



Genome Sequences of Three Colombian *Helicobacter pylori* Strains Isolated from Tolimense Patients

 Alix A. Guevara,^a  Roberto C. Torres,^b John J. Suárez,^a Fabian L. Castro-Valencia,^a Giovanna Parra,^c Javier Torres,^b Luis Carvajal Carmona,^d M. Magdalena Echeverry de Polanco,^a Mabel E. Bohórquez^{a,c}

^aResearch Group in Cytogenetic Phylogeny and Evolution of Populations, Departments of Sciences and Health Sciences, Tolima University, Tolima, Colombia

^bInfectious Diseases Research Unit, Instituto Mexicano del Seguro Social, Mexico City, Mexico

^cMedicine Program, Department of Health Sciences, Tolima University, Tolima, Colombia

^dGenome Center, Department of Biochemistry and Molecular Medicine, School of Medicine, University of California, Davis, California, USA

Alix A. Guevara and Roberto C. Torres contributed equally to this work. Author order was determined alphabetically.

ABSTRACT We present the complete genome sequences of three *Helicobacter pylori* strains isolated from patients who resided in Tolima Department, Colombia, diagnosed with chronic gastritis. The genomes present an average length of 1.6 Mbp and 1,546 genes and correspond to different *H. pylori* subpopulations.

Helicobacter pylori colonizes over 50% of the human population, and it is estimated that in Colombia, 70 to 80% of the adult population is infected (1). Although colonization of the gastric mucosa with *H. pylori* is the main known risk factor for gastric cancer, just a small percentage of infected people develop disease (2). Altered coevolution of the human host and its infecting *H. pylori* strain is associated with increased risk for premalignant gastric lesions (3). In Colombia, genomic studies of infecting *H. pylori* have shown a mixed ancestry between the European, African, and Asian origins, and some isolates diverge from the reported populations and constitute a different subgroup (4–6). We are still learning about the structure of *H. pylori* populations in Colombia, and isolates from more regions need to be studied. This report presents the draft genome sequences of three *H. pylori* strains isolated from patients with gastritis in the department of Tolima.

This study was approved by the Tolima University Bioethics Committee (act number 02 of 31 July 2018). Informed consent and histopathological diagnosis were recorded for all participants. Gastric biopsy specimens were collected from patients at Javeriana Clinic during upper gastrointestinal endoscopy as part of the treatment of dyspepsia. The gastric biopsy specimens were grown on blood agar supplemented with sodium carbonate, hydrolyzed casein, tryptone, activated carbon, 10% fresh horse blood serum, and 1% Vitox and *Campylobacter* selective supplements (Oxoid, Basingstoke, UK) at 37°C for 3 to 15 days under microaerophilic conditions. Each isolate was obtained from a single colony that was grown under the same conditions for 3 days, and genomic DNA was obtained from established growth using a DNeasy blood and tissue kit (Qiagen). Sequencing libraries were prepared with a TruSeq Nano DNA kit (Illumina), and genomes were sequenced using the 2 × 150 paired-end protocol of the Illumina NovaSeq platform (Macrogen, South Korea). Read data sets were trimmed to improve quality with the software package Trimmomatic version 0.39 (7). The genomes were assembled *de novo* with SPAdes version 3.13.1 (8) and annotated with Prokka version 1.12 (9). Ancestry of the samples was determined using fineSTRUCTURE version 4 (10) and ChromoPainter version 2 (11) based on the single nucleotide polymorphisms (SNPs) present in the core genome and using the default parameters. To calculate the

Citation Guevara AA, Torres RC, Suárez JJ, Castro-Valencia FL, Parra G, Torres J, Carvajal Carmona L, Echeverry de Polanco MM, Bohórquez ME. 2020. Genome sequences of three Colombian *Helicobacter pylori* strains isolated from Tolimense patients. Microbiol Resour Announc 9:e00117-20. <https://doi.org/10.1128/MRA.00117-20>.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Copyright © 2020 Guevara et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Mabel E. Bohórquez, mbohorquez@ut.edu.co.

Received 28 February 2020

Accepted 13 April 2020

Published 30 April 2020

TABLE 1 Summary of genome sequences reported

Strain	BioSample accession no.	GenBank accession no.	SRA accession no.	Diagnosis	Host origin department in Colombia	No. of contigs > 0 bp	No. of contigs > 1,000 bp	Genome size (bp) ^a	Coverage (X)	GC content (%) ^b	N ₅₀ (bp) ^b	No. of genes	Population
GCT27	SAMN13950472	CP048601	SRR11183158	Chronic active gastritis	Valle del Cauca	57	32	1,643,791	102	39.000	110,298	1,559	<i>hspAfrica1WAfricaNAmerica</i>
GCT43	SAMN13950473	CP048600	SRR11183157	Chronic active gastritis	Risaralda	68	35	1,642,398	102	39.046	98,340	1,566	<i>hspAfrica1SAfricaMiscAmerica</i>
GCT97	SAMN13950474	CP048599	SRR11183156	Chronic active gastritis	Tolima	60	40	1,656,586	103	38.847	94,297	1,569	<i>hspSWEuropeColombia</i>

^aIncluding contigs of ≥ 0 bp.^bBased on contigs of ≥ 500 bp.

population, we included as donors all those genomes included by Thorell et al. (5), Gutiérrez-Escobar et al. (4), and Muñoz-Ramírez et al. (6).

On average, the genomes have 39% GC content, 1.6 Mbp size, and 1,564 genes. Although the strains are from patients who reside in the same department, the population of each strain was different (Table 1); the GCT27 strain corresponds to a North American subpopulation with African ancestry, the strain GCT43 corresponds to a subpopulation including strains from different regions of Latin America with African ancestry, and the GCT97 strain corresponds to a Colombian subpopulation with European ancestry. These genomes provide information on the genetic population structure and the evolution of Colombian *H. pylori*.

Data availability. The sequence read files and the genome sequences of the strains have been deposited in the GenBank database under the accession numbers shown in Table 1. These sequences represent the first described versions (CP048601.1, CP048600.1, and CP048599.1).

ACKNOWLEDGMENTS

This study was supported by the Research Group in Cytogenetic Phylogeny and Evolution of Populations of Tolima University, Infectious Diseases Research Unit of the Instituto Mexicano del Seguro Social, and by the program for the formation of high-level human capital for the Department of Tolima of Colciencias and Tolima governorate (755-2016).

REFERENCES

1. Bravo LE, Cortés A, Carrascal E, Jaramillo R, García LS, Bravo PE, Badel A, Bravo PA. 2003. Helicobacter pylori: patología y prevalencia en biopsias gástricas en Colombia. *Colomb Med* 34:124–131.
2. Ghoshal UC, Chaturvedi R, Correa P. 2010. The enigma of Helicobacter pylori infection and gastric cancer. *Indian J Gastroenterol* 29:95–100. <https://doi.org/10.1007/s12664-010-0024-1>.
3. Kodaman N, Sobota RS, Mera R, Schneider BG, Williams SM. 2014. Disrupted human-pathogen co-evolution: a model for disease. *Front Genet* 5:290. <https://doi.org/10.3389/fgene.2014.00290>.
4. Gutiérrez-Escobar AJ, Trujillo E, Acevedo O, Bravo MM. 2017. Phylogenomics of Colombian Helicobacter pylori isolates. *Gut Pathog* 9:52. <https://doi.org/10.1186/s13099-017-0201-1>.
5. Thorell K, Yahara K, Berthenet E, Lawson DJ, Mikhail J, Kato I, Mendez A, Rizzato C, Bravo MM, Suzuki R, Yamaoka Y, Torres J, Sheppard SK, Falush D. 2017. Rapid evolution of distinct Helicobacter pylori subpopulations in the Americas. *PLoS Genet* 13:e1006546. <https://doi.org/10.1371/journal.pgen.1006546>.
6. Muñoz-Ramírez ZY, Mendez-Tenorio A, Kato I, Bravo MM, Rizzato C, Thorell K, Torres R, Aviles-Jimenez F, Camorlinga M, Canzian F, Torres J. 2017. Whole genome sequence and phylogenetic analysis show Helicobacter pylori strains from Latin America have followed a unique evolution pathway. *Front Cell Infect Microbiol* 7:50. <https://doi.org/10.3389/fcimb.2017.00050>.
7. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
8. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
9. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
10. Lawson DJ, Hellenthal G, Myers S, Falush D. 2012. Inference of population structure using dense haplotype data. *PLoS Genet* 8:e1002453. <https://doi.org/10.1371/journal.pgen.1002453>.
11. Hellenthal G, Busby GBJ, Band G, Wilson JF, Capelli C, Falush D, Myers S. 2014. A genetic atlas of human admixture history. *Science* 343:747–751. <https://doi.org/10.1126/science.1243518>.