



OPEN

First insight into the whole genome sequence variations in clarithromycin resistant *Helicobacter pylori* clinical isolates in Russia

Daria Starkova^{1,3✉}, Nikita Gladyshev¹, Dmitrii Polev², Alina Saitova², Svetlana Egorova¹ & Alena Svarval¹

Clarithromycin (CLR) is currently a key antibiotic for *Helicobacter pylori* infection treatment, however, the data on CLR resistance patterns in Russia are missing. Here, we applied WGS-based approach to *H. pylori* clinical isolates from Russia to comprehensively investigate sequence variation, identify putative markers of CLR resistance and correlate them with phenotypic susceptibility testing. The phenotypic susceptibility of 44 *H. pylori* isolates (2014–2022) to CLR was determined by disc diffusion method: 23 isolates were CLR-resistant and 21-CLR-susceptible. All isolates were subjected to WGS and submitted to GenBank. Based on complete sequence analysis, we showed that among all sequence variants, the combination of mutations A2146G/A2147G in the 23S rRNA gene is the most reliable for prediction of phenotypic susceptibility. For the first time, the average number of mutations in 106 virulence-associated genes between resistant and susceptible groups were compared. Moreover, this study presents the first WGS insight into genetic diversity of *H. pylori* in Russia with a particular focus on the molecular basis of drug resistance: the novel mutations were described as potential markers for the resistance development. Of these, the most prominent was a frameshift deletion (252:CGGGT) in HP0820 coding region, which is a good candidate for further investigation.

Helicobacter pylori infection is recognized as an important public health concern in more than 50% of the world's population and may be associated with such gastrointestinal diseases as chronic gastritis, gastric/duodenal ulcers, MALT-lymphoma and stomach cancer¹.

Various regimens for *H. pylori* eradication therapy are used worldwide, while the choice of appropriate eradication therapy is determined by the local resistance patterns². Conventional triple therapy (clarithromycin (CLR), proton pump inhibitor (PPI), metronidazole (MTZ) or amoxicillin (AMX)) may be recommended as first-line treatment in areas of low (< 15%) CLR resistance, while quadruple therapy (bismuth, PPI, MTZ, tetracycline) if the local CLR resistance is over 15%^{2–4}. According to the recently published Maastricht VI guidelines, routine susceptibility testing (molecular or after culture) is recommended even before prescribing first-line treatment. However, systematic antimicrobial susceptibility testing (AST) in clinical practice remains controversial due to the practical, economical and logistical issues that must be considered^{3,5}.

To date, Russia belongs to countries with a high prevalence of *H. pylori* infection (65–90% depending on the region)⁶. According to a meta-analysis that summarized studies of *H. pylori* antibiotic resistance in various regions of Russia over 10 years (2011–2020), the resistance rate to CLR reached 10.39%. That indicates a low resistance of the pathogen (< 15%) and allows to consider the triple scheme of treatment as the first-line empirical therapy (based on the absence AST data) in our country⁷. However, the empirical therapy frequently results in treatment failure, which, in turn, leads to the development of antibiotic resistance of the pathogen.

¹Laboratory of Identification of the Pathogens, St. Petersburg Pasteur Institute, St. Petersburg, Mira Street, 14 197101, Russia. ²Metagenomics Research Group, St. Petersburg Pasteur Institute, St. Petersburg, Mira Street 197101, Russia. ³Laboratory of Identification of the Pathogens/Laboratory of Molecular Epidemiology and Evolutionary Genetics, St. Petersburg Pasteur Institute, St. Petersburg, Mira Street 197101, Russia. ✉email: dariastarkova13@gmail.com

During the last decade, many molecular approaches (such as polymerase chain reaction (PCR), Sanger sequencing, whole genome sequencing (WGS) were used to identify known and novel markers for prompt prediction of the phenotypic antimicrobial resistance. It is generally accepted that point mutations A2142G and A2143G (according to Taylor et al. numbering) in V domain of the 23S rRNA gene are mainly responsible for *H. pylori* resistance to CLR⁸. At the same time, point mutations outside these positions (T1942C, G1939A, C2147G, G2172T, T2182C, A2116G, A2144G/T, A2115G, G2111A, T2117C, G2141A, T2182C, T2717C, T2289C, G2224A, C2245T) vary geographically and their role in CLR resistance is not yet clear^{9,10}. Based on WGS results, some sequence changes in *rpl22* (encoding a ribosomal protein that interacts with the 23S rRNA domains) and *infB* (encoding a translation initiation factor, IF-2) genes in CLR-resistant *H. pylori* strains were described¹¹. In addition, four gene clusters of efflux pump systems (HP0605-HP0607 (*hefABC*), HP0971-HP0969 (*hefDEF*), HP1327-HP1329 (*hefGHI*), and HP1489-HP1487) were identified as the resistance-nodulation-cell division (RND) family and are expected to be involved in development of CLR resistance. Some studies suggest that the number of single nucleotide variants (SNVs) in the RND family is significantly higher in CLR-resistant than in CLR-susceptible *H. pylori* isolates^{8,12,13}.

Due to the lack of data on CLR resistance patterns in Russian *H. pylori* clinical isolates, we aimed to comprehensively investigate sequence variations by WGS-based approach to detect putative markers of CLR resistance in *H. pylori* clinical isolates and evaluate the correlation between genotypic and phenotypic AST.

Results

Phenotypic AST of study cohort

For our analysis we selected two types of isolates: resistant (inhibition zone diameter < 10 mm) and susceptible (> 41 mm). Isolates with intermediate susceptibility were not included in our study. Of 44 *H. pylori* clinical isolates, 27 were obtained from newly diagnosed patients and 17—from previously treated patients after eradication failure.

As a result, 23 out of 44 *H. pylori* isolates were phenotypically determined as CLR-resistant, while 21 isolates showed susceptibility to CLR. All isolates from previously treated patients were CLR-resistant by phenotypic AST.

Plasmid detection in *H. pylori* isolates

Of the 44 *H. pylori* genomes analyzed in our study, 40 comprised hypothetical proteins, recognized by annotation tools as plasmid sequences. However, no resistance genes in putative plasmids were found.

Single nucleotide polymorphisms in 23S rRNA genes

Among 44 *H. pylori* isolates, 21 nucleotide substitutions were identified in 23S rRNA gene, however, only one of them—A2147G (aka A2143G)—was significantly associated with phenotypic resistance of *H. pylori* to CLR: of the 23 CLR-resistant isolates, 13 (56.5%) carried the 2147G allele ($p = 0.00025$) (Table 1). However, A2147 allele was related to CLR susceptibility in 95.2% (20/21) of cases, since one strain phenotypically determined as CLR-susceptible possessed A2147G mutation.

The other point mutation—A2146G (aka A2142G)—was found in 13.0% (3/23) of CLR-resistant isolates ($p = 0.33$). The A2146C allele was not detected in our study. All CLR-susceptible isolates (100%; 21/21) possessed A2146 allele. None of the CLR-resistant isolates carried two A2146G/A2147G mutations simultaneously: all 2147G isolates carried the A2146 allele and, vice versa, all 2146G isolates carried the A2147 allele. Seven CLR-resistant isolates had neither A2146G nor A2147G mutations. The agreement between genotypic and phenotypic susceptibility to CLR based on both A2146G/A2147G mutations was 69.6%.

Gene (locus)	Nucleotide change	Amino acid change	CLR-R (n=23), N (%)	CLR-S (n=21), N (%)	p-value
23S rRNA	A2146G	–	3 (13.0)	0	0.23414
	A2147G	–	13 (56.5)	1 (4.8)	0.00025
<i>infB</i> (HP1048)	G2763T	Asp921Glu	2 (8.7)	8 (38.1)	0.03099
<i>omp5</i> (HP0227)	7:insTTTTGA	3:insIleLys	5 (21.7)	0	0.04972
<i>omp32</i> (HP1501)	A2C	Start codon lost	5 (21.7)	0	0.04972
<i>omp25</i> (HP1156)	C1400T	Ala467Val	5 (21.7)	0	0.04972
HP0820	252:delCGGGT/262:delGG or GGG insTTAGCACA	delGly85-fs/delGly88-fs	7 (30.4)	0	0.00943
<i>flaB</i> (HP0115)	C783A	Glu261Asp	6 (26.1)	0	0.02198
<i>flaD</i> (HP0752)	C1997G	Pro666Arg	1 (4.3)	7 (33.3)	0.01900
<i>spaB</i> (HP0600)	G98A	Thr33Ile	12 (52.2)	4 (19.0)	0.03069
<i>res2</i> (HP1521)	C1T	Val1Met	6 (26.1)	1 (4.8)	0.09732
<i>recA</i> (HP0153)	360TGATACA-366AAGCACG	Asp121Ser	8 (34.8)	2 (9.5)	0.07249
<i>rpl10</i> (HP1200)	C257T	Gly86Asp	8 (34.8)	2 (9.5)	0.07249

Table 1. nsSNVs and Indels in coding regions of CLR-resistant and CLR-susceptible *H. pylori* clinical isolates compared to the *H. pylori* 26695 reference genome.

Out of 17 *H. pylori* isolates from previously treated patients, 13 (76.5%) possessed either A2146G or A2147G mutations. This observation is entirely consistent with the general agreement on increasing the resistance to CLR during treatment regimen.

Since the *H. pylori* genome contains two copies of the 23S rRNA gene, we assessed the presence of putative allelic variants due to mutational differences in each copy. A survey of the depth of mapped reads showed that 2146G and 2147G alleles were present in 100% reads of 15/16 CLR-resistant isolates which indicates that mutations has occurred in both copies of 23S rRNA. However, one CLR-resistant isolate (HP220) demonstrated a heterogeneity between two copies of 23S rRNA: the 2147G allele was detected in only 50% reads of one copy, while the second showed 80.4% reads of A2147 allele.

The remaining mutations outside 2146/2147 positions in 23S rRNA gene were found in *H. pylori* isolates regardless of phenotypic susceptibility (Table S1).

Nucleotide variants in *rpl22*, *infB* and RND family genes

Since the presence of four RND efflux clusters (HP0605-HP0607, HP0971-HP0969, HP1327-HP1329, HP1487-HP1489) and certain genetic variants in *rpl22* (HP1314) and *infB* (HP1048) genes have been reported to be presumably associated with CLR-resistance, we analysed the distribution of sequence variations in these genes between two phenotypic groups as well as between the A2146G/A2147G-mutant isolates and the wild type^{8,10,11,13}.

As a result, we found that 12 efflux pump genes were present in all 44 *H. pylori* genomes. Moreover, we did not reveal any significant differences in either individual SNVs or total SNVs number between CLR-resistant and susceptible groups.

Due to analysis of *infB* gene sequence variations we did not identify any of C718T, A1899G, G160A, G60A mutations, which were considered to be involved in increasing antibiotic resistance^{9–11}. Among all types of SNVs, only one—C2763A was significantly predominated in CLR-susceptible isolates ($p = 0.03$; OR: 4.6743) (Table 1). The other synonymous variant G2118A was found specifically in CLR-resistant isolates (26.1%; $p = 0.021$) however, silent mutations are not likely to affect the phenotype. We also did not find significant differences in total number of missense mutations in *infB* gene neither between CLR-resistant and susceptible groups nor A2146G/A2147G-mutant isolates and the wild type.

The most interesting finding among genetic variants in *rpl22* gene is the absence of two types of mutations—3 bp (GTG) deletion and 9 bp (TTCCATGTA) insertion—which have been detected in numerous studies^{8–11,33}. Instead, we identified three missense variants, none of which were associated with phenotypic resistance/susceptibility status (Table S1).

Distributions of virulence-associated genes in *H. pylori* clinical isolates

To investigate whether the presence or absence of virulence-associated genes is related to phenotypic resistance *H. pylori* to CLR, we evaluated the presence of *cag* pathogenicity island (*cagPAI*) genes (*cag1-cag26*), vacuolating cytotoxin gene (*vacA*), ulcer-associated restriction endonuclease (induced by contact with epithelium) and adenine specific DNA methyltransferase genes (*iceA1*, *hpyIM*), urease gene cluster (*ureABCIEFGH*), gamma-glutamyltranspeptidase gene (*ggt*), virulence factor protein (*mviN*), outer membrane protein family (*omp1-omp32*, *ompP1*), toxin-like and iron-regulated outer membrane proteins genes (*fecA*, *frpB*), flagellar family genes (*flaA*, *flaB*, *fla*, *flaA1*, *flgH*, *fliF*, *flgI*, *fliE*, *flgB*, *flgC*, *flgG*, *flhA*, *fliH*, *fliP*, *fliR*, *fliI*, *flgE*, *flgK*, *fliD*, *motA*, *motB*, *fliG*, *fliM*, *flaG*, *hpaA*, *fliN*), heat shock protein family genes (*hslU*, *hslV*, *hspX*, *ibpB*) in two groups of *H. pylori* isolates. Since the *cagA* (*cag26*) gene is considered to be a marker of *cag* pathogenicity island, the presence of *cagPAI* region was assessed by detection of *cagA* gene.

As a result, the majority of virulence genes were certainly observed in all 44 *H. pylori* genomes. Remarkably, only the presence of *cagPAI* in CLR-susceptible group was higher than that in CLR-resistant group (16/21 vs. 13/23), however, not significantly ($p = 0.42$) (Fig. 1). Interestingly, that none of Russian isolates contained *cag2*

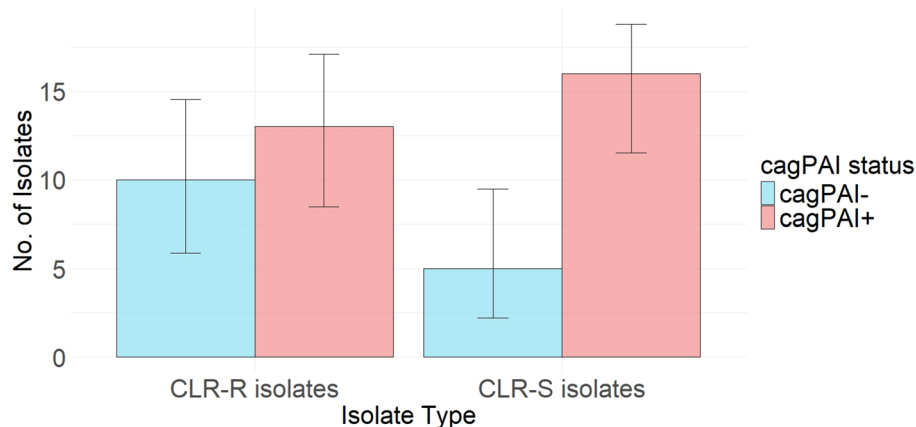


Fig. 1. Comparison of the presence of *cag* pathogenicity island (*cagPAI*) between resistant and susceptible groups of *H. pylori* isolates.

(HP0521) gene and all 44 *H. pylori* isolates regardless of antimicrobial susceptibility harbored A589G mutation in the *cag13* (HP0534) gene leading to the loss of stop codon and splicing with coding region HP0533.

Liu et al. (2022) found that the average frequency of the iron-regulated OMP genes (*fecA* and *frpB*) in the CLR-resistant group was significantly lower than that in the CLR-susceptible group ($p = 0.01$)²⁷. In our study, both genes were present in all *H. pylori* isolates, and the average number of mutations in the *fecA* and *frpB* genes was not associated with any phenotypic group.

For the first time, we also compared the average number of all fsIndel and nonsynonymous variants in 106 genes listed above between CLR-resistant and CLR-susceptible groups of *H. pylori* isolates. There were no statistical differences in the average number of mutations between two phenotypic groups ($p > 0.05$).

Novel sequence variations in coding regions of Russian *H. pylori* isolates

To detect new putative markers of phenotypic resistance/susceptibility to CLR we conducted a global gene analysis of all *H. pylori* isolates compared to the *H. pylori* 26695 reference genome. We investigated sequence variants assessed by the annotation tool as having a significant impact on the translated protein and evaluated their association with the antibiotic phenotype. The majority of variants were equally shared between two phenotypic groups of isolates. However, we were able to detect five novel resistance-associated variants and variants that were predominated in either resistant or susceptible phenotypic group. These findings were summarized in Table 1 and Table S1.

Among OMP (Outer Membrane Protein) family genes the most interesting findings were related to *omp32* (HopW), *omp5* and *omp25* (HopI) genes. Missense mutation A2C resulting in loss of start-codon in the *omp32* gene, conservative inframe insertion 7:insTTTTGA in the *omp5* gene, and nsSNV C1400T in the *omp25* gene were detected in 21.7% (5/23) CLR-resistant isolates per each ($p = 0.049$). None of these mutations were found in CLR-susceptible isolates.

Several unique frameshift (fs) alterations in HP0820 coding region were detected in 30.4% (7/23) CLR-resistant isolates. A fs deletion 252:delCGGGT was linked with another fsIndel in HP0820 onwards: deletion of two (262–263delGG) or three (262–264delGGG) nucleotides and insertion of TTAGCACA fragment.

Among 28 flagellar family genes three nsSNVs in *flaB* gene were found predominantly in resistant group, however, only one of them (C783A) was significantly associated with resistance ($p = 0.021$). Alternatively, we revealed that the prevalence of nsSNV (C1997G) in the *fljD* gene was significantly higher in the CLR-susceptible rather than CLR-resistant group (33.3 vs. 4.3%, respectively) ($p = 0.019$, OR: 0.1057741).

Three multidrug resistance protein genes (*hetA*, *msbA*, *spaB*) were analysed in present study. As a result, the proportion of CLR-resistant *H. pylori* isolates carried missense mutation G98A in multidrug resistance protein SpaB gene was significantly higher than that in CLR-susceptible isolates [52.2 vs. 19.0% ($p = 0.030$; OR: 4.3874)]. Interestingly, using WGS analysis, Liu et al. (2022) showed that the *spaB* was absent in most strains isolated from Shanghai, whereas in our study this gene was certainly observed in all clinical isolates²⁷.

In addition, the same report showed that four variations (S275N, D353N, 120-121delGG and ins116G117) within the integrase/recombinase gene *xerD* were significantly associated with both CLR and MTZ resistance²⁷. In the present study, S275N mutation was found without significant differences between resistant and susceptible phenotypes (82.6 vs. 80.9%; $p > 0.05$) (Table S1). Instead of 120-121GG deletion, we identified missense multiple substitutions of four nucleotides (117-124CTAGGATT → TTAATATC) that were linked together and have prevailed in resistance (34.8%) compared to susceptible (19.0%) group, however, not significantly [$p = 0.32$; OR: 2.1937]. Insertion 116G117 and missense variant D353N were not detected in our study.

Five point substitutions within recombinational repair protein RecA gene were linked together and were carried predominantly by resistant (34.8%) rather than susceptible (9.5%) isolates [$p = 0.072$; OR: 4.674359]. The prevalence of missense variant (CIT) in the type III restriction enzyme gene (*res2*) leading to the replacement of alternative start-codon Val with standard Met was higher in resistant group, however, not significantly [$p = 0.097$; OR: 0.1472963]. It is known that the changing GUG into AUG results in a several-fold increase in the translation efficiency²⁸.

Since *rpl* are considered to be the genes with the greatest value, we initially conducted the analysis of 49 ribosomal proteins genes. As a result, we found one nsSNV (C257T) within *rpl10* gene (encodes ribosomal protein L10/50S) that was predominated in CLR-resistant group, but not significantly [$p = 0.072$; OR: 4.6743].

Further analysis of the distribution of newly detected mutations specific for resistant group in the A2146G/A2147G-mutated *H. pylori* clinical isolates revealed that 6 of 7 (85.7%) HP0820-mutated isolates possessed either 2146G or 2147G alleles (Table 2). In total, more than half (67.8%) of CLR-resistant isolates with mutations in five coding regions (*omp5*, *omp25*, *omp32*, HP0820, *flaB*) were A2146G/A2147G-positive.

Gene (mutation)	A2146G/A2147G- positive isolates (n = 16), N (%)	A2146G/A2147G-negative isolates (n = 7), N (%)	Total no. of CLR-R isolates (n = 23), N (%)
<i>omp5</i> (7:insTTTTGA)	3 (18.8)	2 (28.6)	5 (21.7)
<i>omp32</i> (A2C)	3 (18.8)	2 (28.6)	5 (21.7)
<i>omp25</i> (C1400T)	3 (18.8)	2 (28.6)	5 (21.7)
HP0820 (252:delCGGGT)	6 (37.5)	1 (14.3)	7 (30.4)
<i>flaB</i> (C783A)	4 (25.0)	2 (28.6)	6 (26.1)

Table 2. The distribution of newly detected mutations in A2146G/A2147G-positive and A2146G/A2147G-negative CLR-resistant *H. pylori* clinical isolates.

Phylogenetic analysis and relationship between isolate relatedness and CLR-resistance patterns in *H. pylori* isolates

Based on 153,292 core SNVs of 44 *H. pylori* genomes we constructed the phylogenetic tree based on the maximum likelihood estimation with respect to CLR-resistance patterns (Fig. 2). As a result, no isolates relatedness associated with the antibiotic phenotype was revealed. Instead, these isolates had distinct resistance profiles regardless of their genetic relatedness.

Discussion

Despite insufficient knowledge and interregional dynamic differences in the resistance of *H. pylori* to the main antimicrobials, the level of the primary CLR resistance in Russia is low (10.39%) and allows to consider a standard triple eradication regimen as a first-line therapy. Thus, routine *H. pylori* susceptibility testing usually is not provided in clinical settings, while empirical therapy may result in eradication failure as well as increasing resistance rate. According to the latest data (2020), the prevalence of primary CLR resistance of *H. pylori* in Moscow reached 10.87%, in Smolensk—5.74%, in Kazan—10.0% and in St. Petersburg—22.26%^{29,30}. Such a difference indicates the wide variation in resistance rates in Russia, highlighting the need to ongoing examination of CLR resistance as well as cure rates in each geographic region.

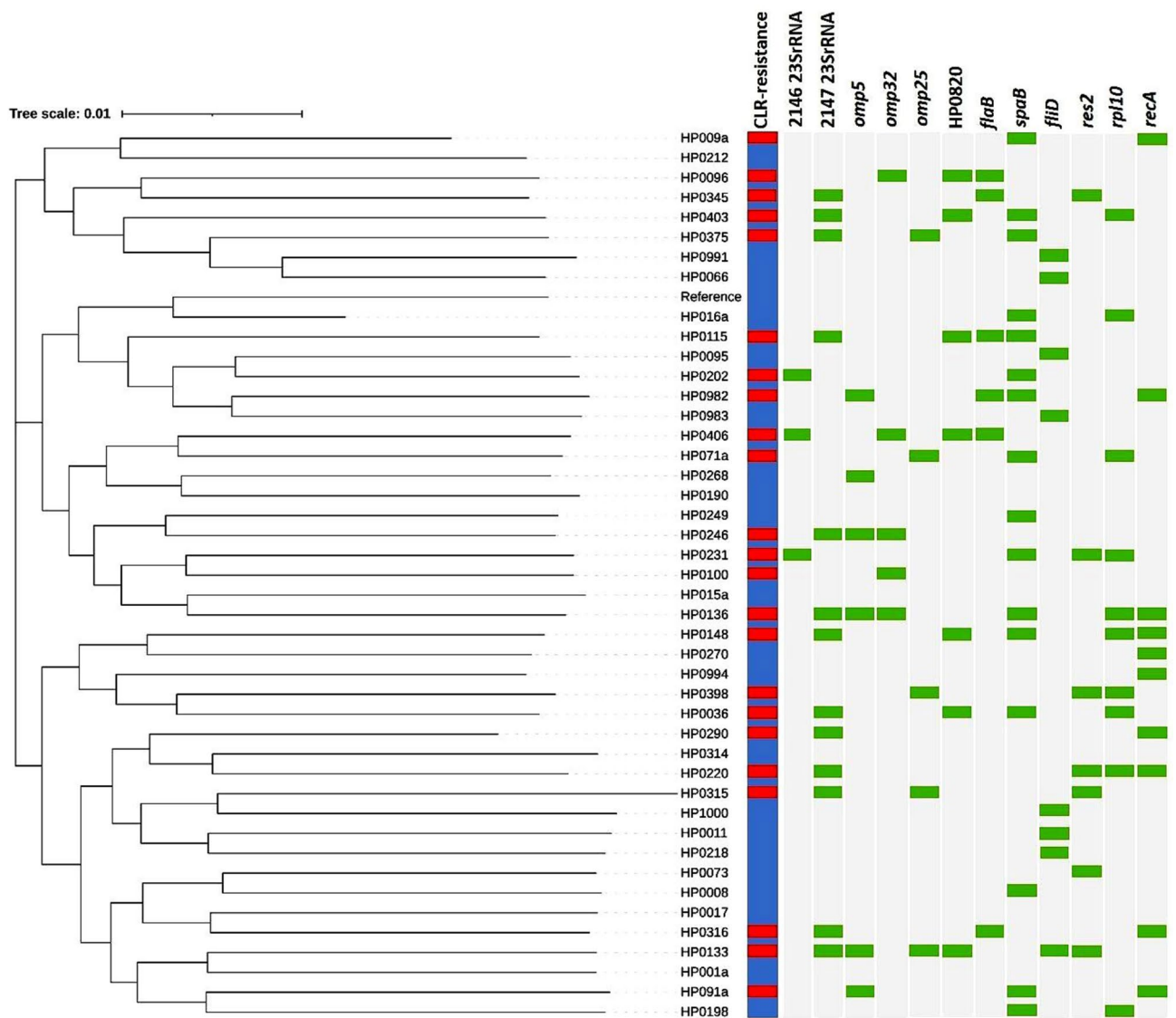


Fig. 2. Phylogenetic analysis of 44 Russian *H. pylori* genomes with respect to CLR-resistance patterns. Maximum likelihood tree based on core SNVs alignment from draft genome mapping against *H. pylori* 26695 reference strain and resistance profile corresponding to specific genetic determinants that differ between CLR-susceptible and CLR-resistant groups of *H. pylori* isolates. The susceptible and resistant patterns are denoted by blue and red rectangles; the presence and absence of the specific loci are denoted by green and grey rectangles, respectively.

For our study, we used conventional disc diffusion method to detect susceptibility of *H. pylori* isolates to CLR. Although the EUCAST has recommended the E-test for MIC breakpoints detection, the high cost of E-test strips makes it difficult to use. Recently, Tang et al. (2020) assessed the disk diffusion technique against E-test and estimated that susceptibility agreement between two methods was 96.0% for CLR¹⁴. It was demonstrated that disc diffusion is simple and affordable alternative to E-test for in vitro CLR susceptibility testing^{31,32}.

The mechanism of drug resistance in *H. pylori* is believed to be due to mutations located in the chromosome and not related to horizontal transfer of plasmids or other mobile elements³³. It is generally accepted that *H. pylori* resistance to CLR is associated with two point mutations A2146C/G and A2147G (in accordance to *H. pylori* reference strain 26695 numbering) in the V domain of 23S rRNA gene that decreases the affinity of drug binding to the ribosome by disrupting the peptidyl transferase loop conformation⁶. In the present study, A2147G was the most frequent mutation found in CLR-resistant *H. pylori* isolates (56.5%) and highly correlated with CLR resistance. However, A2146G mutation has been detected in only 13% CLR-resistant isolates that does not allow us to assess the actual role of this mutation in the resistance development. The combination of mutations A2146G and A2147G was detected in 69.5% isolates, which is consistent with prior studies conducted in Europe demonstrated that both mutations were present in 60–90% of CLR-resistant isolates^{8,33,34}. In addition, we did not find any isolates carrying both mutations simultaneously, which also is in agreement with most previous observations^{8,9,34,35}. The high rate of CLR-resistant isolates without A2146G/A2147G mutations (30.5%) in our study indicates that these isolates may have genetic determinants other than 23S rRNA gene responsible for pathogen resistance development.

Notably, one isolate phenotypically determined as CLR-susceptible in our study possessed A2147G mutation. Such a case may indicate heteroresistance caused either by evolutionary changes in a single strain or by the presence of mixed susceptible and resistant isolates. As shown by our previous study, this isolate was positive for mixed combinations of *vacA* alleles (s1s2 and i1i2) indicating the coexistence of several genetically different *H. pylori* isolates in gastric sites as a result of mixed infection⁶.

Furthermore, there are two copies of the 23S rRNA operon within the *H. pylori* genome. Sequence differences between both copies have rarely been studied and data on whether 2146/2147-mutations in single or both copies of 23S rRNA are necessary for the strain resistance to occur are still controversial³⁵. In this study, 16 CLR-resistant isolates possessed A2146G/A2147G mutations, and only one of them demonstrated a heterogeneity between two copies of 23S rRNA: the 2147G allele was detected in 50% reads of one copy, while the second copy showed 80.4% reads of A2147 allele. This observation suggested that *H. pylori* isolates should be considered resistant if they have the mutation in at least one copy of 23S rRNA gene.

Other genes assumed to be involved in CLR resistance are *rpl22* and *infB*. In fact, several point mutations in *infB* (C718T, A1899G, G160A, G60A) as well as 9-bp insertion and/or 3-bp deletion in *rpl22* genes have been described, which in combination with 23S rRNA mutations were predicted to increase the resistance level of *H. pylori* isolates^{11,36}. Interestingly, no indels in *rpl22* gene were detected in our study: instead, we found 3-bp deletion in *infB* gene primarily in CLR-susceptible isolates, but not significantly. Notably, we did not find any of point mutations in *infB* gene that were detected in other studies. Among all SNVs in *infB*, no substitutions unique for resistant isolates were identified. Altogether, these data provide further evidence that high genetic diversity of *H. pylori* populations is the crucial factor in the searching for putative determinants of resistance.

Another possible mechanism for CLR resistance is multidrug efflux pump systems. It has been reported that four conserved RND families of efflux pump transporters may be responsible for development of macrolide resistance in *H. pylori* by reducing intracellular antimicrobial concentration^{8,10}. As shown by Iwamoto et al. (2014), there is a significant difference in the number of SNVs in *hefABC* (hp0605–hp0607) cluster between CLR-resistant and susceptible groups, as well as CLR-resistant isolates are more prone to single-nucleotide variants in all four clusters of efflux genes¹³. Moreover, Chen et al. (2018) revealed that the number of mutations in the RND family was significantly higher in *H. pylori* isolates harboring the A2147G point mutation in 23S rRNA gene⁹. In our study, we did not find any significant differences in number of SNVs among 12 genes neither between resistant and susceptible phenotypes, nor between isolates carried A2147G mutation and the wild type. Noteworthy, among all types of genetic variants, we found only single-nucleotide substitutions: no deletions or insertions were detected in four clusters of efflux genes. This result is consistent with the hypothesis that SNVs in efflux pump genes may contribute to CLR resistance through a synergistic effect between 23S rRNA mutations and efflux pumps.

Giving the enormous genetic diversity of *H. pylori* and the high heterogeneity of resistance genotypes across geographical regions, identification of novel resistance patterns as well as interpretation of the obtained results appear to be challenging. WGS-based approaches have enabled to provide a comprehensive view of bacterial genotypes and are particularly useful for tracking novel genetic determinants responsible for antimicrobial resistance of the pathogen¹⁰.

To identify novel mutations potentially contributing to antibiotic resistance in *H. pylori* isolates we next sought to analyze genetic variants in other gene families presumably associated with virulence and multidrug resistance. In the current study, the largest number of mutations specific for resistant group was observed among OMP family genes: *omp5*, *omp25* and *omp32*. It is known that OMPs are associated with macrolide resistance in Gram-negative bacteria. Lin et al. (2023) suggested that alteration of outer membrane proteins is a potential mechanism for CLR resistance development in *H. pylori*³⁶. Moreover, aside from mutations, the expression of the *omp25* and *omp32* genes was reported to be down-regulated by long-term low-dose of AMX, which has a significant impact on the treatment efficacy^{36,38}. Interestingly, Thompson et al. (2003) showed the close co-expression of the *cagA* gene with the *omp5* and *omp29* genes, suggesting these OMPs are involved in the secretion and/or activation of CagA³⁷. Among all variants in the OMP family genes we identified, the missense mutation in the anion-selective porin *Omp32*, leading to the loss of the start codon, had the greatest impact on the gene.

Since it is unknown whether the accumulation of such mutations may indirectly be involved in the resistance development, its hypothetical impact needs to be further investigated.

Another interesting finding: resistance-associated fs deletion in the coding region HP0820 (*cj1247c* in *C. jejuni*), which Porcelli et al. (2013) described as a predicted leaderless mRNA in all species containing this gene, which is always upstream of the *uvrC* DNA repair gene. Genes translated from the leaderless mRNAs have been shown to encode the proteins involved in stress responses, as well as outer membrane efflux and multidrug efflux pump proteins³⁹. It is well known that antibiotics create conditions that favor selection of beneficial mutants with elevated mutation rates and this is most often linked to the DNA repair system and SOS response, that can lead to increased activity of associated multidrug efflux pumps. Given the fact that *H. pylori* has a defective DNA mismatch repair system (lacking genes as *mutS*, *mutL*, and *mutH* implicated in this pathway), as well as lack of the SOS regulon, the antibiotic-induced mutagenesis in various types of regulatory genes increases the genetic diversity of bacterial populations and, as a consequence, enhances the evolution of antibiotic resistance^{35,40}.

Missense mutations in two different flagellar genes *flaB* and *fliD* may also be candidates indirectly involved in underlying mechanisms of resistance and/or susceptibility of *H. pylori*. The pathogenesis of *H. pylori* infection is thought to be determined in part by flagellar motility, which has a direct impact on colonization, inflammation, and immune evasion. Flagellar filament FlaB and the filament-capping protein FliD are one of the main structural proteins of the *H. pylori* flagellum and play an important role in bacterial motility^{41,42}. Even though the actual impact of these mutations on resistance development is difficult to assess, the mutational buildup in the genome may be a result of response to antimicrobials and an important predictor of treatment failure.

As for relationship between virulence-associated genes and phenotypic resistance, we determined that there is no significant correlation between the presence of main virulence factors and resistance to CLR. On the one hand, our findings are consistent with majority of previous reports^{27,43,44}. However, it is worth noting that the number of *cagPAI*-negative isolates in the CLR-resistant group predominated roughly twofold compared to those in the susceptible group (10/23 vs. 5/21). Thus, on the other hand, this observation is consistent with other findings showing that CLR resistance is obviously higher in *cagA*-negative than in *cagA*-positive isolates^{45,46}.

The present study has several limitations. First, despite the high level of agreement between disc diffusion method and E-test, the main drawback in this research is the inability to determine MIC values, which does not allow us to reveal a causal relationship between mutations and high/low MIC values of phenotypic resistant isolates. Second, the sample size is rather small and obtained results do not necessarily reflect resistant patterns in general Russian *H. pylori* population. Third, we did not provide further experimental analysis to confirm the involvement of newly detected mutations in CLR resistance development. Nevertheless, in spite of these limitations obtained data provided the foundation for the future investigations.

Conclusions

This study presents the first WGS insight into genetic diversity of *H. pylori* in Russia with a particular focus on the molecular basis of antibiotic resistance. For the first time, the average number of all fsIndel and nonsynonymous variants in 106 virulence-associated genes between resistant and susceptible groups was compared. It has been shown that among all genetic variants we obtained, the combination of mutations A2146G/A2147G in the 23S rRNA gene is the most reliable for prediction of phenotypic AST. Sequencing data did not reveal the involvement of nucleotide variants in any efflux pump, *infB* or *rpl22* genes in phenotypic prediction of CLR resistance, emphasizing the enormous genetic diversity of *H. pylori* populations. The novel mutations in other genes were described in our study as potential markers for the resistance development. Among them, the most prominent mutation is frameshift deletion (252:CGGGT) in HP0820 coding region, which is a good candidate for further investigation.

Materials and methods

Study design and sampling collection

In total, 44 *H. pylori* clinical isolates collected between 2014 and 2022 were retrieved from the bacterial strain collection of the St. Petersburg Pasteur Institute, St. Petersburg, Russia. All strains were isolated from gastric biopsy specimens taken during endoscopy from the separate adult patients with chronic gastritis (n = 32), duodenal ulcer (n = 11) and gastric cancer (n = 1). The patients were 18 men (40.9%) and 26 women (59.1%); median age 45.8 ± 4.5 years (range 22–70 years). For this study, *H. pylori* isolates were deliberately selected based on their CLR-resistance phenotype and hence are not an epidemiologically representative sample of *H. pylori* primary antibiotic resistance.

Endoscopic biopsy specimens were homogenized and used for the culture on *H. pylori*-selective medium (Columbia agar base with the addition of 5–7% defibrinated horse blood and 1% IsoVitalex solution). The bacterial cultures were incubated at 37 °C under microaerophilic conditions (10% CO₂, 85% N₂, 5% O₂, GasPak 100, BD Biosciences, USA). After 7–10 days incubation *H. pylori* colonies were identified by microscopy of Gram-staining culture smears and biochemical tests (urease, catalase, and oxidase). Final identification was made using MALDI-TOF MS (Autof MS1000, China). The *H. pylori* cultures were stored at –80 °C for further examination.

Ethics approval

The retrospective study was approved by the Independent Ethics Committee of the St. Petersburg Pasteur Institute, Russia (protocol No. 50/04-2019, 22.06.2020). All patient-related data were treated anonymously. All methods in our study were conducted in accordance with relevant guidelines and regulations.

Phenotypic antimicrobial susceptibility testing (AST)

The susceptibility of *H. pylori* isolates to CLR was determined by conventional disc diffusion method. The test was performed by direct suspension of a 72-h cultivated *H. pylori* strains in Müller-Hinton broth adjusted to a density of 0.5 OD₆₀₀ or $1-2 \times 10^8$ cells CFU/ml according to the McFarland scale. Bacterial suspension (0.1 ml) was inoculated onto plates with fastidious Mueller Hinton agar medium supplemented with 5% defibrinated horse blood and evenly distributed over the surface with a spatula. Immediately after inoculation, 6-mm-diameter discs containing CLR (15 µg) were placed on the plate surface (1 disk per plate). After incubation at 37 °C under microaerophilic conditions for 72 h, the inhibition zone diameters were measured in millimeters (mm). *H. pylori* NCTC 12,823 strain was used as a control for susceptibility to CLR. To date, there are no standardized cutoff values for disk diffusion method. Using the linear regression analysis, Tang et al. (2021) calculated the inhibition zone diameter breakpoint to be 41 mm for CLR corresponding to the EUCAST MIC cutoff value of 0.5 mg/L¹⁴. Such inhibition zone diameter, obtained using the 15-µg CLR disks, also correlate well with the MIC by E-test ($r = -0.894$).

DNA extraction and sequencing

The total DNA of *H. pylori* isolates was extracted using the QIAamp DNA Mini Kit (QIAGEN GmbH, Germany) according to the manufacturer's guidelines. The DNA concentration of each sample was quantified on a Qubit 4.0 fluorometer. Whole-genome shot-gun DNA libraries were prepared using MGIEasy FS DNA Library Prep Set and then sequenced on a DNBSEQ-G50 sequencer (MGI Tech Co. Ltd, Beijing, China) in a 2*100 paired-end (PE) mode.

Bioinformatics analysis

The raw paired-end reads were initially analyzed using FastQC software (v.0.12.1; Babraham Institute, Cambridge, UK) to assess their acceptability for further analysis. The reads were trimmed to remove adapters and low-quality sequences (Q-score < 20) and filtered by Trim Galore! (version 0.6.7). Bacterial genomes were assembled de novo using SPAdes assembler software (version 3.13.1)¹⁵, and the results were evaluated with QUAST (version 5.2.0)¹⁶.

To identify plasmid-encoded resistance genes, trimmed FASTQ files were analyzed by plasmidSPAdes pipeline (Version 3.15.4)¹⁷. Next, high-quality reads were merged by PEAR software¹⁸ installed via Conda environment¹⁹ and aligned to the *H. pylori* 26695 reference genome, available at NCBI GenBank under the accession number AE000511.1. To evaluate the genetic variations between *H. pylori* isolates and reveal potential genotype-to-phenotype correlations, the insertions/deletions (indels) and SNVs were called from alignments using Snippy pipeline v.4.6.0 (<https://github.com/tseemann/snippy>) with the following parameters: minimum base quality score—30, minimum mapping quality—60, minimum coverage—20, proportion for variant evidence—0.9²⁰. CLR-resistant isolates were also assessed for any novel point mutations. Putative impact of identified variants was predicted using SnpEff software v.4.3²¹. Aligned nucleotide sequences were visually analyzed using UGENE v. 38.1²². Particular genome features were retrieved from the UniProt platform (<https://www.uniprot.org/>)²³. Mutations clustering analysis and a maximum likelihood tree was constructed by using the RAXML-NG is a phylogenetic tree inference tool based on core genome obtained using Snippy core²⁴. The reliability of the phylogenetic tree branches was evaluated by bootstrapping method with 1000 replications.

All genome assemblies were deposited to the NCBI and available under BioProject “Whole-genome sequence variations in Russian *Helicobacter pylori* isolates” (Accession: PRJNA1011037 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1011037>)).

Statistical analysis

All statistics and data visualization were performed in R programming language, Environment version 4.3.1 with the following packages: tidyverse v. 2.0.0., epitools v. 0.5-10.1, psych v. 2.3.6^{25,26}.

The association between genotypic and phenotypic groups was screened using Chi-square and Fisher's exact tests. The significance level of the differences was set at $\alpha = 0.05$. We used two-by-two tables to calculate the odds ratio (OR).

Data availability

All data of this study are presented in the article and supplementary material. The datasets analysed during the current study are available in the NCBI repository and available under BioProject “Whole-genome sequence variations in Russian *Helicobacter pylori* isolates” (Accession: PRJNA1011037 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1011037>)).

Received: 26 January 2024; Accepted: 22 August 2024

Published online: 29 August 2024

References

- Mannion, A. et al. *Helicobacter pylori* antimicrobial resistance and gene variants in high- and low-gastric-cancer-risk populations. *J. Clin. Microbiol.* **59**(5), e03203–e3220. <https://doi.org/10.1128/JCM.03203-20> (2021).
- Suzuki, S., Kusano, C., Horii, T., Ichijima, R. & Ikehara, H. The ideal *Helicobacter pylori* treatment for the present and the future. *Digestion* **103**(1), 62–68. <https://doi.org/10.1159/000519413> (2022).
- Malfertheiner, P. et al. European *Helicobacter* and microbiota study group. Management of *Helicobacter pylori* infection: The Maastricht VI/Florence consensus report. *Gut* <https://doi.org/10.1136/gutjnl-2022-327745> (2022).
- Thung, I. et al. Review article: The global emergence of *Helicobacter pylori* antibiotic resistance. *Aliment Pharmacol. Ther.* **43**(4), 514–533. <https://doi.org/10.1111/apt.13497> (2016).

5. Mestrovic, A. *et al.* Personalized approach in eradication of *Helicobacter pylori* infection. *Antibiotics (Basel)* **12**(1), 7. <https://doi.org/10.3390/antibiotics12010007> (2023).
6. Svarval, A. V., Starkova, D. A., Ferman, R. S. & Narvskaya, O. V. Genetic polymorphisms of *Helicobacter pylori* clinical isolates in St. Petersburg, Russia. *Russ. J. Inf. Immunol.* **12**(2), 315–322. <https://doi.org/10.15789/2220-7619-GPO-1744> (2022).
7. Andreev, D. N., Maev, I. V. & Kucheryavy, Y. A. *Helicobacter pylori* resistance in the Russian Federation: A meta-analysis of studies over the past 10 years. *Ther. Arch.* **92**(11), 24–30. <https://doi.org/10.26442/00403660.2020.11.000795> (2020).
8. Marques, A. T., Vitor, J. M. B., Santos, A., Oleastro, M. & Vale, F. F. Trends in *Helicobacter pylori* resistance to clarithromycin: from phenotypic to genomic approaches. *Microb. Genom.* **6**(3), e000344. <https://doi.org/10.1099/mgen.0.000344> (2020).
9. Chen, J. *et al.* Application of next-generation sequencing to characterize novel mutations in clarithromycin-susceptible *Helicobacter pylori* strains with A2143G of 23S rRNA gene. *Ann. Clin. Microbiol. Antimicrob.* **17**(1), 10. <https://doi.org/10.1186/s12941-018-0259-8> (2018).
10. Saracino, I. M. *et al.* Next generation sequencing for the prediction of the antibiotic resistance in *Helicobacter pylori*: A literature review. *Antibiotics (Basel)* **10**(4), 437. <https://doi.org/10.3390/antibiotics10040437> (2021).
11. Binh, T. T. *et al.* Discovery of novel mutations for clarithromycin resistance in *Helicobacter pylori* by using next-generation sequencing. *J. Antimicrob. Chemother.* **69**(7), 1796–1803. <https://doi.org/10.1093/jac/dku050> (2014).
12. Hirata, K. *et al.* Contribution of efflux pumps to clarithromycin resistance in *Helicobacter pylori*. *J. Gastroenterol. Hepatol.* **25**(1), 75–79. <https://doi.org/10.1111/j.1440-1746.2009.06220.x> (2010).
13. Iwamoto, A. *et al.* Whole-genome sequencing of clarithromycin resistant *Helicobacter pylori* characterizes unidentified variants of multidrug resistant efflux pump genes. *Gut Pathog.* **6**, 27. <https://doi.org/10.1186/1757-4749-6-27> (2014).
14. Tang, X. *et al.* Re-assessment of the disk diffusion technique for routine antimicrobial susceptibility testing for *Helicobacter pylori*. *Helicobacter* **25**(4), e12703. <https://doi.org/10.1111/hel.12703> (2020).
15. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes de novo assembler. *Curr. Protoc. Bioinform.* **70**, e102. <https://doi.org/10.1002/cpbi.102> (2020).
16. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**(8), 1072–1075. <https://doi.org/10.1093/bioinformatics/btt086> (2013).
17. Antipov, D. *et al.* PlasmidSPAdes: Assembling plasmids from whole genome sequencing data. *Bioinformatics* **32**, 3380–3387 (2016).
18. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: A fast and accurate Illumina paired-end read merger. *Bioinformatics* **30**(5), 614–620. <https://doi.org/10.1093/bioinformatics/btt593> (2014).
19. Grüning, B. *et al.* Bioconda Team. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat. Methods* **15**(7), 475–476. <https://doi.org/10.1038/s41592-018-0046-7> (2018).
20. Seemann, T. *Snippy: Fast Bacterial Variant Calling from NGS Reads*. <https://github.com/tseemann/snippy> (2018).
21. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92. <https://doi.org/10.4161/fly.19695> (2012).
22. Okonechnikov, K., Golosova, O. & Fursov, M. UGENE team. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **28**(8), 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091> (2012).
23. UniProt Consortium. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, 506–515. <https://doi.org/10.1093/nar/gky1049> (2019).
24. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**(9), 1312–1313 (2014).
25. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**(43), 1686. <https://doi.org/10.21105/joss.01686> (2019).
26. Revelle, W. & Revelle, M. W. *Package 'Psych' the Comprehensive R Archive Network* (2015).
27. Liu, Y. *et al.* Antimicrobial resistance patterns and genetic elements associated with the antibiotic resistance of *Helicobacter pylori* strains from Shanghai. *Gut Pathog.* **14**(1), 14. <https://doi.org/10.1186/s13099-022-00488-y> (2022).
28. Looman, A. C. & van Knippenberg, P. H. Effects of GUG and AUG initiation codons on the expression of lacZ in *Escherichia coli*. *FEBS Lett.* **197**(1–2), 315–320. [https://doi.org/10.1016/0014-5793\(86\)80349-0](https://doi.org/10.1016/0014-5793(86)80349-0) (1986).
29. Maev, I. V. *et al.* Antibiotic resistance of *Helicobacter pylori* in the European part of the Russian Federation: First results. *Ther. Arch.* **92**(8), 24–28. <https://doi.org/10.26442/00403660.2020.08.000761> (2020).
30. Maev, I. V. *et al.* *Helicobacter pylori* resistance to clarithromycin in the Russian Federation. *Effect. Pharmacother.* **16**(30), 16–23. <https://doi.org/10.33978/2307-3586-2020-16-30-16-22> (2020).
31. Zhong, Z. *et al.* A retrospective study of the antibiotic-resistant phenotypes and genotypes of *Helicobacter pylori* strains in China. *Am. J. Cancer Res.* **11**(10), 5027–5037 (2021).
32. Mansour-Ghanaei, F., Poostizadeh, G., Joukar, F. & Siavoshi, F. Efficacy of disc diffusion and agar dilution methods in evaluating *Helicobacter pylori* susceptibility to antibiotics. *Middle East J. Dig. Dis.* **14**(2), 207–213. <https://doi.org/10.34172/mejdd.2022.274> (2022).
33. Lauener, F. N. *et al.* Genetic determinants and prediction of antibiotic resistance phenotypes in *Helicobacter pylori*. *J. Clin. Med.* **8**(1), 53. <https://doi.org/10.3390/jcm8010053> (2019).
34. Bińkowska, A., Biernat, M. M., Łaczmanski, Ł. & Gościński, G. Molecular patterns of resistance among *Helicobacter pylori* strains in south-western Poland. *Front. Microbiol.* **9**, 3154. <https://doi.org/10.3389/fmicb.2018.03154> (2018).
35. Fernández-Caso, B., Miqueleiz, A. & Alarcón, T. Whole genome sequencing for studying *Helicobacter pylori* antimicrobial resistance. *Antibiotics (Basel)* **12**(7), 1135. <https://doi.org/10.3390/antibiotics12071135> (2023).
36. Lin, Y., Shao, Y., Yan, J. & Ye, G. Antibiotic resistance in *Helicobacter pylori*: From potential biomolecular mechanisms to clinical practice. *J. Clin. Lab Anal.* **37**(7), e24885. <https://doi.org/10.1002/jcla.24885> (2023).
37. Thompson, L. J. *et al.* Gene expression profiling of *Helicobacter pylori* reveals a growth-phase-dependent switch in virulence gene expression. *Infect. Immun.* **71**(5), 2643–2655. <https://doi.org/10.1128/IAI.71.5.2643-2655.2003> (2003).
38. Godoy, A. P. *et al.* Differentially expressed genes in response to amoxicillin in *Helicobacter pylori* analyzed by RNA arbitrarily primed PCR. *FEMS Immunol. Med. Microbiol.* **50**(2), 226–230. <https://doi.org/10.1111/j.1574-695X.2006.00209.x> (2007).
39. Porcelli, I., Reuter, M., Pearson, B. M., Wilhelm, T. & van Vliet, A. H. Parallel evolution of genome structure and transcriptional landscape in the Epsilon proteobacteria. *BMC Genomics* **14**, 616. <https://doi.org/10.1186/1471-2164-14-616> (2013).
40. Revitt-Mills, S. A. & Robinson, A. Antibiotic-induced mutagenesis: Under the microscope. *Front. Microbiol.* **11**, 585175. <https://doi.org/10.3389/fmicb.2020.585175> (2020).
41. Gu, H. Role of flagella in the pathogenesis of *Helicobacter pylori*. *Curr. Microbiol.* **74**(7), 863–869. <https://doi.org/10.1007/s00284-017-1256-4> (2017).
42. Gibson, K. H. *et al.* Control of the flagellation pattern in *Helicobacter pylori* by FlhF and FlhG. *J. Bacteriol.* **205**(9), e0011023. <https://doi.org/10.1128/jb.00110-23> (2023).
43. Vianna, J. S. *et al.* The interplay between mutations in cagA, 23S rRNA, gyrA and drug resistance in *Helicobacter pylori*. *Rev. Inst. Med. Trop. Sao Paulo* **60**, e25. <https://doi.org/10.1590/s1678-9946201860025> (2018).
44. Oktem-Okullu, S. *et al.* Analysis of correlation between the seven important *Helicobacter pylori* virulence factors and drug resistance in patients with gastritis. *Gastroenterol. Res. Pract.* **2020**, 3956838. <https://doi.org/10.1155/2020/3956838> (2020).
45. Haddadi, M. H., Negahdari, B., Asadolahi, R. & Bazargani, A. *Helicobacter pylori* antibiotic resistance and correlation with cagA motifs and homB gene. *Postgrad. Med.* **132**(6), 512–520. <https://doi.org/10.1080/00325481.2020.1753406> (2020).

46. Brennan, D. E., Dowd, C., O'Morain, C., McNamara, D. & Smith, S. M. Can bacterial virulence factors predict antibiotic resistant *Helicobacter pylori* infection?. *World J. Gastroenterol.* **24**(9), 971–981. <https://doi.org/10.3748/wjg.v24.i9.971> (2018).

Acknowledgements

We are thankful to Dr. Igor Mokrousov for his valuable assistance and technical guidance during this work.

Author contributions

N. Gladyshev—bioinformatics, phylogenetics and statistical analysis; A. Svarval—material preparation, data collection, culture methods; D. Polev/A. Saitova—sequencing, bioinformatics; S. Egorova—supervision; D. Starkova—bioinformatics, study design and writing the first draft manuscript; All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-70977-4>.

Correspondence and requests for materials should be addressed to D.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024