🔓 | Bacteriology | Research Article

Resistome and virulome determination in *Helicobacter pylori* using next-generation sequencing with target-enrichment technology

Léo Gillet,[1] Lucie Bénéjat,[1] Quentin Jehanne,[1] Pierre-Louis Maunet,[1] Claudie Perreau,[1] Astrid Ducournau,[1] Johanna Aptel,[1] Marine Jauvain,[1,2] Philippe Lehours[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 12.

ABSTRACT The identification of *Helicobacter pylori* infection from gastric biopsy samples requires PCR or bacterial cultures. However, it is difficult to culture *H. pylori* because it is a fragile bacterium. Next-generation sequencing (NGS) allows direct assessment of the resistome and virulome. Here we describe a new NGS method for studying the resistome and virulome of *H. pylori* directly from gastric biopsies, based on enrichment analyses and targeted sequencing of *H. pylori* DNA. In all, 19 DNA samples from human gastric biopsies that tested positive for *H. pylori* were analyzed. The Agilent SureSelectXT target-enrichment protocol was used with a custom bait library prior to sequencing using the Agilent MagnisDx NGS Library Prep System. NGS sequencing was performed on the Illumina iSeq 100 sequencer using RNA probes for virulence, resistance, and molecular typing genes. The method yielded significant results with a limit of detection of around $1.8e^5$ CFU per mL *H. pylori*. Mutations in the *23S rDNA* sequence associated with macrolide resistance and in the quinolone resistance-determining region of *gyrase A* associated with levofloxacin resistance were correctly identified. The results of MLST phylogeny analyses performed after target-enrichment were consistent with those obtained via conventional Sanger sequencing. Among the *cagA*-positive isolates, the gene was detected correctly, and the *vacA* genotype was determined. In conclusion, our enrichment method enables rapid assessment of the resistome and virulome of *H. pylori* directly from fresh gastric biopsies.

IMPORTANCE *Helicobacter pylori*, a bacterium that infects at least 50% of the world population, is often treated by probabilistic antimicrobial therapies due to the lack of antimicrobial resistance data provided by clinical laboratories to clinicians. However, targeted antimicrobial therapies are increasingly recommended to achieve efficient eradication with a limited impact on the gut microbiota and with fewer adverse events for the patient. Recent advancements in next-generation sequencing strategies have opened new opportunities in the diagnosis of *H. pylori* infection. The significance of our research is the development of a novel next-generation sequencing strategy based on target-enrichment. This approach enables the identification of the resistome and the virulome of *H. pylori* directly from gastric biopsies, providing clinicians with a broad overview of therapeutic options.

KEYWORDS next-generation sequencing, *H. pylori*, resistome, gastric biopsy, target-enrichment

*H*elicobacter pylori is a Gram-negative, microaerophilic, coiled, and flagellated bacterium that exclusively infects the human stomach. Its role in gastroduodenal diseases, notably gastric ulcers, was established, and in 1994, the International Agency

[This article was published on 5 March 2025 with errors in the byline. The byline was corrected in the current version, posted on 11 March 2025.]

for Research on Cancer classified the bacterium as a class I carcinogen (1). It is the most common cause of cancers of infectious origin, which makes stomach cancer the sixth most common, and the fourth most deadly, cancer (2). Eradication of the bacterium cures gastritis and may alter the long-term course of complications of the disease (3).

Virulence factors of *H. pylori* enable its survival, colonization, and persistence in the acidic environment of the stomach. The *cagA* gene is part of a pathogenicity island (*cagPAI*) that codes for about 30 proteins that make up the type IV secretion system (T4SS). CagA is an oncoprotein and the major virulence factor that drives inflammation and proliferation and inhibits cell polarity and apoptosis. Its biological activity is linked to the presence of repeated C-terminal phosphorylation motifs (4). The *vacA* gene, which encodes the VacA cytotoxin, is present in all *H. pylori* strains but is not always expressed (5) due to genetic polymorphisms in the N-terminal signal sequences (s1 or s2), the middle region (m1 or m2), and the intermediate region (i1 or i2). The s1/m1 polymorphism is correlated with elevated risks of gastric ulcers and cancer (6).

Invasive and noninvasive diagnostic methods are available for the detection of *H. pylori*, depending on whether gastric biopsy is required (7). Noninvasive methods include the urea breath test, serology, and stool tests for *H. pylori* antigens. Invasive methods are the most sensitive and specific (8) and require the collection of gastric biopsies during upper digestive endoscopy. The biopsies can be sent to a pathology laboratory to analyze tissue lesions and detection of *H. pylori*. The disadvantage of histology is that, unlike noninvasive methods, it does not allow investigation of antimicrobial resistance.

The bacteriological examination of biopsies by culture on agar media has the highest specificity and enables antimicrobial susceptibility testing (AST) (8). However, not all laboratories can cultivate *H. pylori* because it grows slowly and its culture is problematic, requiring expertise. After extraction of DNA from a biopsy, molecular detection of the bacterium can be performed using commercially available, real-time PCR kits that can detect macrolide resistance more rapidly and more sensitively than culture (9, 10).

Given the increasing resistance of *H. pylori* to antimicrobials, *in vitro* AST prior to treatment is needed for antibiotic susceptibility-guided eradication therapy (7). If AST is not performed, empirical treatment will be initiated (11). Unfortunately, even in specialized laboratories, PCR-positive biopsies for *H. pylori* can remain negative in culture.

New methods using next-generation sequencing (NGS) have been developed to alleviate this problem. In 2018, Lauener et al. showed that it is possible to detect mutations responsible for resistance to clarithromycin, levofloxacin, rifampicin, and tetracycline using bacterial DNA and NGS, with more than 99% congruence with phenotypic results (12). An NGS method, commercially available only in the United States, applied directly to fresh gastric biopsies, paraffin-embedded biopsies, or stool samples enables assessment of the *H. pylori* resistome (13) (14). This important technological advance allows the implementation of targeted therapies based on molecular rather than phenotypic data.

Other NGS strategies use target-enrichment, which allows larger numbers of genes to be sequenced. The target-enrichment strategy has been used for whole-human-exome sequencing as well as research on human cancers and genetic diseases, and its flexibility allows the design of probes for any target (15). This technique involves the development of specific DNA or RNA probes that bind to the desired regions by complementarity, enabling specific sequencing. It can be used to isolate and specifically amplify regions of the genome of a pathogen in biological samples, which can be subjected to high-quality sequencing. It has been evaluated and validated for *Chlamydia trachomatis* (16) (17) and *Mycobacterium tuberculosis* (18).

In this study, we developed a target-enrichment technique to evaluate the resistome and virulome of *H. pylori* directly from gastric biopsies. The results were compared to those obtained using conventional PCR and Sanger sequencing methods. We validated the method for *H. pylori* diagnosis as it provides reliable results.

MATERIALS AND METHODS

Samples

In all, 19 gastric biopsies received at the National Reference Center for Campylobacter and Helicobacter (NRCCH) (http://www.cnrch.fr) between April and December 2022 were included in the study. All of the biopsies were from patients living in France, with a mean age of 49.5 ± 18 years and a sex ratio of approximately 0.46 (Table 1). The biopsies were confirmed to be positive for *H. pylori* in culture and via real-time PCR.

Upon receipt, according to our routine protocol, the biopsies were ground in 500 µL nutrient broth and stored at −80°C. A portion of the suspension was treated with 20 µL proteinase K (Roche Diagnostica, Meylan, France) in 180 µL ATL Qiagen buffer at 56°C for 3 h, and 200 µL was used on ground biopsies before DNA extraction on a MagNA Pure 96 system (Roche Diagnostics) and PCR on Eurogentec strips (Liège, Belgium) using an LC480 (Roche Diagnostics) (19). The threshold cycle values (Ct) determined via real-time PCR ranged from 18.3 to 27.5 (Table 1).

Bacterial culture was performed in parallel according to internal laboratory procedures (20). The strains were cultured on in-house blood agar containing antimicrobials as described previously (20). The media were incubated in a microaerobic atmosphere in a special workstation (Baker Ruskinn, Concept Ruskinn, Bridget, UK) at 37°C for 10 days. After 2 days, the plates were observed daily, and colonies were tested for oxidase, catalase, and urease and underwent morphological observation. AST was performed on Mueller–Hinton agar containing 10% horse blood, using Etest strips (bioMérieux) for clarithromycin, levofloxacin, rifampicin, and tetracycline. The cutoff values were those recommended by the French Microbiology Society Antibiogram Committee (CASFM) (CA-SFM/EUCAST: Société Française de Microbiologie Ed; 2024: pp. 1–177). For each strain, the MICs of clarithromycin, levofloxacin, rifampicin, and tetracycline were

TABLE 1 Description of cases

Sample	Sex	Age	Previous eradication	Pathology, symptoms	*H. pylori* ct	*23S rDNA* genotype determined by RT-PCR[a]	AST resistance marker[b,c]
1	M	60	No	Gastritis	26.6	A2142G or A2143G	Cla
2	M	84	Yes	MALT lymphoma	25.7	A2142G or A2143G	Cla, Lev
3	F	71	No	Gastritis	19.6	WT + A2142G or A2143G	Cla
4	F	34	No data	No data	19.5	WT	Lev
5	F	41	No	Other inflammatory bowel diseases (IBD)	19.5	WT	Cla
6	M	34	No	Gastritis	22.8	A2142G or A2143G	Cla
7	F	39	No	Gastroesophageal reflux disease (GERD)	23.4	WT	None
8	F	22	No	Gastritis	25.9	A2142C	Cla, Lev
9	F	31	No	Emesis	19.5	WT	None
10	F	49	Yes	Gastritis	18.3	WT + A2142G or A2143G	Cla
11	F	27	No	Gastritis	19.3	WT	None
12	F	62	No	Epigastralgia	19.8	A2142G or A2143G	Cla, Lev
13	M	41	No	Gastritis	19.1	WT	None
14	F	58	No	Dyspepsia	20.3	WT	None
15	F	32	Yes	Gastritis	23.3	WT + A2142G or A2143G	Cla
16	F	71	Yes	Gastritis	23.3	WT	Lev
17	F	54	No	Epigastralgia	20.1	WT + A2142G or A2143G	Cla
18	M	72	No	Positive serology	25	A2142G or A2143G	Cla, Lev
19	M	59	No	Gastritis	27.5	A2142G or A2143G	Cla, Lev

[a]WT, wild type.
[b]AST, antimicrobial susceptibility testing.
[c]Cla, clarithromycin; Lev, levofloxacin; Rif, rifampicin; Tet, tetracycline.

measured by two independent readers. Quality control was performed using *H. pylori* strain CCUG 17874.

End-point PCR

DNA was extracted from gastric biopsies to determine the *vacA* and *cagA* genotypes via PCR, to detect resistance markers (quinolone resistance-determining region [QRDR] of *gyrA* for levofloxacin resistance), and to extract genes for MLST typing, using primers described previously (https://pubmlst.org/organisms/helicobacter-pylori/primers) (Table S1). The reaction consisted of 5 µL 5 × PCR buffer (Promega, Madison, WI), 2.5 µL MgCl$_2$ (Promega), 0.5 µL dNTPs (Promega), 0.125 µL GoTaq G2 Hot-Start Polymerase (Promega), 1 µL each primer at 10 µM (Eurofins Genomics, Ebersberg, Germany), and 5 µL extracted DNA in a final volume of 25 µL. Amplification parameters consisted of 1 cycle at 95°C for 5 min, followed by 40 cycles at 95°C for 30 s, 54°C for 30 s, and 72°C for 90 s, and finally 1 cycle at 72°C for 5 min. The amplified fragments were resolved on 2% agarose gels containing green Midori Pre-stain (Nippon Genetics Europe, Düren, Germany). The PCR products of *gyrA* and MLST typing genes were sequenced by Eurofins Genomic. The sequences were analyzed using the NCBI Nucleotide BLAST (21) and aligned with MultAlin (22) if necessary.

Probe design for target-enrichment

An RNA probe library of 13,245 120-mer probes with a total size of 108 kbp was designed by Agilent using the selected regions of interest of the *H. pylori* reference strain J99 (23) and the reference strain B38 (24). The use of these genotypes resulted in a library containing probes that can hybridize to all s/i/m *vacA* genotypes. To enrich the library with probes specific for *cagA* EPIYA phosphorylation motifs, we used reference strains with multiple repeats of the C motif (involved in host interaction), NCTC 11637, and MT5114 (25) (which has the D motif common in Asian *H. pylori* strains), F32, and F17 (26, 27, 28). The probe library enabled the detection of mutations associated with resistance to clarithromycin (*23S rDNA*), levofloxacin (*gyrA*), rifampicin (*rpoB*), and tetracycline (*16S rDNA*) (Table 2). All of the regions of interest typically used in *H. pylori* MLST were targeted (29).

Automated capture-based library preparation

Most steps of the SureSelectXT HS2 workflow were automated using the Agilent MagnisDx library preparation platform from enzymatic fragmentation to the capture of hybridized libraries. The samples were diluted to 50 to 100 ng total DNA in a 14 µL input volume. Following Agilent's recommendation, the enzymatic fragmentation duration was set to 15 min and the read-length was two times 150 bp. After multiple calibration tests (data not shown), sequencing quality was maximized by setting 20 pre-capture and 24 post-capture PCR cycles.

NGS short-read sequencing

The fragment size distribution and molarity of the libraries were evaluated using the Agilent TapeStation 4150 System and High-Sensitivity D1000 assay. Each sample (2 µL) was diluted 1:10 and mixed with 2 µL High-Sensitivity D1000 Sample Buffer. The

TABLE 2 Mutations associated with phenotypic resistance to antibiotics

Gene	Mutations	Associated resistance
16S rDNA	AGA$_{926-928}$TTC, AGA$_{926-928}$ATC, AGA$_{926-928}$TTA, AGA$_{926-928}$TGC, AGA$_{926-928}$AGC, AGA$_{926-928}$ATA	Tetracycline
23S rDNA	A2142C, A2142G, A2142T, A2143C, A2143G	Clarithromycin
QRDR (*gyrA*)	D86N, N87K, N87I, N87Y, A88P, A88V, D91G, D91N, D91Y	Levofloxacin
rpoB	L525I, L525P, S526T, Q527H, Q527K, Q527R, F528E, D530A, D530E, D530G, D530N, D530V, T539A, H540C, H540N, H540Y, S545L, L547F, I586N, I586L, T588NL	Rifamycin

libraries were pooled at an equimolar concentration of 0.125 nM. Twenty microliters was harvested for multiplex sequencing on the Illumina iSeq 100 Sequencer at the NRCCH.

Bioinformatics workflow

Raw reads were mapped to the host genome using Bowtie2 v0.7.17 and the *Homo sapiens* GRCh38 reference genome. Unmapped reads were trimmed and cleaned using fastp v0.23.4 (30). A deduplication step, taking advantage of the duplex metabarcoding of the inserts, was used to generate consensus sequences and reduce false-positive calls. Reads specific to *H. pylori* were selected using Kraken2 v1.2.2 (31), a taxonomic classification system that employs exact k-mer matches, with the PlusPF database, which contains RefSeq bacterial, plasmid, viral, human, protozoal, and fungal sequences. The reads were mapped onto *H. pylori* reference genome J99 (assembly ASM98269v1) (23) using Bowtie2 (32), and the consensus sequences of the targeted regions were extracted using Freebayes v1.3.6 (33) and bcftools v1.1.19 (34). A homemade suite of tools was used with those consensus sequences to detect genomic markers and assess the resistome. In parallel, a *de novo* assembly was generated using the NCBI SAUTE v2.5.1 (35), to generate complete sequences of *cagA* and *vacA*. MLST sequences were concatenated and used as input for STRUCTURE v2.3.4 (36), attributing each sample to a range of annotated populations. They were compared to the automatically updated and annotated database of allele sequences, available on PubMLST (37). Phylogenetic trees were generated using RAxML-NG v1.2.2 (38) using the GTR + Gamma model. A set of homemade scripts was developed to compile the results and generate PDF and HTML reports, as well as a circular genome representation, using the Circos graphic library v0.69–9 (39).

Limit of detection

One *H. pylori*-negative gastric biopsy sample (by PCR and culture) was ground, homogenized, and spiked with *H. pylori* CCUG 17874 (40, 41) at 1.8 to $1.8e^7$ CFU per mL and enzymatically digested overnight at 56°C in 180 µL ATL buffer (Qiagen) and 20 µL proteinase K (Roche Diagnostics). The libraries were prepared three times and sequenced using the same protocol as the clinical samples to identify cutoff Ct and bacterial charge values that ensured a good sequencing quality and accurate evaluation of resistome and virulome.

Serial dilution of mixed samples

To evaluate the ability of our technique to detect mixed infection with susceptible (S) and resistant (R) strains, artificial mixes of the clarithromycin- and levofloxacin-susceptible *H. pylori* reference strain CCUG 17874 with clarithromycin- and levofloxacin-resistant clinical isolate (AST data + PCR data) were prepared. Each strain was suspended at a turbidity of 1 McFarland in brain–heart infusion broth and transferred to 1:1 (500 µL each strain), 1:5 (800 µL the susceptible strain, 200 µL the resistant strain), 1:10 (900 µL S, 100 µL R), and 1:100 (990 µL S, 10 µL R) mixtures in duplicate. The mixtures were centrifuged for 10 min at 13,000 RPM before extraction, following the same protocol as that used for the spiked samples. Mutations in *23S rDNA* were detected using real-time PCR, and mutations in the QRDR region of *gyrA* were detected via PCR and Sanger sequencing, using the primer pairs given in Table S1.

Whole-genome sequencing of *cagA*-positive samples

cagA-positive strains were sent for whole-genome sequencing at the GenoBioMICS sequencing platform (CHU Henri Mondor, Créteil, France) to determine the EPIYA motifs of CagA. Total DNA was extracted using mechanical (bead-beating), enzymatic, and chemical methods on the Revvity Chemagic Prime instrument. Libraries were prepared using the Illumina DNA Prep Kit on the Tecan DreamPrep NGS system and sequenced using an Illumina NovaSeq 6000 sequencer. Bioinformatic analysis was performed using the pipeline described above, with the exception of the genome assembly step, which

was conducted using NCBI SKESA v2.5.1 (35). A supplementary assembly annotation step was added to extract *cagA* sequences using Prokka v1.14.5 (42).

RESULTS

Sensitivity and limit of detection

The number of reads attributed to *H. pylori* and the coverage of regions at minimum depth (5× +metabarcoding consensus) are shown in Fig. 1. The percentage of reads attributed to *H. pylori* was <10% for DNA extracted from gastric biopsies spiked with <$1.8e^5$ UFC per mL. As expected, the reduction in bacterial load was accompanied by an increase in the percentage of human reads Fig. 2. The sequencing data of the spiked gastric biopsies and the samples are listed in Table S2.

Limits of detection for *23S rDNA* and *gyrA* mutations in mixed populations

To assess the sensitivity of the technique for detecting antibiotic-resistant mixed populations, a series of mixtures of different *H. pylori* strains were prepared (one sensitive and the other resistant to clarithromycin and levofloxacin) and sequenced using the same library preparation protocol. At a 1:1 susceptible:resistant strain ratio, approximately half of the sequencing reads at position A2143 of the *23S rDNA* and G271 of *gyrA* had SNPs associated with resistance to clarithromycin (A2143G) and levofloxacin (D91Y), respectively. The other sequencing reads corresponded to wild-type genotypes.

As the concentration of resistant bacteria in the mixtures decreased, the proportion of mutated alleles decreased. Mutated alleles were detected at a 1:10 susceptible:resistant strain ratio. At a 1:100 ratio, few reads had mutations associated with resistance to both antimicrobials; these were not detected using our technique (Table 3).
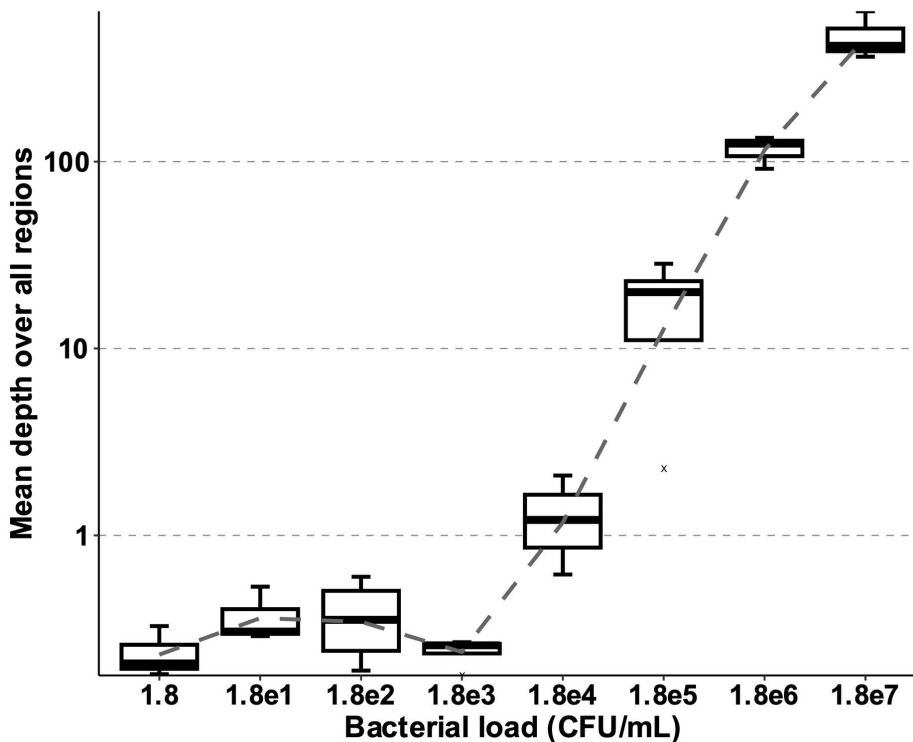

FIG 1 Limit of detection in serial dilutions of gastric biopsies spiked with the *H. pylori* reference strain CCUG 17874. Values are average mean depths overall targets, with a threshold of 10 reads per base.
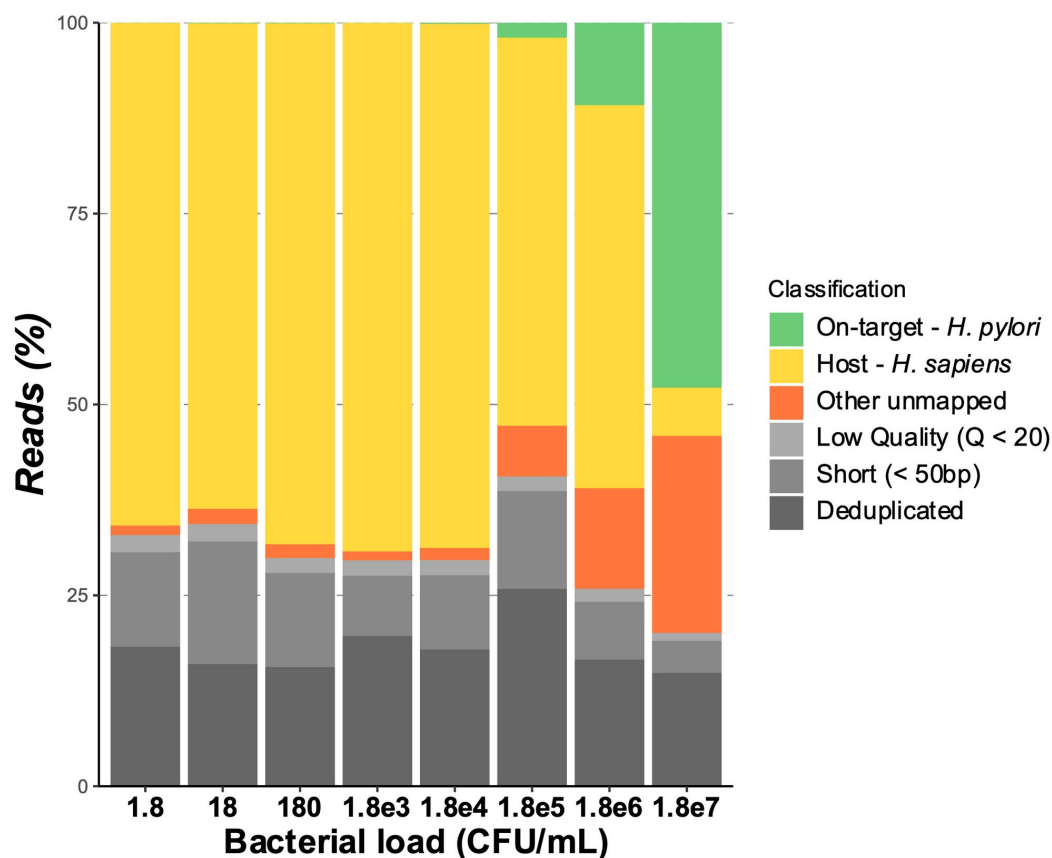
FIG 2 Classification of sequencing reads depending on the bacterial load (CFU per mL). Although other *unmapped* reads may correspond to other sequenced organisms either by accidental contamination or capture of conserved genomic sequences, metagenomic classification of raw reads using Kraken2 did not reveal the presence of particular species but detected a handful of species at low levels and unclassifiable reads.

Detection of antimicrobial resistance markers in clinical isolates

To evaluate our method using clinical specimens, 19 samples were tested. All samples tested positive via real-time PCR for *H. pylori* with Ct values < 27 (Table 1), consistent with the sensitivity mentioned above. Mutations in *23S rDNA* identified via target-enrichment were compared to those determined through real-time PCR. The correlation was perfect in 17 of the 19 samples tested, and all genotypes were correctly identified (Table 4). The main mutation identified was A2143G. Base depth was insufficient for samples 2 and 7. A mixed population of wild-type and mutant *H. pylori* was detected in samples 10, 12, 15, and 17, with the mutant population representing 10% to 70% of the called bases. Surprisingly, only A2143G was detected in sample 3, and the double population originally detected via real-time PCR was absent. The presence of a mixed population (S + R) was confirmed *in vitro* via AST using Etest on the strain isolated from this sample and by genome sequencing (WT +A2143G 60%) (data not shown). In sample 12, the mutations A2142G (85% of bases called) and A2143G (13% of bases called) were identified, indicating a mixture of clarithromycin-resistant *H. pylori* populations.

The mutations in the QRDR region of *gyrA* identified via target-enrichment were compared to those identified through endpoint PCR and Sanger sequencing. Excluding sample 11 with insufficient base depth, the correlation was perfect in all the remaining cases (Table 5), and all of the mutations were correctly identified. The mutations associated with levofloxacin resistance were D91N (two samples), N87K (four samples), and N87I (one sample). In samples 5 and 19, 53% and 25% of the called bases were identified, respectively. This indicates the presence of a mixture of wild-type and mutant

TABLE 3 Percentages of mutant alleles in mixtures of susceptible and resistant strains

Theoretical ratio (percentage of resistant individuals)	23S A2143G allele percentage	Base depth at position	gyrA D91Y allele percentage	Base depth at position
1:1 (50%)	49%	318	49%	582
1:5 (20%)	13%	264	13%	635
1:10 (10%)	5%	177	6%	536
1:100 (1%)	<0.01%	445	<0.01%	517

H. pylori. No evidence of mixed populations was found in these samples via *in vitro* AST using Etest. No mutation of *gyrA* associated with levofloxacin resistance was missed in these two samples using target-enrichment.

No resistance marker for tetracycline and rifampicin was found in the 19 samples (data not shown). This was expected because no *H. pylori* isolates resistant to these drugs were included in this study. However, the average 5× coverages over the sites searched for potential antimicrobial markers in the 19 samples for *16S rDNA* and *rpoB* were 94.7% (± 23%) and 84.7% (± 36%), respectively (Table S3). This indicates perfect sequencing of genes involved in tetracycline and rifampicin resistance. The exceptions were samples 2, 7, 8, and 16, which had poor sequencing quality over the targeted regions.

MLST of clinical samples

Geographic attribution of samples was computed from the concatenated sequences of the *H. pylori* typing scheme using STRUCTURE and a database of annotated sequences from PubMLST. The results of end-point PCR followed by Sanger sequencing were compared to those of target-enrichment (Table 6). The correlation between the two approaches was perfect. Most of the biopsies (14 of 19, 73.7%) contained hpEurope *H. pylori*. Three biopsies contained hpEurope and hpAfrica1 hybrids (3 of 19, 15.8%) and *H. pylori* strains with hpAfrica1 represented 25% to 50% of the profile. The two other cases (2 of 19, 10.5%) were attributed to hpAfrica1.

TABLE 4 23S rDNA mutations detected via real-time PCR and target-enrichment analysis[a]

Sample	Observed mutation by RT-PCR	Observed mutation by target-enrichment
1	A2142G or A2143G	A2143G
2	A2142G or A2143G	Insufficient base depth
3	WT + A2142G or A2143G	A2143G
4	WT	WT
5	WT	WT
6	A2142G or A2143G	A2143G
7	WT	Insufficient base depth
8	A2142C	A2142C
9	WT	WT
10	WT +A2142G or A2143G	WT + A2143G (10% of called bases)
11	WT	WT
12	A2142G or A2143G	A2142G (85%), A2143G (13%)
13	WT	WT
14	WT	WT
15	WT + A2142G or A2143G	WT + A2143G (70%)
16	WT	WT
17	WT + A2142G or A2143G	WT + A2143G (32%)
18	A2142G or A2143G	A2143G
19	A2142G or A2143G	A2143G

[a]For samples with mixed populations, such as samples 10, 12, 15, and 17, the percentages of variants are shown.

TABLE 5 *gyrA* mutations in gastric samples detected via end-point PCR and Sanger sequencing and target-enrichment analysis[a]

Sample	Observed mutation by PCR & Sanger sequencing	Observed mutation by target-enrichment
1	WT	WT
2	D91N	D91N
3	WT	WT
4	N87K	N87K
5	N87K	WT + N87K (53%)
6	WT	WT
7	WT	WT
8	N87K	N87K
9	WT	WT
10	WT	WT
11	WT	Insufficient base depth
12	N87I	N87I
13	WT	WT
14	WT	WT
15	WT	WT
16	N87K	N87K
17	WT	WT
18	D91N	D91N
19	N87K	WT + N87K (25%)

[a]For samples with mixed populations, such as samples 5 and 19, the percentages of variants are shown.

Virulence typing

There was no discrepancy in the detection of *cagA* via target-enrichment compared to end-point PCR (Table 7). The *cagA* sequence assembly was of insufficient length to detect all EPIYA phosphorylation motifs in samples 1 and 16, and discrepancies in the number of EPIYA motifs were found for samples 11 and 12 (1C versus 2C motifs).

For *vacA* genotypes, the correlation was perfect for 15 of the 19 samples. Only the m2 genotype was identified in sample 8. No prediction was made for samples 2, 7, and 19 because of insufficient assembly length.

DISCUSSION

Our NGS approach using target-enrichment technology yielded reliable results for the *H. pylori* resistome and virulome of gastric biopsy specimens.

The use of NGS with gastric biopsies has previously been evaluated (13, 14). However, using classical library preparation approaches, the first barrier to sufficient sequencing quality is the extraction of enough good-quality DNA suitable for NGS. Our strategy requires 50 to 100 ng of total DNA extracted from a biopsy sample, which is feasible in routine practice. The sensitivity of the technique limits its use to biopsies containing >10^4 CFU/mL, which corresponds to a Ct value of 27 in real-time PCR routinely used by the NRCCH (19). This detection limit would enable the characterization of approximately 68% of biopsies received annually by our reference center that are PCR-positive but culture-negative (data not shown). A large amount of human DNA is present in gastric biopsies and is amplified with *H. pylori* DNA. Therefore, depletion of human DNA using techniques such as nuclease digestion (43, 44) or methylation-based separation (45) prior to *H. pylori* capture could optimize the amplification of bacterial DNA.

In 2014, Christiansen et al. published a similar technique using the same preparation protocol and probes targeting the whole genome of *Chlamydia trachomatis*, the culture of which is difficult. The detection sensitivity was 10-fold higher than the culture of urine and vaginal swabs (16, 17). Moreover, adding multiple reference genomes of *C. trachomatis* facilitated the capture of more diverse strains (46), prompting us to add

TABLE 6 MLST typing and hpGroup determined via end-point PCR plus Sanger sequencing and target-enrichment analysis using STRUCTURE

Sample	Assigned geographic profile using Sanger sequencing	Assigned geographic profile using target-enrichment
1	*hpEurope* (65%), *hpAfrica1* (35%)	*hpEurope* (70%), *hpAfrica1* (30%)
2	*hpEurope*	*hpEurope*
3	*hpEurope*	*hpEurope*
4	*hpEurope*	*hpEurope*
5	*hpAfrica1*	*hpEurope*
6	*hpEurope*	*hpEurope*
7	*hpEurope*	*hpEurope*
8	*hpEurope*	*hpEurope*
9	*hpAfrica1*	*hpAfrica1*
10	*hpEurope* (50%), *hpAfrica1* (50%)	*hpEurope* (50%), *hpAfrica1* (50%)
11	*hpEurope*	*hpEurope*
12	*hpAfrica1*	*hpAfrica1*
13	*hpAfrica1*	*hpEurope* (75%), *hpAfrica1* (25%)
14	*hpEurope*	*hpEurope*
15	*hpEurope*	*hpEurope*
16	*hpEurope*	*hpEurope*
17	*hpEurope*	*hpEurope*
18	*hpEurope*	*hpEurope*
19	*hpEurope*	*hpEurope*

probes that would ensure the detection of *cagA* genotypes found in Asian strains. Application of this technique to *Mycobacterium tuberculosis* enables the sequencing of strains directly from sputum samples, with accuracy sufficient to identify heterozygous positions at up to a 95:5 ratio and mixed infections at up to a 90:10 ratio.

TABLE 7 Detection of the *cagA* and *vacA* genotypes via target-enrichment analysis

Sample	*cagA* presence determined by PCR on strains and phosphorylation motifs by WGS	*cagA* presence and phosphorylation motifs determined by target-enrichment	*vacA* genotypes using Sanger sequencing	*vacA* genotypes using target-enrichment
1	Pos,[a] ABC	Pos, no motif	s2/i2/m2	s2/i2/m2
2	Neg[b]	Neg	s2/i2/m2	Insufficient quality
3	Neg	Neg	s2/i2/m2	s2/i2/m2
4	Neg	Neg	s2/i2/m2	s2/i2/m2
5	Pos, ABC	Pos, ABC	s1/i1/m1	s1/i1/m1
6	Neg	Neg	s2/i2/m2	s2/i2/m2
7	Neg	Neg	s2/i2/m2	Insufficient quality
8	Neg	Neg	s1/i2/m2	m2
9	Neg	Neg	s2/i2/m2	s2/i2/m2
10	Pos, ABC	Pos, ABC	s1/i1/m2	s1/i1/m2
11	Pos, BC	Pos, BCC	s1/i2/m2	s2/i2/m2
12	Pos, ABC	Pos, ABCC	s1/i1/m1	s1/i1/m1
13	Pos, ABC	Pos, ABC	s1/i1/m1	s1/i1/m1
14	Neg	Neg	s2/i2/m2	s2/i2/m2
15	Neg	Neg	s2/i2/m2	s2/i2/m2
16	Pos, ABC	Pos, no motif	s1/i1/m1	s1/i1/m1
17	Pos, ABC	Pos, ABC	s1/i1/m2	s1/i1/m2
18	Neg	Neg	s2/i2/m2	s2/i2/m2
19	Neg	Neg	s2/i2/m2	Insufficient quality

[a]Pos: positive.
[b]Neg: negative.

Our results are in line with these previous studies because our technique detected a mutation in a mixture in which 10% of the population was resistant to clarithromycin. This is particularly useful for diagnosis prior to targeted eradication therapy. The ability to detect mixed populations of resistant and susceptible *H. pylori* is of particular interest because of its ability to rapidly acquire resistance to antimicrobials due to its high point-mutation frequency (47). This is also an advantage compared to conventional AST or Sanger sequencing techniques, which are less efficient at detecting mixed populations. NGS approaches can provide important information for clinicians for targeted eradication strategies using tailored treatment.

The results of resistome determination via target-enrichment analysis were consistent with those of Sanger sequencing and with the resistance phenotypes obtained *in vitro*. The technique was particularly reliable for markers of resistance to clarithromycin, levofloxacin, rifampicin, and tetracycline. The major mutations detected by Sanger and target-enrichment sequencing were A2143G in *23S rDNA* (associated with clarithromycin resistance) and N87K (associated with levofloxacin resistance), as described by others (48). The absence of resistance to rifampicin and tetracycline is expected, given the low prevalence of resistance to these two antibiotics in *H. pylori* in France (NRCCH annual report; https://www.cnrch.fr/).

Our pipeline database was based on the major mutations in *23S rDNA*, *gyrA*, *rpoB*, and *16S rDNA* associated with clarithromycin, levofloxacin, rifabutin, and tetracycline resistance, respectively. It was also enriched by *rpoB* mutations in French rifampicin-resistant clinical isolates collected over the past 8 years by our reference center. The database was created in collaboration with the *Helicobacter pylori* Genome Project (https://dceg.cancer.gov/research/how-we-study/genomic-studies/h-pylori-genome-project), a global multidisciplinary consortium to study the population structure, drug resistance, and pathogenesis of *H. pylori* (49). Therefore, we believe that our database enables reliable detection of the resistome of *H. pylori* for clarithromycin, levofloxacin, rifabutin, and tetracycline, and at a larger scale than described previously (13, 14).

The bait design involved mutations in *rdxA*, *frxA*, and *fdxB* (associated with metronidazole resistance (50, 51) as well as *pbp1* (amoxicillin resistance [52]). Nevertheless, in the absence of proven associations with antimicrobial resistance, we do not present the results for these genes. Our strategy and pipeline can be applied upon publication of mutations or genes associated with resistance to these antimicrobials and the baits will be adapted if necessary.

The lack of predictions for amoxicillin and metronidazole, unlike other studies (13, 14), is not a major problem for clinicians because resistance in these is rare (amoxicillin) or is counterbalanced by an increased dose (amoxicillin can be adapted to body weight) and/or duration of eradication treatment (metronidazole) (53).

Nevertheless, our target-enrichment approach to the virulome needs to be modified. Although *cagA* detection was good, improvements are planned to increase the sensitivity of the technique for detecting CagA phosphorylation motifs. We had to sequence all *cagA*-positive isolates because the multiplex PCR of Argent et al. (54) provided less reliable results (data not shown) than whole-genome sequencing. Baits designed for the capture of *htrA*, a major virulence factor and an essential gene for the survival of *H. pylori* (55–57), could be added to the library to provide further insight into the virulence profile. Although we were able to detect most *vacA* genotypes, improvement is needed because of the high diversity of this gene. The design of new probes for less well-covered areas is therefore needed. Nevertheless, the evaluation of the virulome is of lesser clinical value than the assessment of the resistome. In addition, no consensus guidelines suggest that antibiotic therapy be adjusted based on the virulome. However, determination of the virulome, in particular *cagA* (and the hypervirulent *htrA* genotype), may be of clinical relevance to *H. pylori* infections in patients with gastric adenocarcinoma, a first-degree family history of gastric cancer, or advanced preneoplastic lesions.

MLST classification of the infecting strain can be seen as "icing on the cake." The majority of patients in this study were infected with European or African strains, in proportions usually observed in the French population. The correlation between host genotype and infecting strain has been proposed by De Sablet as being, in the event of discordance, a negative factor in disease progression (58). It would be interesting to correlate the *cagA* genotype, in particular the presence of an Asian motif, with the MLST profile. Asian strains are more virulent than others (because of the D motif) and are associated with a higher frequency of severe disease.

Our results lead to an important question: where should the target-enrichment strategy for the routine diagnosis of *H. pylori* infection be placed? In our laboratory, access to culture is not a problem, and its performance is good (up to 10%–15% false-negative results); therefore, the target-enrichment strategy could be reserved for PCR-positive but culture-negative biopsies from patients for whom eradication is expected to be of marked clinical benefit (those with ulcers, preneoplastic lesions, cancer, or familial history of cancer). This would provide information important for selecting the optimum *H. pylori* eradication strategy. For users who want to bypass *H. pylori* culture, the implementation of our method would be feasible using Agilent's MagnisDx NGS Library Prep System.

In conclusion, we describe a high-performance NGS approach for investigating the resistome and virulome of *H. pylori* from gastric biopsy samples.

ACKNOWLEDGMENTS

AUTHOR AFFILIATIONS

[1]CHU de Bordeaux, CNR des Campylobacters et des Hélicobacters, Bordeaux, France
[2]INSERM U1312, UMR BRIC-Team 4, Bordeaux, France

AUTHOR ORCIDs

Léo Gillet http://orcid.org/0009-0008-8854-6732
Philippe Lehours http://orcid.org/0000-0001-6077-1825

AUTHOR CONTRIBUTIONS

Léo Gillet, Methodology, Software, Validation, Writing – original draft | Lucie Bénéjat, Formal analysis, Methodology, Visualization, Writing – original draft | Quentin Jehanne, Methodology, Software, Validation, Visualization, Writing – review and editing | Pierre-Louis Maunet, Methodology | Claudie Perreau, Methodology | Astrid Ducournau, Methodology | Marine Jauvain, Validation, Writing – review and editing.

DATA AVAILABILITY

Raw sequencing data generated using the target-enrichment library preparation method as well as whole-genome sequencing data of *cagA*-positive samples are available in the NCBI Sequence Read Archive under BioProject number PRJNA1149679.

ETHICS APPROVAL

All diagnostic methods were performed routinely. All patients were investigated in a hospital and private clinic setting, according to good clinical practices. In this routine process, consent for the endoscopic procedure and biopsy collection is always provided in writing and is kept in the patient's medical record. Informed consent for the use of gastric DNA was not obtained from the patients. Therefore, to protect subject anonymity, all information that could indirectly identify patient data was removed from the present study.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Table S1 (Spectrum03298-24-s0001.docx). PCR primers used as controls.
Table S2 (Spectrum03298-24-s0002.xlsx). Mean depth and coverage over antimicrobial-resistance marker sites.
Table S3 (Spectrum03298-24-s0003.xlsx). Fivefold coverage over sites searched for potential antimicrobial markers for 16S rDNA, 23S rDNA, gyrA, and rpob.

REFERENCES

1. Cancer IA for R on. 1994. Schistosomes, liver flukes, and *Helicobacter pylori*, p 177–241. In Monographs WI (ed), Monographs on the evaluation of carcinogenic risk to humans: schistosomes liver flukes, and *Helicobacter pylori*. Lyon, France.
2. Mégraud F, Bessède E, Varon C. 2015. *Helicobacter pylori* infection and gastric carcinoma. Clin Microbiol Infect 21:984–990. https://doi.org/10.1016/j.cmi.2015.06.004
3. Sugano K, Tack J, Kuipers EJ, Graham DY, El-Omar EM, Miura S, Haruma K, Asaka M, Uemura N, Malfertheiner P, faculty members of Kyoto Global Consensus Conference. 2015. Kyoto global consensus report on *Helicobacter pylori* gastritis. Gut 64:1353–1367. https://doi.org/10.1136/gutjnl-2015-309252
4. Tran SC, Bryant KN, Cover TL. 2024. The *Helicobacter pylori cag* pathogenicity island as a determinant of gastric cancer risk. Gut Microbes 16:2314201. https://doi.org/10.1080/19490976.2024.2314201
5. Atherton JC, Sharp PM, Cover TL, Gonzalez-Valencia G, Peek RM Jr, Thompson SA, Hawkey CJ, Blaser MJ. 1999. Vacuolating cytotoxin (vacA) alleles of *Helicobacter pylori* comprise two geographically widespread types, m1 and m2, and have evolved through limited recombination. Curr Microbiol 39:211–218. https://doi.org/10.1007/s002849900447
6. Rhead JL, Letley DP, Mohammadi M, Hussein N, Mohagheghi MA, Eshagh Hosseini M, Atherton JC. 2007. A new *Helicobacter pylori* vacuolating cytotoxin determinant, the intermediate region, is associated with gastric cancer. Gastroenterology 133:926–936. https://doi.org/10.1053/j.gastro.2007.06.056
7. Malfertheiner P, Megraud F, Rokkas T, Gisbert JP, Liou J-M, Schulz C, Gasbarrini A, Hunt RH, Leja M, O'Morain C, Rugge M, Suerbaum S, Tilg H, Sugano K, El-Omar EM. 2022. Management of *Helicobacter pylori* infection: the maastricht VI/Florence consensus report. Gut 71:1724–1762. https://doi.org/10.1136/gutjnl-2022-327745
8. Mégraud F, Lehours P. 2007. *Helicobacter pylori* detection and antimicrobial susceptibility testing. Clin Microbiol Rev 20:280–322. https://doi.org/10.1128/CMR.00033-06
9. Bénéjat L, Ducournau A, Lehours P, Mégraud F. 2018. Real-time PCR for *Helicobacter pylori* diagnosis. The best tools available. Helicobacter 23:e12512. https://doi.org/10.1111/hel.12512
10. Bénéjat L, Ducournau A, Martins CD, Bessède E, Lehours P. 2023. RIDAGENE *Helicobacter pylori* PCR on the ELITe InGenius system. Eur J Clin Microbiol Infect Dis 42:593–596. https://doi.org/10.1007/s10096-023-04563-3
11. Schuetz AN, Theel ES, Cole NC, Rothstein TE, Gordy GG, Patel R. 2024. Testing for *Helicobacter pylori* in an era of antimicrobial resistance. J Clin Microbiol 62:e0073223. https://doi.org/10.1128/jcm.00732-23
12. Lauener FN, Imkamp F, Lehours P, Buissonnière A, Benejat L, Zbinden R, Keller PM, Wagner K. 2019. Genetic determinants and prediction of antibiotic resistance phenotypes in *Helicobacter pylori*. J Clin Med 8:53. https://doi.org/10.3390/jcm8010053
13. Argueta EA, Alsamman MA, Moss SF, D'Agata EMC. 2021. Impact of antimicrobial resistance rates on eradication of *Helicobacter pylori* in a US population. Gastroenterology 160:2181–2183. https://doi.org/10.1053/j.gastro.2021.02.014
14. Herzlinger M, Dannheim K, Riaz M, Liu E, Bousvaros A, Bonilla S. 2023. *Helicobacter pylori* antimicrobial resistance in a pediatric population from the New England region of the United States. Clin Gastroenterol Hepatol 21:3458–3460. https://doi.org/10.1016/j.cgh.2023.02.026
15. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. 2010. Target-enrichment strategies for next-generation sequencing. Nat Methods 7:111–118. https://doi.org/10.1038/nmeth.1419
16. Christiansen MT, Brown AC, Kundu S, Tutill HJ, Williams R, Brown JR, Holdstock J, Holland MJ, Stevenson S, Dave J, Tong CW, Einer-Jensen K, Depledge DP, Breuer J. 2014. Whole-genome enrichment and sequencing of *Chlamydia* trachomatisdirectly from clinical samples. BMC Infect Dis 14:591. https://doi.org/10.1186/s12879-014-0591-3
17. Seth-Smith HMB, Bénard A, Bruisten SM, Versteeg B, Herrmann B, Kok J, Carter I, Peuchant O, Bébéar C, Lewis DA, Puerta T, Keše D, Balla E, Zákoucká H, Rob F, Morré SA, de Barbeyrac B, Galán JC, de Vries HJC, Thomson NR, Goldenberger D, Egli A. 2021. Ongoing evolution of *Chlamydia trachomatis* lymphogranuloma venereum: exploring the genomic diversity of circulating strains. Microb Genom 7:000599. https://doi.org/10.1099/mgen.0.000599
18. Lozano N, Lanza VF, Suárez-González J, Herranz M, Sola-Campoy PJ, Rodríguez-Grande C, Buenestado-Serrano S, Ruiz-Serrano MJ, Tudó G, Alcaide F, Muñoz P, García de Viedma D, Pérez-Lago L. 2021. Detection of minority variants and mixed infections in *Mycobacterium tuberculosis* by direct whole-genome sequencing on noncultured specimens using a

specific-DNA capture strategy. mSphere 6:e0074421. https://doi.org/10.1128/mSphere.00744-21

19. Bénéjat L, Ducournau A, Domingues-Martins C, Lecoeur M, Blosse A, Mégraud F, Bessède E, Lehours P. 2021. Adaptation of an in-house PCR for the detection of *Helicobacter pylori* and the mutations associated with macrolide resistance into ready-to-use PCR microwell strips. Helicobacter 26:e12855. https://doi.org/10.1111/hel.12855

20. Lehours P, Mégraud F. 2021. Culture-based antimicrobial susceptibility testing for *Helicobacter pylori*, p 45–50. In Smith SM (ed), *Helicobacter pylori*. Springer US, New York, NY.

21. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

22. Corpet F. 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16:10881–10890. https://doi.org/10.1093/nar/16.22.10881

23. Alm RA, Ling L-SL, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, Caruso A, Uria-Nickelsen M, Mills DM, Ives C, Gibson R, Merberg D, Mills SD, Jiang Q, Taylor DE, Vovis GF, Trust TJ. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature New Biol 397:176–180. https://doi.org/10.1038/16495

24. Thiberge J-M, Boursaux-Eude C, Lehours P, Dillies M-A, Creno S, Coppée J-Y, Rouy Z, Lajus A, Ma L, Burucoa C, Ruskoné-Foumestraux A, Courillon-Mallet A, De Reuse H, Boneca IG, Lamarque D, Mégraud F, Delchier J-C, Médigue C, Bouchier C, Labigne A, Raymond J. 2010. From array-based hybridization of *Helicobacter pylori* isolates to the complete genome sequence of an isolate associated with MALT lymphoma. BMC Genomics 11:368. https://doi.org/10.1186/1471-2164-11-368

25. Mannion A, Dzink-Fox J, Shen Z, Piazuelo MB, Wilson KT, Correa P, Peek RM, Camargo MC, Fox JG. 2021. *Helicobacter pylori* antimicrobial resistance and gene variants in high- and low-gastric-cancer-risk populations. J Clin Microbiol 59:e03203-20. https://doi.org/10.1128/JCM.03203-20

26. Azuma T, Yamakawa A, Yamazaki S, Ohtani M, Ito Y, Muramatsu A, Suto H, Yamazaki Y, Keida Y, Higashi H, Hatakeyama M. 2004. Distinct diversity of the cag pathogenicity island among Helicobacter pylori strains in Japan. J Clin Microbiol 42:2508–2517. https://doi.org/10.1128/JCM.42.6.2508-2517.2004

27. Kawai M, Furuta Y, Yahara K, Tsuru T, Oshima K, Handa N, Takahashi N, Yoshida M, Azuma T, Hattori M, Uchiyama I, Kobayashi I. 2011. Evolution in an oncogenic bacterial species with extreme genome plasticity: *Helicobacter pylori* East Asian genomes. BMC Microbiol 11:104. https://doi.org/10.1186/1471-2180-11-104

28. Kojima KK, Furuta Y, Yahara K, Fukuyo M, Shiwa Y, Nishiumi S, Yoshida M, Azuma T, Yoshikawa H, Kobayashi I. 2016. Population evolution of *Helicobacter pylori* through diversification in DNA methylation and interstrain sequence homogenization. Mol Biol Evol 33:2848–2859. https://doi.org/10.1093/molbev/msw162

29. Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Mégraud F, Otto K, Reichard U, Katzowitsch E, Wang X, Achtman M, Suerbaum S. 2003. Traces of human migrations in *Helicobacter pylori* populations. Science 299:1582–1585. https://doi.org/10.1126/science.1080857

30. Chen S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. Imeta 2:e107. https://doi.org/10.1002/imt2.107

31. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:257. https://doi.org/10.1186/s13059-019-1891-0

32. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923

33. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing (2). arXiv. https://doi.org/10.48550/ARXIV.1207.3907

34. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10:giab008. https://doi.org/10.1093/gigascience/giab008

35. Souvorov A, Agarwala R, Lipman DJ. 2018. SKESA: strategic k-mer extension for scrupulous assemblies. Genome Biol 19:153. https://doi.org/10.1186/s13059-018-1540-z

36. Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587. https://doi.org/10.1093/genetics/164.4.1567

37. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1

38. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455. https://doi.org/10.1093/bioinformatics/btz305

39. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645. https://doi.org/10.1101/gr.092759.109

40. O'Toole PW, Logan SM, Kostrzynska M, Wadström T, Trust TJ. 1991. Isolation and biochemical and molecular analyses of a species-specific protein antigen from the gastric pathogen *Helicobacter pylori*. J Bacteriol 173:505–513. https://doi.org/10.1128/jb.173.2.505-513.1991

41. Vandamme P, Falsen E, Rossau R, Hoste B, Segers P, Tytgat R, De Ley J. 1991. Revision of *Campylobacter, Helicobacter,* and Wolinella taxonomy: emendation of generic descriptions and proposal of *Arcobacter* gen. nov. Int J Syst Bacteriol 41:88–103. https://doi.org/10.1099/00207713-41-1-88

42. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153

43. Hasan MR, Rawat A, Tang P, Jithesh PV, Thomas E, Tan R, Tilley P. 2016. Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. J Clin Microbiol 54:919–927. https://doi.org/10.1128/JCM.03050-15

44. Nelson MT, Pope CE, Marsh RL, Wolter DJ, Weiss EJ, Hager KR, Vo AT, Brittnacher MJ, Radey MC, Hayden HS, Eng A, Miller SI, Borenstein E, Hoffman LR. 2019. Human and extracellular DNA depletion for metagenomic analysis of complex clinical infection samples yields optimized viable microbiome profiles. Cell Rep 26:2227–2240. https://doi.org/10.1016/j.celrep.2019.01.091

45. Shi Y, Wang G, Lau H-H, Yu J. 2022. Metagenomic sequencing for microbial DNA in human samples: emerging technological advances. IJMS 23:2181. https://doi.org/10.3390/ijms23042181

46. Bowden KE, Joseph SJ, Cartee JC, Ziklo N, Danavall D, Raphael BH, Read TD, Dean D. 2021. Whole-genome enrichment and sequencing of *Chlamydia trachomatis* directly from patient clinical vaginal and rectal swabs. mSphere 6:e01302-20. https://doi.org/10.1128/mSphere.01302-20

47. Fernández-Caso B, Miqueleiz A, Alarcón T. 2023. Whole genome sequencing for studying *Helicobacter pylori* antimicrobial resistance. Antibiotics (Basel) 12:1135. https://doi.org/10.3390/antibiotics12071135

48. Hulten KG, Genta RM, Kalfus IN, Zhou Y, Zhang H, Graham DY. 2021. Comparison of culture with antibiogram to next-generation sequencing using bacterial isolates and formalin-fixed, paraffin-embedded gastric biopsies. Gastroenterology 161:1433–1442. https://doi.org/10.1053/j.gastro.2021.07.012

49. Martínez-Martínez FJ, Chiner-Oms Á, Furio V, Yamaoka Y, Dekker JP, Megraud F, Bénéjat L, Ducournau A, Giese A, Jehanne Q, Jauvain M, Camargo MC, Comas I, Lehours P. 2023. Genomic determinants of antibiotic resistance in *Helicobacter pylori* treatment and surveillance. Microb Health Dis. https://doi.org/10.2139/ssrn.5064086

50. Kwon DH, El-Zaatari FA, Kato M, Osato MS, Reddy R, Yamaoka Y, Graham DY. 2000. Analysis of rdxA and involvement of additional genes encoding NAD(P)H flavin oxidoreductase (FrxA) and ferredoxin-like protein (FdxB) in metronidazole resistance of *Helicobacter pylori*. Antimicrob Agents Chemother 44:2133–2142. https://doi.org/10.1128/AAC.44.8.2133-2142.2000

51. Han F, Liu S, Ho B, Yan Z, Yan X. 2007. Alterations in rdxA and frxA genes and their upstream regions in metronidazole-resistant *Helicobacter pylori* isolates. Res Microbiol 158:38–44. https://doi.org/10.1016/j.resmic.2006.10.001

52. Rimbara E, Noguchi N, Kawai T, Sasatsu M. 2007. Correlation between substitutions in penicillin-binding protein 1 and amoxicillin resistance in *Helicobacter pylori*. Microbiol Immunol 51:939–944. https://doi.org/10.1111/j.1348-0421.2007.tb03990.x

53. Graham DY, Liou J-M. 2022. Primer for development of guidelines for *Helicobacter pylori* therapy using antimicrobial stewardship. Clin

Gastroenterol Hepatol 20:973–983. https://doi.org/10.1016/j.cgh.2021.03.026

54. Argent RH, Zhang YL, Atherton JC. 2005. Simple method for determination of the number of *Helicobacter pylori* CagA variable-region EPIYA tyrosine phosphorylation motifs by PCR. J Clin Microbiol 43:791–795. https://doi.org/10.1128/JCM.43.2.791-795.2005

55. Harrer A, Boehm M, Backert S, Tegtmeyer N. 2017. Overexpression of serine protease HtrA enhances disruption of adherens junctions, paracellular transmigration and type IV secretion of CagA by *Helicobacter pylori*. Gut Pathog 9:40. https://doi.org/10.1186/s13099-017-0189-6

56. Zarzecka U, Modrak-Wójcik A, Figaj D, Apanowicz M, Lesner A, Bzowska A, Lipinska B, Zawilak-Pawlik A, Backert S, Skorko-Glonek J. 2019. Properties of the HtrA protease from bacterium *Helicobacter pylori* whose activity is indispensable for growth under stress conditions. Front Microbiol 10:961. https://doi.org/10.3389/fmicb.2019.00961

57. Zarzecka U, Tegtmeyer N, Sticht H, Backert S. 2023. Trimer stability of *Helicobacter pylori* HtrA is regulated by a natural mutation in the protease domain. Med Microbiol Immunol 212:241–252. https://doi.org/10.1007/s00430-023-00766-9

58. 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. https://doi.org/10.1136/gut.2010.234468