

MIT Open Access Articles

Draft Genome Sequences of Helicobacter pylori Strains Isolated from Regions of Low and High Gastric Cancer Risk in Colombia

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Sheh, A., M. B. Piazuelo, K. T. Wilson, P. Correa, and J. G. Fox. "Draft Genome Sequences of Helicobacter Pylori Strains Isolated from Regions of Low and High Gastric Cancer Risk in Colombia." *Genome Announcements* 1, no. 5 (August 29, 2013): e00736–13–e00736–13.

As Published: <http://dx.doi.org/10.1128/genomeA.00736-13>

Publisher: American Society for Microbiology

Persistent URL: <http://hdl.handle.net/1721.1/89169>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: Creative Commons Attribution 3.0



Draft Genome Sequences of *Helicobacter pylori* Strains Isolated from Regions of Low and High Gastric Cancer Risk in Colombia

Alexander Sheh,^a M. Blanca Piazuelo,^b Keith T. Wilson,^{b,c,d,e} Pelayo Correa,^b James G. Fox^{a,f}

Division of Comparative Medicine, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA^a; Division of Gastroenterology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA^b; Department of Cancer Biology, Vanderbilt University School of Medicine, Nashville, Tennessee, USA^c; Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, Tennessee, USA^d; Veterans Affairs Tennessee Valley Healthcare System, Nashville, Tennessee, USA^e; Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA^f

The draft genome sequences of six Colombian *Helicobacter pylori* strains are presented. These strains were isolated from patients from regions of high and low gastric cancer risk in Colombia and were characterized by multilocus sequence typing. The data provide insights into differences between *H. pylori* strains of different phylogeographic origins.

Received 15 August 2013 Accepted 21 August 2013 Published 19 September 2013

Citation Sheh A, Piazuelo MB, Wilson KT, Correa P, Fox JG. 2013. Draft genome sequences of *Helicobacter pylori* strains isolated from regions of low and high gastric cancer risk in Colombia. *Genome Announc.* 1(5):e00736-13. doi:10.1128/genomeA.00736-13.

Copyright © 2013 Sheh et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to James G. Fox, jgfox@mit.edu.

Helicobacter pylori is a Gram-negative, microaerophilic, helix-shaped bacterium that colonizes the stomachs of at least half of the world's human population (1). In most cases, *H. pylori* can persist in the human stomach without health consequences, though it is a risk factor for chronic gastritis, gastric or duodenal ulcers, and gastric cancer (2).

In the present report, we announce the genome sequencing of six strains of *Helicobacter pylori*. These strains were isolated from three patients from the low-gastric-cancer-risk region of Tumaco, Colombia (strains PZ5004, PZ5024, and PZ5026), and three patients from the high-gastric-cancer-risk region of Pasto, Colombia (strains PZ5056, PZ5080, and PZ5086). These strains have been previously classified by multilocus sequence typing (MLST) and found to be of African (PZ5004 and PZ5024) and of European (PZ5026, PZ5056, PZ5080, and PZ5086) origins (3). Microarray analysis of these six strains demonstrated that clustering of transcriptomes also sorted strains based on their phylogeographic origins, with greater expression of virulence genes, such as *cagA* and *vacA*, in European strains (4). European strains also induced greater interleukin 8 (IL-8) expression while reducing apoptosis (4). The genome data provide insights about the genomic diversity of *H. pylori* from two sites with different incidence risks for gastric cancer and may help determine the underlying causes for differential transcription and virulence. The RNase-treated DNAs from the six isolates were sequenced using an Illumina MiSeq system as described previously (4). The 150-bp paired-end sequencing reads generated by MiSeq were assembled into contigs using Velvet (5). Sequences were annotated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (6). Pertinent statistics are summarized in Table 1. Overall, average GC contents of 38.65% and 1,660 coding DNA sequences (CDS) were found. The median coverage depth ranged from 128- to 246-fold.

Nucleotide sequence accession numbers. The genome sequences of *Helicobacter pylori* strains PZ5004, PZ5024, PZ5026,

TABLE 1 Pertinent statistics for sequenced Colombian strains

Strain	NCBI accession number	Genome length (bp) (Velvet)	No. of contigs (Velvet)	Median coverage depth (<i>n</i> -fold) (Velvet)	GC content (%) (Velvet)	No. of genes (PGAAP)	No. of CDS (PGAAP)
PZ5004	ASZF00000000	1,569,902	303	246	38.75	1,727	1,662
PZ5024	ASYS00000000	1,496,849	413	145	38.31	1,706	1,643
PZ5026	ASYT00000000	1,604,992	253	212	38.73	1,731	1,668
PZ5056	ASYU00000000	1,578,164	335	128	38.72	1,753	1,695
PZ5080	ASYV00000000	1,597,127	283	194	38.62	1,738	1,677
PZ5086	ASYW00000000	1,547,845	295	207	38.76	1,672	1,616

PZ5056, PZ5080, and PZ5086 were deposited at GenBank with the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants P01CA028842 (to J.G.F., K.T.W., and P.C.), P01CA026731 (to J.G.F.), and P30ES002109 (to J.G.F.).

REFERENCES

- Kusters JG, van Vliet AH, Kuipers EJ. 2006. Pathogenesis of *Helicobacter pylori* infection. *Clin. Microbiol. Rev.* 19:449–490.
- Fox JG, Wang TC. 2007. Inflammation, atrophy, and gastric cancer. *J. Clin. Invest.* 117:60–69.
- de Sablet T, Piazuelo MB, Shaffer CL, Schneider BG, Asim M, Chaturvedi R, Bravo LE, Sicinschi LA, Delgado AG, Mera RM, Israel DA, Romero-Gallo J, Peek RM, Jr, Cover TL, Correa P, Wilson KT. 2011. Phylogeographic origin of *Helicobacter pylori* is a determinant of gastric cancer risk. *Gut* 60:1189–1195.
- Sheh A, Chaturvedi R, Merrell DS, Correa P, Wilson KT, Fox JG. 2013. Phylogeographic origin of *Helicobacter pylori* determines host-adaptive responses upon coculture with gastric epithelial cells. *Infect. Immun.* 81:2468–2477.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829.
- Klimke W, Agarwala R, Badretin A, Chetvernin S, Ciuffo S, Fedorov B, Kiryutin B, O'Neill K, Resch W, Resenchuk S, Schafer S, Tolstoy I, Tatusova T. 2009. The National Center for Biotechnology Information's Protein Clusters Database. *Nucleic Acids Res.* 37:D216–D223.