a circle with a diameter proportional to the genetic diversity within it. The centres of the circles are joined by a phylogenetic tree showing the relationships between the four subpopulations. Bacteria in each subpopulation are found predominantly in people who originate from the regions shown. **(b)** A population-level phylogenetic tree of the *H. pylori* geographic subpopulations shown in (a). **(c)** A median-joining network of human populations derived from mitochondrial DNA [14]. Such a network shows alternative potential evolutionary relationships between clusters. Each circle represents a cluster of mitochondrial types with a diameter proportional to the frequency of that type within the subpopulations. All non-African populations are derived from one African lineage; the network of relationships within this lineage is magnified (top). (a,b) Adapted from [5]; (c) adapted from [14].

obtained using mitochondrial DNA variation (Figure 1c) [14,15]. The modern human gene pool, as inferred from mitochondrial DNA and corroborated by Y chromosome studies, is thought to have had an African origin approximately 150,000-200,000 years ago [14,15,17,18]. The original human population then spread and diversified throughout Africa for nearly 100,000 years, before expanding into Western Asia and Europe and into Southern and Eastern Asia approximately 50,000-60,000 years ago, replacing the existing archaic populations of humans in these regions. Subsequent migrations spread into Australasia by 40,000 years ago, then to the Pacific Islands, and later into North America, approximately 15,000 years ago (Figure 2) [14,15,17,18]. The remarkable similarity between this view of human history and the results from studies of *H. pylori* have led Falush and colleagues [5] as well as others [2] to conclude that *H. pylori* evolution has followed the path of modern human expansion and migration. This work thus provides another type of data for analysis of human evolution and migration, independent of mitochondrial DNA and Y chromosomes, that will be valuable in further studies. Unfortunately, however, estimating the divergence dates from *H. pylori* is particularly difficult, owing to the extremely high rate of recombination [10]. Additional sampling and analytical techniques may be required to further test the migratory hypothesis.

Other pathogens have also been proposed to follow evolutionary histories similar to their hosts. One of the most interesting examples, though with a non-human host, is that of aphids, a bacterium found within them, and two plasmids associated with the bacterium. Funk *et al.* [19] found that the inferred intraspecific phylogenies of these four genomes were completely congruent. Returning to human pathogens, the human polyomavirus JC virus (JCV) can be divided into genotypes that correspond to the major continental land masses [6]. Like *H. pylori*, JCV - which can cause progressive multifocal leukoencephalopathy (loss of myelination in the central nervous system) - is very widespread amongst humans as a result of familial transmission. A total of 12 known subtypes have been defined, with European, African, and Asian distributions [7]. Although direct inferences of an African origin for JCV are problematic because there is no suitable outgroup with which to root the phylogenetic tree, when an African origin is assumed, a reasonable evolutionary history can be hypothesized (Figure 3) [7]. As with *H. pylori*, inferring molecular divergence dates is currently problematic for JCV. Further investigation into the evolutionary history of these human pathogens is therefore necessary.

Elucidating the patterns of evolution of human pathogens may ultimately provide additional evidence not only about their history but also about human evolution and history. This will be especially true for pathogens such as *H. pylori* that have a predominantly mother-child mode of transmission, mimicking mitochondrial DNA evolution. *H. pylori*'s causative role in several chronic stomach conditions is just one of many current examples of the known or suspected

Figure 2
A map of the pattern of expansion and migration of modern humans throughout the world, derived from studies of mitochondrial DNA and Y chromosomes [14,15,17,18]. Numbers indicate the approximate time (in years before the present) when modern humans first appeared in the indicated region.
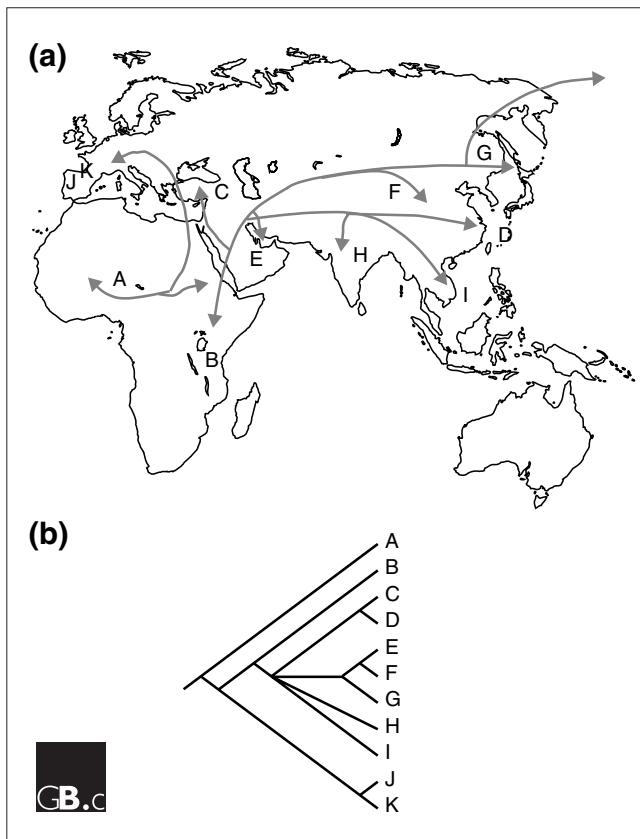
Figure 3
Relationships of human polyoma JC virus (JCV) subtypes found in humans from different parts of the world [7]. Letters refer to individual subtypes.
(a) The hypothesized pattern of spread of JCV subtypes through the world (excluding the Americas); **(b)** an inferred phylogeny of JCV subtypes, assuming an African origin for the virus. Adapted from [7].

role of an infectious agent leading to chronic diseases. Bacteria are suspected to be involved in the development of arteriosclerosis, stroke, and Crohn's disease, whereas viruses are known to lead to AIDS and the various forms of chronic hepatitis. Cervical cancer, hepatocellular carcinoma, Burkitt's lymphoma, Kaposi's sarcoma, and perhaps diabetes mellitus are also either known or suspected to be of viral origin. If some of the ubiquitous chronic diseases prove to be of bacterial or viral origin, worldwide surveys of these pathogens in human populations should be carried out immediately, so that knowledge about the evolution and diversity of these pathogens can be incorporated into the research programs designed to ameliorate the conditions that they cause.

References
1. Darwin C: *The Descent of Man and Selection in Relation to Sex*. 6th edn. London: J. Murray; 1872.
2. Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappuoli R: ***Helicobacter pylori* virulence and genetic geography.** *Science* 1999, **284:**1328-1333.
3. Suerbaum S, Achtman M: **Evolution of *Helicobacter pylori*: the role of recombination.** *Trends Microbiol* 1999, **7:**182-184.
4. Achtman M, Azuma T, Berg DE, Ito Y, Morelli G, Pan Z-J, Suerbaum S, Thompson SA, van der Ende A, van Doorn L-J: **Recombination and clonal groupings within *Helicobacter pylori* from different geographical regions.** *Mol Microbiol* 1999, **32:**459-470.
5. Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, *et al.*: **Traces of human migrations in *Helicobacter pylori* populations.** *Science* 2003, **299:**1582-1585.
6. Agostini HT, Deckhut A, Jobes DV, Girones R, Schlunck G, Prost MG, Frias C, Pérez-Trallero E, Ryschkewitsch CF, Stoner GL: **Genotypes of JC virus in East, Central and Southwest Europe.** *J Gen Virol* 2001, **82:**1221-1331.
7. Sugimoto C, Hasegawa M, Kato A, Zheng H-Y, Ebihara H, Taguchi F, Kitamura T, Yogo Y: **Evolution of human polyomvirus JC: implications for the population history of humans.** *J Mol Evol* 2002, **54:**285-297.
8. Marshall BJ, Warren JR: **Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration.** *Lancet* 1984, **1:**1311-1315.
9. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, *et al.*: **The complete genome sequence of the gastric pathogen *Helicobacter pylori*.** *Nature* 1997, **388:**539-547.
10. Falush D, Kraft C, Taylor NS, Correa 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.
11. Rothenbacher D, Bode G, Berg G, Knayer U, Gonser T, Adler G, Brenner H: ***Helicobacter pylori* among preschool children and their parents: evidence of parent-child transmission.** *J Infect Dis* 1999, **179:**398-402.
12. Tindberg Y, Bengtsson C, Granath F, Blennow M, Nyrén O, Granström M: ***Helicobacter pylori* infection in Swedish school children: lack of evidence of child-to-child transmission outside the family.** *Gastroenterology* 2001, **121:**310-316.
13. Giles RE, Blanc H, Cann HM, Wallace DC: **Maternal inheritance of human mitochondrial DNA.** *Proc Natl Acad Sci USA*, 1980, **77:**6715-6719.
14. Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, Laos S, Parik J, Watkins WS, Dixon ME, *et al.*: **Deep common ancestry of Indian and western Eurasian mtDNA lineages**. *Curr Biol*, 1999, **9:**1331-1334.
15. Ingman M, Kaessmann H, Paabo S, Gyllensten U: **Mitochondrial genome variation and the origin of modern humans.** *Nature* 2000, **408:**708-713.
16. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **155:**945-959.
17. Hammer MF, Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegura SL: **Hierarchical patterns of global human Y-chromosome diversity.** *Mol Biol Evol* 2001, **18:**1189-1203.
18. Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL: **The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations.** *Ann Hum Genet* 2001, **65:**43-62.
19. Funk DJ, Helbling L, Wernegreen JJ, Moran NA: **Intraspecific phylogenetic congruence among multiple symbiont genomes.** *Proc R Soc Lond B Biol Sci* 2000, **267:**2517-2521.