genomeA
Journals.ASM.org

CrossMark
click for updates

Genome Sequencing of 10 *Helicobacter pylori* Pediatric Strains from Patients with Nonulcer Dyspepsia and Peptic Ulcer Disease

Alexandra Nunes,[a] Raquel Rocha,[b] Filipa F. Vale,[c] Luís Vieira,[d] Daniel A. Sampaio,[d] Ricardo Dias,[e] João P. Gomes,[a] Mónica Oleastro[b]

Bioinformatics Unit[a] and National Reference Laboratory of Gastrointestinal Infections,[b] Department of Infectious Diseases, National Institute of Health, Lisbon, Portugal; Host-Pathogen Interactions Unit, Research Institute for Medicines (iMed-ULisboa), Institute of Molecular Medicine, Faculty of Pharmacy, University of Lisbon, Lisbon, Portugal[c]; Innovation and Technology Unit, Department of Human Genetics, National Institute of Health, Lisbon, Portugal[d]; Center for Biodiversity, Functional and Integrative Genomics, Faculty of Sciences, University of Lisbon, Lisbon, Portugal[e]

We present draft genome sequences of 10 *Helicobacter pylori* clinical strains isolated from children. This will be important for future studies of comparative genomics in order to better understand the virulence determinants underlying peptic ulcer disease.

Address correspondence to Mónica Oleastro, monica.oleastro@insa.min-saude.pt

Long-term infection with *Helicobacter pylori* is the main cause of peptic ulcer disease (PUD), the most common severe condition associated with this bacterium. It affects 15–20% of infected adults, causing them considerable morbidity and mortality (1). Considering the high prevalence rates of infection worldwide, PUD is a major public health problem with high annual costs of treatment. In children with no other etiology for the disease, this rare event occurs shortly after infection and is thus a good model for the search of virulence biomarkers.

In this study, 10 strains were isolated from antral biopsies collected from Portuguese children with ages ranging from 7 to 15 years. Five specimens were obtained from nonulcer dyspepsia cases, while the remaining five were isolated from children exhibiting gastric or duodenal ulceration (Table 1). Briefly, biopsies were ground with a tissue homogenizer and inoculated onto both non- and selective media at 37°C under a microaerobic environment, for up to 14 days of incubation. *H. pylori* identification was performed according to conventional tests: colony and Gram stain morphologies, catalase, oxidase, and hydrolysis of urea. For each strain, total DNA was extracted using the QIAmp DNA minikit according to the manufacturer's instructions. After fragmentation, dual-indexed Illumina libraries of genomic DNA were prepared with the Nextera XT Index kit and then subjected to cluster generation and paired-end sequencing (2 × 250 bp or 2 × 300 bp) on a MiSeq Illumina platform. The resulting mean depth of coverage ranged from 39- to 156-fold (Table 1). After trimming with the FASTX tool (http://hannonlab.cshl.edu/fastx_toolkit), high-quality reads were *de novo* assembled using Velvet version 1/2/10 (2). Contigs were visually inspected with the Tablet version 1.14.04.10 graphical viewer (3) and properly corrected. Annotation was performed using the RAST server (4) and the NCBI Prokaryotic Genomes Annotation Pipeline version 2.3. Putative plasmids were identified by checking contigs >1,000 bp with high depth of coverage and evidence of circularity. Search for putative prophages was performed using the PHAge Search Tool (PHAST) (5).

Overall, the draft genomes of the 10 *H. pylori* strains varied from

TABLE 1 Description of the 10 *H. pylori* strains and of the Portuguese children from whom each was isolated[a]

Strain	Strain features						Patients[b]		
	Yr of isolation	Accession no. (no. of contigs)	Total size (bp)	Mean depth coverage (per base)	Putative plasmid (size)	No. of CDSs (pseudogenes)	Age (yr)	Gender	Clinical symptom
499/02	2002	JTDG00000000 (73)	1,666,743	126×	No	1,518 (66)	11	M	GU
1089/03	2003	JSUY00000000 (62)	1,587,159	134×	No	1,448 (75)	10	M	DU
1152/04	2004	JSUZ00000000 (59)	1,588,646	156×	No	1,449 (77)	10	M	DU
1198/04	2004	JSXT00000000 (57)	1,618,345	72×	pHPY1198 (2,794 bp)	1,485 (64)	15	M	DU
1846/05	2005	JSXV00000000 (58)	1,635,807	52×	pHPY1846 (3,161 bp)	1,500 (70)	13	M	DU
173/00	2000	JSXX00000000 (79)	1,562,065	32×	pHPY173 (4,194 bp)	1,438 (71)	14	M	NUD
207/99	1999	JSXU00000000 (47)	1,544,449	142×	No	1,395 (73)	7	F	NUD
228/99	1999	JSXY00000000 (54)	1,614,472	39×	pHPY228 (4,194 bp)	1,487 (59)	8	M	NUD
655/99	1999	JSXB00000000 (45)	1,615,666	130×	No	1,475 (63)	11	M	NUD
1786/05	2005	JSXW00000000 (41)	1,616,052	50×	pHPY1786 (9,960 bp)	1,467 (77)	11	M	NUD

[a] All strains are indicated by their collection number. All patients were from Portugal.
[b] M, male; F, female; NUD, nonulcer dyspepsia; DU, duodenal ulcer; GU, gastric ulcer.

1.54 to 1.67 Mb, with an average G+C content of 39% (Table 1). All ulcerogenic strains possessed a complete *cag* pathogenicity island, revealing a highly conserved gene content and gene order. Five different putative plasmids (sizes 2,794–9,960 bp), which are apparently cryptic based on BLAST results, were found for two duodenal ulcer and three nonulcer strains. With the exception of pHPY1198, all plasmids likely display more than 10 copies *per* chromosome. Two of the ulcerogenic strains, 1846/05 and 499/02, were found to carry a putative intact prophage of 16.7 Kb (GC% 36.7) and 15.3 Kb (GC% 37.4), respectively, both comprising 23 CDSs.

In order to better understand the genetic diversity among *H. pylori* strains that cause peptic ulcer disease in children, an in-depth comparative analysis between these ulcerogenic and nonulcerogenic draft genomes is under way and will be published in a subsequent report.

Nucleotide sequence accession numbers. The draft genome sequences of each *H. pylori* strain were deposited in GenBank under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

REFERENCES

1. **Testerman TL, Morris J.** 2014. Beyond the stomach: an updated view of *Helicobacter pylori* pathogenesis, diagnosis, and treatment. World J Gastroenterol **20:**12781–12808. http://dx.doi.org/10.3748/wjg.v20.i36.12781.
2. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.
3. **Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D.** 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform **14:**193–202. http://dx.doi.org/10.1093/bib/bbs012.
4. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9:**75. http://dx.doi.org/10.1186/1471-2164-9-75.
5. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS.** 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39:**W347–W352. http://dx.doi.org/10.1093/nar/gkr485.