Helicobacter pylori infections in the Bronx, New York: Surveying Antibiotic Susceptibility and Strain Lineage by Whole-genome Sequencing


1 Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY
2 Department of Pathology, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY
3 Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, VA, USA
4 Provincial Public Health Laboratory, Eastern Health Microbiology Services, St. John’s, New Foundland and Labrador, Canada
5 Department of Medicine, Division of Gastroenterology, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY
6 Department of Pediatrics, Division of Gastroenterology and Nutrition, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY

*Co-last authors

Corresponding Author
Wendy A. Szymczak, PhD, D(ABMM), MT(ASCP)
Montefiore Medical Center
Bronx, NY 10467
wszymcza@montefiore.org
Abstract

Emergence of drug resistance in *Helicobacter pylori* has resulted in a greater need for susceptibility-guided treatment. While alleles associated with resistance to clarithromycin and levofloxacin have been defined, there are limited data regarding molecular mechanisms underlying resistance to other antimicrobials. Using *H. pylori* isolates from 42 clinical specimens, we compared phenotypic and whole genome sequencing (WGS)-based detection of resistance. Phenotypic resistance correlated with the presence of alleles of 23S rRNA (A2142G/A2143G) for clarithromycin (kappa 0.84, 95% CI: 0.67-1.0) and gyrA (N87I/N87K/D91Y/D91N/D91G/D99N) for levofloxacin (kappa 0.90, 95% CI: 0.77-1.0).

Phenotypic resistance to amoxicillin in three isolates correlated with mutations in *pbp1*, *pbp2*, and/or *pbp3* within coding regions near known amoxicillin binding motifs. All isolates were phenotypically susceptible to tetracycline, although four bore a mutation in 16S rRNA (A926G).

For metronidazole, nonsense mutations and R16H substitutions in *rdxA* correlated with phenotypic resistance (kappa = 0.76, 95% CI: 0.56-0.96). Previously identified mutations in the *rpoB* rifampicin resistance-determining region (RRDR) were not present, but 14 novel mutations outside the RRDR were found in rifampicin resistant isolates. WGS also allowed for strain lineage determination, which may be important for future studies in associating precise MICs with specific resistance alleles. In summary, WGS allows for broad analyses of *H. pylori* isolates and our findings support the use of WGS for detection of clarithromycin and levofloxacin resistance. Additional studies are warranted to better define mutations conferring resistance to amoxicillin, tetracycline, and rifampin, but combinatorial analyses for *rdxA* gene truncations and R16H mutations have utility for determining metronidazole resistance.
Introduction

*Helicobacter pylori* colonizes approximately half of the population globally and causes gastritis in nearly all infected individuals (1). In some, infection will progress resulting in dyspepsia, formation of gastric and duodenal ulcers, mucosa-associated lymphoid tissue lymphoma, and gastric cancer. In one study conducted in Japan, 2.9% of patients with gastric ulcers, dyspepsia, or gastric hyperplasia in conjunction with *H. pylori* infection developed gastric cancer (2). *H. pylori* is the only bacterium classified as a Type 1 carcinogen, with pathogen eradication resulting in reduction of cancer risk (3-5). Anti-*H. pylori* therapy involves multidrug regimens that consist of at least two antibiotics and proton pump inhibitors (PPI), with or without bismuth (3, 5).

Treatment of primary *H. pylori* infections is largely empiric, but susceptibility-guided therapy is recommended after initial treatment failures or if the local level of resistance to clarithromycin or metronidazole is greater than 15% (5). Two meta-analyses that included seventeen randomized controlled trials revealed that susceptibility-guided treatment is superior to empirical therapy with two antibiotics and a PPI (6, 7). However, phenotypic susceptibility testing for *H. pylori* is not offered by the majority of North American hospital laboratories due to challenges in performing the assay. *H. pylori* culture and susceptibility testing can take weeks because of the slow growth and fastidious nature of the organism (3).

Molecular detection of *H. pylori* drug resistance may provide a more rapid and amenable means for susceptibility determination. Real-time PCR and DNA hybridization-based assays have been shown to correlate with phenotypic results for the detection of point mutations mediating clarithromycin (93-97% concordance) (8-13) and levofloxacin (80-92%)
(10, 11) resistance, and targeted sequencing of the *H. pylori* 23S rRNA gene also shows high concordance with phenotypic results for clarithromycin (14). Molecular methods can be superior to phenotypic methods for detection of subpopulations of resistant strains, and sequencing offers an advantage over PCR-based assays since all known mutations within a targeted gene or whole genome can be assessed (15). Recently, the agreement between whole genome sequencing (WGS) and E-test results for clarithromycin, levofloxacin, and rifampicin was found to be strong; however, the study found no agreement for metronidazole (16).

In this study, we performed WGS of *H. pylori* isolates originating from 42 clinical specimens to determine strain lineage in our local population of isolates and to compare WGS-mediated resistance detection to phenotypic E-test results for all antimicrobials with \( \text{Clinical Laboratory Standards Institute (CLSI) or European Committee for Antimicrobial Susceptibility Testing (EUCAST)} \) breakpoints. Additionally, WGS led to identification of novel mutations that have not previously been associated with drug resistance.
Methods

H. pylori isolates

H. pylori isolates were recovered from gastric biopsy specimens submitted to the clinical microbiology laboratory at Montefiore Medical Center, Bronx, NY, for routine patient care between May 14th, 2018 and May 8th, 2019. Routine testing consisted of culture and phenotypic susceptibility testing for clarithromycin, amoxicillin, levofloxacin, and tetracycline. Isolates, which consisted of multiple colonies, were then stored in Cryosaver Brucella glycerol and bead-containing tubes (Hardy Diagnostics) at -80°C. Isolates that remained viable after freezing were whole genome sequenced and phenotypic susceptibility testing for rifampin and metronidazole was performed for research purposes only. Routine susceptibility data and corresponding patient demographics were retrieved retrospectively from the electronic medical records. Primary infection was defined as no previous documentation in the notes or previous test results suggestive of H. pylori infection. Testing reviewed included results of H. pylori stool antigen, rapid urease point-of-care, cultures, prescription history, and pathology reports. Only the first isolate for each patient recovered during the study period was included. This study was approved by the Albert Einstein College of Medicine IRB (2018-9702). The need for informed consent was waived by the IRB since loss to follow-up of patients was expected given the use of remnant isolates.

Phenotypic antimicrobial susceptibility testing

Phenotypic susceptibility testing was performed using 3-4 day old cultures of H. pylori when sufficient growth was obtained to prepare a 2 McFarland suspension (17). H. pylori clinical isolates and quality control strain ATCC 43504 were inoculated into Mueller Hinton Blood Agar...
Plates containing 5% sheep blood (BD BBL, Sparks, Maryland) and Epsilon tests (E-tests, Biomerieux, France) were used for determination of the minimum inhibitory concentration (MIC) after 72-96 hours of incubation at 37°C under microaerophilic conditions. CLSI M45 3rd edition breakpoints were used to determine susceptibility to clarithromycin. Interpretations for other antibiotics were based on EUCAST 2019 version 9.0 guidelines, solely for the purposes of this study.

Isolation of genomic DNA and whole genome sequencing.

*H. pylori* clinical isolates were sub-cultured on blood agar plates for genomic DNA isolation. Three plates were used for each isolate after 3-4 days of incubation under microaerophilic conditions at 37°C. The isolates were resuspended in sterile saline solution and pelleted. *H. pylori* clinical isolates were resuspended in sterile saline solution and pelleted. Genomic DNA was prepared using PureLink genomic DNA isolation kit (Invitrogen Inc.) following manufacturer’s instructions. DNA quality was assessed using Nanodrop One (Thermo Scientific, Inc). All DNA quantifications were done by Qubit™ dsDNA high sensitivity assay kit (Invitrogen Inc). Paired end genomic library was prepared using Nexera XT library preparation kit (Illumina, San Diego, CA) as per manufacturer’s instructions. Following clean up using Ampure XP magnetic beads (Beckman Coulter, USA), genomic libraries were pooled in equimolar concentration (2 nM each) with 20 pM PhiX control and sequencing was performed using MiSeq™ Reagent Kit v3 (150-cycle) (2 x 75 bp) on an Illumina Miseq platform (Illumina, Inc). Quality of read files were evaluated using FastQC v0.11.7 (18). The read files and adaptors were trimmed using Trim Galore version (v0.6.1) and Cutadapt version (v2.2). Sequence coverage for
all the isolates were assessed using Miseq reporter by keeping *H. pylori* 26695 as reference.

Coverage ranged between 28x-210x with an average depth of 80x.

**Genome assembly and annotation**

Short read assembly was performed using SPAdes (19) at the Pathosystems Resource Integration Center (PATRIC) online server (https://www.patricbrc.org/) (20). The number of contigs obtained after assembly ranged from 33 to 72 with the exception of one isolate (MHP39) with 188 contigs. The assembled contigs were annotated by RASTtk (21) using PATRIC annotation server with default parameters. Genome annotations revealed the presence of 1573 to 1702 ORFs encoding proteins and RNAs among the sequenced clinical isolates. Pan genome of sequenced isolates constituted 2031 genes while the core genome consisted of 1177 genes.

Quality metrics data for all sequenced genomes including coverage, genome length, N50 and L50 values are detailed in Supplementary data set 1. *H. pylori* isolates from 42 clinical specimens were used in this study; however, it was found upon subculture of the frozen clinical isolates that one specimen (MH9) consisted of two *H. pylori* strains that were sequenced independently (MH9C and MH9L).

**WGS and antimicrobial resistance analyses**

In order to identify drug resistance mutations, sequence data corresponding to the 23S rRNA gene for clarithromycin, *gyrA* and *gyrB* for levofloxacin, *pbp1*, *pbp2* and *pbp3* for amoxicillin, *rdxA*, *frxA*, *fdxB*, *mdaB*, *omp11* and *rpsU* for metronidazole, 16S rRNA gene for tetracycline, and *rpoB* for rifampin were retrieved from the annotated genomes and analyzed by performing multiple sequence alignment with MUSCLE using MEGA 10.0.5 and Geneious R11 by keeping a sensitive strain (*H. pylori* 26695) as reference (22). Defined drug resistance mutations were
used to determine the correlation between phenotypic and genotypic results (23). The
presence of novel variants associated with phenotypic resistance was evaluated separately.
Isolates with discrepant phenotype and sequence results were examined for allele variations by
read mapping trimmed short reads to *H. pylori* 26695 using “Map to reference” option with a
threshold of 95 in Geneious R11. We observed at least 200X to a maximum of 1200X depth for
all drug resistant mutations.

**Epidemiology and lineage analyses**

Multi-locus sequence typing (MLST) on all isolates was performed using a web based MLST tool
hosted by Center of Genetic Epidemiology, Denmark ([http://www.genomicepidemiology.org](http://www.genomicepidemiology.org)). A
phylogenetic tree was constructed using PATRIC’s ‘Codon Tree’ service, which uses the amino
acid and nucleotide sequences from defined number of PATRIC’s global Protein Families (24).
The protein families, known as PGFams, were picked randomly to build an alignment and then
to generate a tree based on the differences within those selected sequences. Both the protein
(amino acid) and gene (nucleotide) sequences were used for each of the selected genes from
the PGFams. Protein sequences were aligned using MUSCLE (22), and the nucleotide coding
gene sequences were aligned using the Codon align function of BioPython (25). A concatenated
alignment of all proteins and nucleotides were written to a phylip formatted file, and then a
partitions file for RaxML (26) was generated, describing the alignment in terms of the proteins
and then the first, second and third codon positions. This phylip file was used as an input file for
FigTree v1.4.4 to build an unrooted phylogenetic tree. Support values were generated using
100 rounds of the “Rapid” bootstrapping option (27) of RaxML. Sequences from 512 genes
(128869 amino acids and 386361 nucleotides) were used to generate the tree, which included
102 genomes (Supplementary data set 2 & Supplementary data set 3). These included the 43 *H. pylori* genomes sequenced in this study, and sixty-six additional genomes that were described in Kumar et al. (28).

**Statistical analyses**

Cohen’s kappa coefficient was used to assess the agreement between phenotypic resistance and drug resistance mutations. A value of <0.4 was considered low agreement; 0.4 – 0.6, moderate agreement; 0.61 – 0.8 substantial agreement; and 0.81 – 1.0 near or perfect agreement. SAS version 9.4 (SAS Institute Inc., Cary, NC) was used to estimate kappa coefficient values. Prism version 7 was used to perform Fisher’s exact test for contingency analyses followed by Bonferroni-Dunn to correct for multiple comparisons. A p-value of ≤ 0.05 was considered statistically significant.

**Data Availability**

All 43 sequenced genome of *H. pylori* strains isolated from 42 clinical specimens in this study were assembled, annotated and submitted to NCBI Bioproject database under Bioproject ID PRJNA566177. The accession numbers are as follows WACT00000000, WACU00000000,

WACV00000000, WACW00000000, WACX00000000, WACY00000000, WACZ00000000,

WADA00000000, WADB00000000, WADC00000000, WADD00000000, WADE00000000,

WADF00000000, WADG00000000, WADH00000000, WADI00000000, WADJ00000000,

WADK00000000, WADL00000000, WADM00000000, WADN00000000, WADO00000000,

WADP00000000, WADQ00000000, WADR00000000, WADS00000000, WADT00000000,

WADU00000000, WADV00000000, WADW00000000, WADX00000000, WADY00000000,
189 WADZ00000000, WAEA00000000, WAEB00000000, WAEC00000000, WAED00000000,
190 WAEE00000000, WAEF00000000, WAEG00000000, WAEH00000000, WAEI00000000,
191 WAEJ00000000.
192
193
Results

Patient characteristics

H. pylori clinical isolates \((n=54)\) recovered from patient biopsy specimens between May 14\(^{th}\), 2018 and May 8\(^{th}\), 2019 were stored for this study. Isolates from 42 specimens were recoverable after storage \((n=42, 79\%)\), and the corresponding patient demographics and clinical characteristics were retrieved for analyses. For this study population in the Bronx (New York), the majority of isolates were from adults, but 21\% percent of isolates were from pediatric patients under the age of 21 (Table 1). Most of the isolates were from females \((76\%)\) and the most common race recorded in the medical record was “other” \((38\%)\), which likely reflects those that identify as of Hispanic or Latino heritage in the Bronx. Review of the medical records revealed that nearly half \((48\%)\) of isolates were from patients with a documented previous H. pylori infection. For patients with corresponding histopathology evaluation, most exhibited active, chronic gastritis in the antrum and/or fundus \((74\%)\) without evidence of progression to intestinal metaplasia. The most common antibiotic combinations for empiric therapy or for treatment of the most recently documented H. pylori infection prior to specimen collection \((n=23)\) were amoxicillin in combination with clarithromycin \((n=10)\) followed by amoxicillin with levofloxacin \((n=5, \text{Table 1}).\)

Phenotypic antimicrobial resistance of H. pylori isolates

The most common phenotypic resistance observed was to levofloxacin \((57\% \text{ of isolates}),\) metronidazole \((45\%)\) and clarithromycin \((31\%).\) In contrast, phenotypic resistance to amoxicillin was rare \((7\%)\) and none of the isolates were found to be phenotypically resistant to...
tetracycline. Using EUCAST guidelines for susceptibility to rifampicin (MIC ≤1 μg/ml as susceptible), 19% of our isolates were resistant. A recent study has suggested that an MIC of ≤4 μg/ml for susceptibility is more appropriate (29), and using this higher breakpoint, 7% of the isolates were rifampicin resistant. Co-resistance to both levofloxacin and clarithromycin was observed in 21% of the isolates. Resistance to clarithromycin was more common in isolates obtained from patients with previous documented *H. pylori* infection (45% secondary/tertiary vs. 18% primary infections) but the association was not statistically significant (*p*=0.1). Resistance to levofloxacin was similar for isolates from previous and primary infections (60% vs. 50%, *p*=0.6).

**Genotypic determination of antimicrobial resistance**

The sequences of drug resistance related genes were retrieved from the annotated genomes to identify potential drug resistance mutations. Mutations in genes previously reported to be associated with amoxicillin, tetracycline, clarithromycin, levofloxacin, metronidazole and rifampicin resistance were examined and the correlation with phenotypic results determined (23, 30, 31). Three isolates were phenotypically resistant to amoxicillin (MHP11, MHP12, MHP22) and all exhibited point mutations in penicillin binding protein genes *pbp1* (N107R, A201V, V250I, S543T), *pbp2* (I259T) and/or *pbp3* (D2N, A50S, F490Y, A541T and V374I) that were not present in susceptible isolates (Table 2). Although none of the amino acid substitutions were located in regions previously reported to confer amoxicillin resistance such as SXXK/XTG in Pbp1A (32), all three isolates exhibited at least one substitution nearby a SXXK/XTG motif in Pbp1A, Pbp2, or Pbp3 (Supplementary table 1). All isolates were
tetracycline susceptible phenotypically, but four isolates (MHP28, MHP32, MHP38, MHP44) had point mutations in 16S rRNA that are associated with resistance (A926G) (33). The 23S rRNA mutations A2142G and A2143G correlated with phenotypic E-test results for clarithromycin (kappa 0.84, 95% CI: 0.67-1.0, Table 3). All thirteen clarithromycin phenotypic resistant isolates demonstrated point mutations in 23S rRNA (A2142G/A2143G). However, we also identified three isolates that were resistant by WGS but phenotypically sensitive (MHP28, MHP34, MHP50) (Table 2). None of our isolates harbored A2142C, A2115G, G2141A, C2147G, T2190C mutations that have been previously associated with clarithromycin resistance (12, 34). We did observe C2195T mutations in three isolates (MHP41, 50 and 52) among which MHP41 and MHP50 were sensitive phenotypically and MHP52 had A2142G mutations. We did not consider this mutation in our analyses as previous studies have only found associations with resistance when present in combination with other known clarithromycin mutations (35, 36). Twenty-three isolates were phenotypically resistant to levofloxacin, and resistance mutations in the Quinolone Resistance-Determining Region (QRDR) in gyrA (N87I/N87K/D91Y/D91N/D91G/D99N) were found in 22 isolates. Isolate (MHP11) carried R484K mutation in both gyrB in addition to gyrA mutation. For one phenotypically resistant isolate (MHP35), mutations in gyrA or gyrB were not identified (Table 2). Agreement between methods for levofloxacin was strong (kappa 0.9, 95% CI: 0.77-1.0, Table 3). In addition we identified heteroresistant strains in one of our isolates (MHP9) wherein one strain was resistant to clarithromycin (MHP9C) and the other was resistant to levofloxacin (MHP9L, Table 2).

None of the previously reported drug resistant mutations within the rifampicin resistance-determining region (RRDR) of rpoB were present in our isolates. However, five of 13
phenotypically resistant isolates with MIC > 1 μg/ml and five of seven with MICs ≥ 3 μg/ml exhibited mutations outside the rpoB RRDR. No mutations in rpoB were present in isolates with MICs < 3 μg/ml (Table 2). Correlation between methods was poor (kappa 0.46, 95% CI: 0.18-0.75) when the EUCAST breakpoint of MIC > 1 μg/ml was used to define phenotypic resistance, but improved when an MIC >4 μg/ml was used as a breakpoint (kappa 0.73, 95% CI: 0.37-1.0, Table 3) (29).

For metronidazole resistance analyses, all sequences were assessed for the presence of intact rdxA and fxaA, either of which convert the pro-drug form of metronidazole into its active form. rdxA is believed to be the predominant gene responsible for mediating metronidazole resistance, with the role of fxa being discrepant among studies (37). Eleven of 19 phenotypically resistant isolates carried nonsense mutations in rdxA at amino acid positions 50, 72, 75, 76, 112, 119, 143 and 197, which leads to the synthesis of truncated/non-functional protein (Supplementary Table 2) (38). The presence of rdxA truncations alone was present in only 58% of the resistant isolates.

Various point mutations in rdxA have been associated with metronidazole resistance, but results are not consistent between studies (23). We examined our isolates for point mutations that may be associated with metronidazole resistant and found a mutation at position R16 that was associated with six of the eight phenotypic resistance isolates that lacked a truncation in rdxA (Table 2, Supplementary Table 2). This mutation has been previously associated with metronidazole phenotypic resistance (37, 39, 40). The detection of either a
truncated *rdxA* gene or a mutation at position R16 correlated (kappa 0.76) with 88% of phenotypic testing results (Table 2).

**H. pylori** lineage tree analyses

WGS data was used to construct a phylogenetic tree to determine where the 43 clinical strain genomes from the 42 clinical specimens clustered when compared to previously defined *H. pylori* sequence types (28). The hpEastAsia, hpSouthIndia, hpWestAfrica and hpAfrica2 sequence types are monophyletic with a common ancestor and each having support values of 100 at the defining branch point. The hpAmerind and hpEurope genomes were found to be paraphyletic wherein the isolates were descended from multiple ancestors. Three of the 43 genomes sequenced here (MHP4, MHP3 and MHP10) cluster with the hspEastAsia clade, 12 with the hpEurope clade, and 28 with the hpWestAfrica clade. Interestingly, the heteroresistant isolate MHP9 was comprised of two strains from different lineages; strain MHP9C from hpWestAfrica and MHP9L from hpEurope (Figure 1). None of the genomes we sequenced were closely related to the hpAmerind, hpSouthIndia or hpAfrica1 clades.

Clarithromycin, levofloxacin and amoxicillin phenotypically resistant strains were found to be spread across all identified lineages. All clarithromycin resistant strains from West African (*n*=6) and East Asian lineage (*n*=2) were co-resistant to levofloxacin.
Drug resistance among clinical isolates of *H. pylori* is a well-recognized issue and susceptibility-guided treatment of *H. pylori* is crucial to attain better treatment response among infected individuals (41). WHO has listed *H. pylori* as a high priority pathogen for which there is an urgent need to develop new antimicrobials (42). We assessed the utility of WGS as a method for detecting molecular determinants of resistance by comparing to phenotypic testing results. We found strong correlation for clarithromycin (93%) and levofloxacin (95%) between methods. Additionally, we identified mutations in our clinical isolates that correlated with metronidazole phenotypic resistance and identified mutations outside of the *rpoB* RRDR and novel mutations in penicillin binding proteins that may contribute to rifampicin and amoxicillin resistance, respectively.

WGS allowed us to explore putative mechanisms underlying phenotypic resistance in our *H. pylori* isolates. For amoxicillin, point mutations in penicillin binding protein genes have been predicted to result in amoxicillin resistance in *H. pylori*, but most studies report novel mutations that require validation (23, 43). We identified 11 amino acid substitutions in *Pbp1A/Pbp2/Pbp3* exclusively in three amoxicillin resistant isolates and seven of these substitutions were found adjacent to SXN/SXXK/KTG motifs. In *Pbp1A*, mutations in these motifs have been found to confer resistance (32), and mutations adjacent to SKN<sub>402</sub>−<sub>404</sub> and KTG<sub>555</sub>−<sub>557</sub> have been previously identified among resistant isolates (44). Our data provides further evidence that the identification of mutations near-by SXN/SXXK/KTG motifs may contribute to resistance. Another observation was the detection of a D99N substitution (*gyrA*)
in a levofloxacin phenotypically resistant isolate (MHP11). A previous study from Nepal found that mutation at this position confers resistance, but a D99V substitution was noted (40).

WGS revealed clinically significant mutations in isolates testing phenotypically sensitive, which highlights an advantage of WGS over phenotypic testing. We noted point mutations in 23S rRNA (A2142G/A2143G) in three strains that were phenotypically sensitive to clarithromycin (MHP28, MHP34, MHP50) and a mutation in gyrA at position N87K in an isolate sensitive to levofloxacin (MHP35). The A2143G mutation is associated with clarithromycin MIC values ranging from susceptible to resistant, and patients with this mutation have a 60% increased risk for treatment failure (45). Environmental conditions and the genetic context (ie. strain lineage) may also influence the phenotypic susceptibility result (46). Another possible explanation for discrepancies between methods is that we may have selectively sub-cultured colonies with drug resistance while reviving from stocks containing mixed populations. Nevertheless, the existence of heteroresistance is common during H. pylori infections (15).

In addition, we detected mutations in 16S rRNA gene (A926G) in four tetracycline sensitive isolates, but the significance of this finding is unclear since this mutation has been associated with both low level resistance and phenotypic susceptibility (16, 47, 48). In a recent study that examined the correlation between detection of mutations directly from formalin fixed, paraffin-embedded tissue, only one treatment failure was noted in six patients with the A926G mutation, but studies with larger numbers of patients treated with tetracycline-containing regimens are needed to understand the clinical significance (49).

Based on EUCAST guidelines, we categorized 13 isolates as rifampicin resistant (MIC >1), but did not identify any significant amino acid substitutions in the rpoB RRDR region, even for
isolates with an MIC >4 μg/ml. However, we could detect amino acid substitutions outside the rpoB RRDR exclusively in phenotypically resistant isolates with MICs ≥3 μg/ml. Mutations outside of the RRDR have been previously associated with elevated rifampicin MICs between 2-4 μg/ml (29). As previously suggested, the lack of detection of any mutations in our isolates with an MIC < 3 μg/ml provides further support that the EUCAST breakpoint may need to be re-evaluated. Further studies are needed to demonstrate that the point mutations identified outside the RRDR in our clinical isolates are responsible for the resistance phenotype observed. Isolates from all 42 specimens were tested phenotypically for metronidazole resistance and 19 had metronidazole MICs >256 μg/ml. The presence of one or both intact genes (rdxA and frxA) did not always indicate low MICs, signifying that there may be other point mutations that could contribute to resistance (50). Few mutations were identified among our isolates carrying intact rdxA and frxA. R16H and V57A substitutions in rdxA, R58C, A85V, I117M and E169K in frxA seem to be potential drug resistant mutations which were present only in drug resistant isolates with intact rdxA/frxA (Supplementary Table 2). However, R16H mutation in rdxA was the most common mutation found among our metronidazole resistant isolates with intact rdxA and has also been previously reported to be a significant resistant mutation (37, 39, 40, 48). We found that the combination of R16H and nonsense mutations in rdxA to be the best correlate of phenotypic resistance in our isolates. However, three isolates (MHP12, MHP15 & MHP31) had nonsense mutations in rdxA but were found to be sensitive by phenotypic E-test supporting the fact that metronidazole resistance is multifaceted and is not yet clearly understood (50). We did not find any exclusive mutations in genes rpsU, mdaB, omp11 or fdxB that have previously associated with metronidazole resistance among the metronidazole
resistant isolates (data not shown). Additional studies on other possible mechanisms such as overexpression of efflux pumps or defects in membrane porins might provide some novel insights into metronidazole resistance (50). A recent study comparing WGS to phenotypic testing using *H. pylori* isolates from a collection in Switzerland demonstrated an agreement between methods of 99% for clarithromycin and levofloxacin, 100% agreement for rifampicin, but no agreement for tetracycline (A926 mutation) or metronidazole (16). Taken together, our data and this previous study support the use of WGS for susceptibility determination for clarithromycin and levofloxacin but do not support a role for the A926G mutation in conferring tetracycline resistance. In contrast to this previously published study, we found that mutations outside the *rpoB* RRDR were associated with elevated MICs to rifampicin suggesting that a complete catalogue of resistance determinants for *H. pylori* must be established. Our study also adds novel mutations associated with amoxicillin resistance and provides molecular correlates of metronidazole resistance that best predict phenotypic testing results. Continued WGS-based characterization of *H. pylori* isolates will support future studies to determine which genetic determinants contribute to eradication failure. MLST analyses revealed the presence of unique alleles among all the sequenced isolates which is not uncommon in *H. pylori* and none of the isolates belonged to known sequence types (STs) (51). Phylogenetic analysis with reference genomes from known lineages revealed the existence of East Asian, European and West African (Africa 1) lineages in the Bronx. This observation is in agreement with the earlier studies from North America wherein West African
lineage is predominant and may be due to migration of people from different ethnic backgrounds (52). *H. pylori* infection usually occurs during childhood resulting from transmission through the oral-oral or oral-fecal route with several studies demonstrating transmission between members of the same family or between individuals sharing the same house (53-57). The presence of two different lineages (MHP9C & MHP9L) with different resistance patterns in one patient sample is not surprising since co-infection with multiple strains in one host can occur and in-host evolution can occur due to recombination events between *H. pylori* strains (54, 58). Resistant strains among all three lineages were identified, however clarithromycin/levofloxacin co-resistant strains were largely observed from East Asian and West African lineage. Analyses of a larger cohort of isolates is needed to confirm if any association between strain lineage and resistance exists.

In summary, WGS provides a more detailed analysis of *H. pylori* resistance for a broader array of antimicrobials that may be of clinical value. Strengths of our study are the inclusion of *H. pylori* isolates from pediatric patients and adults, patients with recurrent *H. pylori* infections where phenotypic resistance was greatest, and patients with diverse racial backgrounds from our community in the Bronx. Limitations of our study are a low number of isolates that were phenotypically resistant to amoxicillin or tetracycline and because of the limited availability of patient follow-up data, we could not determine if the detection of resistance by WGS was associated with patient outcomes. Because of the use of remnant specimens and retrospective chart review, we cannot confirm that all strains originated from unrelated individuals. Strains transmitted within families would be expected to be similar, which would reduce the significance of our findings. However, when MLST was performed, we found unique alleles
among all strains suggesting that transmission among individuals in our cohort would be unlikely. Our findings support the utilization of WGS for the determination of *H. pylori* resistance to clarithromycin and levofloxacin. Our data suggests that the single nucleotide mutation A926G is not associated with tetracycline phenotypic resistance, and provides potential genetic determinants of resistance for amoxicillin, rifampicin and metronidazole resistance for further investigation.

**Acknowledgements**

We would like to thank Yungtai Lo, PhD, Associate Professor of Epidemiology and Population Health, Albert Einstein College of Medicine, for his assistance with statistical analyses.
References


17. CLSI. Methods for antimicrobial dilution and disk susceptibility testing of infrequently isolated or fastidious bacteria. 3rd ed. CLSI guideline M45. Wayne, PA: Clinical and Laboratory Standards Institute; 2016.


obacter pylori strains isolated from...


Figure legends

Figure 1. Phylogenetic tree based on the amino acid and nucleotide sequences of 512 genes shared across 102 *Helicobacter pylori* genomes. The names of some of the genomes, originally identified in Kumar et al. (2015) are colored based on the phylogenetic lineages that they described. The 43 genomes in this analysis are colored black.
Table 1. Demographics and clinical characteristics of patients (n=42) with H. pylori-sequenced isolates.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>Mean - year</td>
<td>56</td>
</tr>
<tr>
<td>Distribution – no. (%)</td>
<td></td>
</tr>
<tr>
<td>&lt;21</td>
<td>9 (21)</td>
</tr>
<tr>
<td>&gt;21</td>
<td>33 (79)</td>
</tr>
<tr>
<td><strong>Gender – no. (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>32 (76)</td>
</tr>
<tr>
<td>Male</td>
<td>10 (24)</td>
</tr>
<tr>
<td><strong>Race – no. (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>13 (31)</td>
</tr>
<tr>
<td>Asian</td>
<td>3 (7)</td>
</tr>
<tr>
<td>White</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Other</td>
<td>16 (38)</td>
</tr>
<tr>
<td>Unavailable or declined</td>
<td>8 (19)</td>
</tr>
<tr>
<td><strong>Infection Status – no. (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>22 (52)</td>
</tr>
<tr>
<td>Secondary or tertiary</td>
<td>20 (48)</td>
</tr>
<tr>
<td><strong>Histopathology – no. (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Evaluation performed</td>
<td>31 (74)</td>
</tr>
<tr>
<td>Helicobacter-associated inactive, chronic gastritis</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Helicobacter-associated active, chronic gastritis</td>
<td>23 (74)</td>
</tr>
<tr>
<td>Helicobacter-associated active, chronic gastritis w/focal intestinal metaplasia</td>
<td>7 (23)</td>
</tr>
<tr>
<td><strong>Antibiotics used in H. pylori treatment regimens prior to biopsy – no. (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Amoxicillin, clarithromycin</td>
<td>10 (44)</td>
</tr>
<tr>
<td>Amoxicillin, levofloxacin</td>
<td>5 (22)</td>
</tr>
<tr>
<td>Amoxicillin, doxycycline or tetracycline</td>
<td>4 (17)</td>
</tr>
<tr>
<td>Clarithromycin, doxycycline</td>
<td>2 (9)</td>
</tr>
<tr>
<td>Doxycycline, levofloxacin, metronidazole</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Moxifloxacin, rifabutin</td>
<td>1 (4)</td>
</tr>
</tbody>
</table>
Table 2. *H. pylori* isolate E-test MIC values and resistance mutations/genes for amoxicillin, clarithromycin, levofloxacin and tetracycline, metronidazole and rifampicin

<table>
<thead>
<tr>
<th>Si No.</th>
<th>Isolate ID</th>
<th>MIC</th>
<th>Mutations</th>
<th>Primary [P]/ Recurrent [R]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AMX (MICs)</td>
<td>CLR (MICs)</td>
<td>LVX (MICs)</td>
<td>TE (MICs)</td>
</tr>
<tr>
<td>1</td>
<td>0.056</td>
<td>0.064</td>
<td>0.031</td>
<td>0.38</td>
</tr>
<tr>
<td>2</td>
<td>0.023</td>
<td>0.047</td>
<td>0.023</td>
<td>0.047</td>
</tr>
<tr>
<td>3</td>
<td>0.026</td>
<td>0.023</td>
<td>0.023</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>0.012</td>
<td>&gt;256</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
<tr>
<td>5</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.19</td>
</tr>
<tr>
<td>6</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.19</td>
</tr>
<tr>
<td>7</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.25</td>
</tr>
<tr>
<td>8</td>
<td>0.012</td>
<td>&gt;256</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
<tr>
<td>9</td>
<td>0.064</td>
<td>1</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
<tr>
<td>10</td>
<td>0.016</td>
<td>0.125</td>
<td>0.75</td>
<td>0.5</td>
</tr>
<tr>
<td>11</td>
<td>0.019</td>
<td>&gt;256</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
<tr>
<td>12</td>
<td>0.016</td>
<td>0.125</td>
<td>0.5</td>
<td>&gt;256</td>
</tr>
<tr>
<td>13</td>
<td>0.012</td>
<td>0.125</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0.016</td>
<td>0.031</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
<tr>
<td>15</td>
<td>0.023</td>
<td>0.125</td>
<td>0.125</td>
<td>&gt;256</td>
</tr>
<tr>
<td>16</td>
<td>0.012</td>
<td>0.125</td>
<td>0.125</td>
<td>&gt;256</td>
</tr>
<tr>
<td>17</td>
<td>0.016</td>
<td>0.125</td>
<td>0.125</td>
<td>&gt;256</td>
</tr>
<tr>
<td>18</td>
<td>0.016</td>
<td>0.125</td>
<td>0.125</td>
<td>&gt;256</td>
</tr>
<tr>
<td>19</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>20</td>
<td>0.016</td>
<td>&gt;256</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
<tr>
<td>21</td>
<td>0.023</td>
<td>0.047</td>
<td>0.047</td>
<td>0.047</td>
</tr>
<tr>
<td>22</td>
<td>0.016</td>
<td>&gt;256</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
<tr>
<td>23</td>
<td>0.016</td>
<td>&gt;256</td>
<td>&gt;256</td>
<td>&gt;256</td>
</tr>
</tbody>
</table>

**Mutations**

- **AMX**: D1372T, D1372N, A550G, F450Y, A543T
- **CLR**: D13, N87K
- **LVX**: A2143G
- **TE**: T485S
- **RIF**: A2143G
- **MET**: N87K

**Resistance**

- **p(pbp1A, 2 & 3)**: S543T
- **p(pbp2)**: D93E, S91F
- **p(pbp3)**: A541T
- **p(pbp1A)**: I259T
- **16S rRNA**: A926G
- **23S rRNA**: A2142G
- **rpoB**: D5149N, F971Y, K1137R

---

**References**

1. H. pylori isolate E-test MIC values and resistance mutations/genes for amoxicillin, clarithromycin, levofloxacin and tetracycline, metronidazole and rifampicin.
3. E-test results are shown in MICs (micrograms per milliliter) for each antibiotic.
4. Resistance mutations are reported for *p(pbp1A, 2 & 3)*, *p(pbp2)*, *p(pbp3)*, *p(pbp1A)*, and mutations in *16S rRNA* and *23S rRNA*.
5. Primary (P) and recurrent (R) resistance patterns are indicated.
6. **S543T** (pbp1A): S543T substitution in *pbp1A* gene.
7. **D1372T** (pbp1A): D1372T substitution in *pbp1A* gene.
8. **D1372N** (pbp1A): D1372N substitution in *pbp1A* gene.
10. **F450Y** (pbp1A): F450Y substitution in *pbp1A* gene.
11. **A543T** (pbp1A): A543T substitution in *pbp1A* gene.
15. **I259T** (pbp1A): I259T substitution in *pbp1A* gene.

---

**Notes**

- Full resistance patterns are indicated with **>256** for all antibiotics.
- Resistance patterns are interpreted based on breakpoints provided by the manufacturer.
- Resistance mutations are interpreted based on known resistance mechanisms.
- Primary (P) and recurrent (R) resistance patterns are indicated for each isolate.

---

**Acknowledgments**

- The authors acknowledge the support of the National Institutes of Health (NIH) for funding this study.
- Special thanks to the collaborating institutions for providing bacterial isolates.

---

**Author Contributions**

- The senior author designed the study and supervised data collection.
- The junior author performed the laboratory analyses and data interpretation.
- All authors contributed to the manuscript writing and critical revision.

---

**Conflicts of Interest**

- The authors declare no conflicts of interest.

---

**References**

Isolates with phenotypically resistant MIC values are highlighted in red. Discrepancies among WGS results when compared to phenotypic resistance were highlighted in blue. Blank cells indicate that no mutations associated with resistance in 23S rRNA gene for clarithromycin, gyrA and gyrB for levofloxacin, *pbp1*, *pbp2* and *pbp3* for amoxicillin, 16S rRNA for tetracycline, or *rdxA* for metronidazole.

MHP9 is a heteroresistant isolate comprised of strain MHP9C (resistant to clarithromycin) and MHP9L (resistant to levofloxacin).
Table 3. Agreement between phenotypic MIC values and resistance genotype

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>MIC µg/ml Breakpoint</th>
<th>Genotype</th>
<th>Phenotype</th>
<th>Kappa Coefficient (95% C.I)</th>
<th>Fisher’s Exact p value</th>
<th>Adjusted p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amoxicillin</td>
<td>&gt;0.125</td>
<td>Sensitive</td>
<td>39</td>
<td>0</td>
<td>1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resistant</td>
<td>0</td>
<td>3</td>
<td>(1.0 – 1.0)</td>
<td>0.0007</td>
</tr>
<tr>
<td>Clarithromycin</td>
<td>&gt;0.5</td>
<td>Sensitive</td>
<td>26</td>
<td>0</td>
<td>0.84</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resistant</td>
<td>3</td>
<td>13</td>
<td>(0.67 – 1.0)</td>
<td>0.0007</td>
</tr>
<tr>
<td>Levofoxacin</td>
<td>&gt;1</td>
<td>Sensitive</td>
<td>17</td>
<td>1</td>
<td>0.9</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resistant</td>
<td>1</td>
<td>3</td>
<td>(0.77 – 1.0)</td>
<td>0.0007</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>&gt;1</td>
<td>Sensitive</td>
<td>38</td>
<td>0</td>
<td>-</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resistant</td>
<td>4</td>
<td>0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Metronidazole</td>
<td>&gt;8</td>
<td>Sensitive</td>
<td>20</td>
<td>2</td>
<td>0.76</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resistant</td>
<td>3</td>
<td>17</td>
<td>(0.56-0.96)</td>
<td>0.0007</td>
</tr>
<tr>
<td>Rifampicin</td>
<td>&gt;1</td>
<td>Sensitive</td>
<td>29</td>
<td>8</td>
<td>0.46</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resistant</td>
<td>0</td>
<td>5</td>
<td>(0.18-0.75)</td>
<td>0.0105</td>
</tr>
<tr>
<td></td>
<td>&gt;4</td>
<td>Sensitive</td>
<td>37</td>
<td>0</td>
<td>0.73</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resistant</td>
<td>2</td>
<td>3</td>
<td>(0.37 – 1.0)</td>
<td>0.0063</td>
</tr>
</tbody>
</table>

* Genetic resistance was defined as mutations near putative amoxicillin binding motifs.

* Genetic resistance was defined as truncation in rdxA or R16H mutation.

* Genetic resistance was defined as any mutation in rpoB.