Figure S1 Abundance of cagA, vacA, oipA, and iceA genes in different clinical cases: height of the columns represents the occurrence of each clinical case. The height of each colored partition represents the number of positive cases for each virulence gene.



Figure S2 PFGE fingerprint heatmap. It showed the genetic fingerprint of 22 samples in comparison to each other in the form of heatmaps. Samples are presented on the vertical line (Y-axis) and different band classes are presented on the horizontal line (X-axis). Figure 2a represents the rate of presence or absence of bands in every band class along the entire process in different isolates. Figure S2b represents the same results of figure S2a but on the level of band intensity.

Table S1 Helicobacter pylori strains from different geographical locations used to build the phylogenic tree

|  |  |  |  |
| --- | --- | --- | --- |
| Genome name | BioProject Accession | Geographic Location | Date of publication |
| Helicobacter pylori NCTC 11637 = CCUG 17874 = ATCC 43504 strain NCTC11637 | PRJEB6403 | Australia  | 2015-2018 |
| Helicobacter pylori strain BM013B | PRJNA245314 | Australia: Perth  | 2014 |
| Helicobacter pylori strain HP14039 | PRJNA374603 | Australia: Perth | 2018 |
| Helicobacter pylori strain ausabrJ05 | PRJNA245115 | Australia, Jigalong, | 2016 |
| Helicobacter pylori XZ274 | PRJNA79003 | China | 2012 |
| Helicobacter pylori strain GCT 27 | PRJNA604073 | Colombia: Valle del Cauca | 2020 |
| Helicobacter pylori strain 25-A-EK9 | PRJNA490474 | Germany: Magdeburg | 2019 |
| Helicobacter pylori strain 21-F-EK1 | PRJNA490474 | Germany: Magdeburg | 2019 |
| Helicobacter pylori strain 24-A-EK1 | PRJNA490474 | Germany: Magdeburg | 2019 |
| Helicobacter pylori strain 23-A-EK1 | PRJNA490474 | Germany: Magdeburg | 2019 |
| Helicobacter pylori strain L7 | PRJNA245115 |  India, Ladakh | 2016 |
| Helicobacter pylori oki154 | PRJNA215152 | Japan: Okinawa | 2014 |
| Helicobacter pylori strain TN2wt | PRJNA244276 | Japan: Oita | 2019 |
| Helicobacter pylori strain F78 | PRJDB4296 | Japan: Fukui | 2016 |
| Helicobacter pylori strain C-Mx-2010-5 | PRJNA203445 | Mexico: Mexico City | 2015 |
| Helicobacter pylori strain NCTC13345 | PRJEB6403 | Nigeria | 2014 |
| Helicobacter pylori Rif1 | PRJNA175544 | Russia: Mosco | 2012 |
| Helicobacter pylori HUP-B14 | PRJNA95277 | Spain | 2012 |
| Helicobacter pylori SouthAfrica20 | PRJNA209860 | South Africa | 2013 |
| Helicobacter pylori strain CC33C | PRJNA245115 | South Africa Cape Town | 2016 |
| Helicobacter pylori strain DU15 | PRJNA245115 | South Korea, Seoul, | 2016 |
| Helicobacter pylori 26695-1MET | PRJNA270699 | USA: Texas | 2015 |
| Helicobacter pylori J99 | PRJNA234 | USA | 2003 |
| Helicobacter pylori 26695 | PRJNA233 | United Kingdom | 2003 |

Table S2 list of resistance genes predicted in both samples M1 and M2.

|  |  |  |  |
| --- | --- | --- | --- |
| Resistance gene | Function | Antibiotic class | Classification |
| CmeDEF | Multidrug efflux system, membrane fusion component => CmeD, CmeE, CmeF | Penams, Peptide antibiotics | efflux pump conferring antibiotic resistance |
|  |  |  |  |
| MurA | UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7) | Fosfomycin | antibiotic target in susceptible species |
| gyrB | DNA gyrase subunit B (EC 5.99.1.3) | Quinolones, Fluoroquinolones | antibiotic target in susceptible species |
| 7a-HSDH-like | Enoyl-[acyl-carrier-protein] reductase (EC 1.3.1.-), 7-alpha-HSDH-like => refractory to triclosan | Triclosan | antibiotic target replacement protein |
| gidB | 16S rRNA (guanine (527)-N (7))-methyltransferase (EC 2.1.1.170) | Aminoglycosides | gene conferring resistance via absence |
| Ddl | D-alanine--D-alanine ligase (EC 6.3.2.4) | Cycloserine | antibiotic target in susceptible species |
| kasA | 3-oxoacyl-[acyl-carrier-protein] synthase, KASII (EC 2.3.1.179) | Isoniazid, Triclosan | antibiotic target in susceptible species |
| EF-Tu | Translation elongation factor Tu | Elfamycins | antibiotic target in susceptible species |
| gyrA | DNA gyrase subunit A (EC 5.99.1.3) | Quinolones, Fluoroquinolones | antibiotic target in susceptible species |
| Dxr | 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) | Fosmidomycin | antibiotic target in susceptible species |
| Alr | Alanine racemase (EC 5.1.1.1) | Cycloserine | antibiotic target in susceptible species |
| S12p | SSU ribosomal protein S12p (S23e) | Aminoglycosides | antibiotic target in susceptible species |
| Rho | Transcription termination factor Rho | Bicyclomycins | antibiotic target in susceptible species |
| Iso-tRNA | Isoleucyl-tRNA synthetase (EC 6.1.1.5) | Mupirocin | antibiotic target in susceptible species |
| hp1181 | efflux pump conferring antibiotic resistance | Multidrug efflux transporter | efflux pump conferring antibiotic resistance |
| hp1184 | Protein homolog model | Multi-antimicrobial extrusion protein | efflux pump conferring antibiotic resistance |
| PgsA | CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5) | Peptide antibiotics | protein altering cell wall charge conferring antibiotic resistance |
| inhA, fabI | Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) | Isoniazid, Ethionamide, Triclosan | antibiotic target in susceptible species |
| S10p | SSU ribosomal protein S10p (S20e) | Tetracyclines, Glycylcyclines | antibiotic target in susceptible species |
| EF-G | Translation elongation factor G | Fusidic acid | antibiotic target in susceptible species |

Table S3 List of resistance for antibiotics as detected from ResFinder and PointFinder tools for samples M1 and M2

|  |  |  |  |
| --- | --- | --- | --- |
| Antimicrobial | Class | WGS-predicted phenotype M1 | WGS-predicted phenotype M2 |
| Vancomycin | Glycopeptide | No resistance | No resistance |
| Mupirocin | pseudomonic acid | No resistance | No resistance |
| Tobramycin | Aminoglycoside | No resistance | No resistance |
| Virginiamycin m | streptogramin a | No resistance | No resistance |
| Isepamicin | Aminoglycoside | No resistance | No resistance |
| Virginiamycin s | streptogramin b | No resistance | No resistance |
| Hydrogen peroxide | Peroxides | No resistance | No resistance |
| Butirosin | Aminoglycoside | No resistance | No resistance |
| Ampicillin | beta-lactam | No resistance | No resistance |
| Astromicin | Aminoglycoside | No resistance | No resistance |
| Lividomycin | Aminoglycoside | No resistance | No resistance |
| Sulfamethoxazole | folate pathway antagonist | No resistance | No resistance |
| Temocillin | beta-lactam | No resistance | No resistance |
| Trimethoprim | folate pathway antagonist | No resistance | No resistance |
| Oleandomycin | macrolide | No resistance | No resistance |
| Unknown macrolide | macrolide | No resistance | No resistance |
| Florfenicol | phenicol | No resistance | No resistance |
| Fluoroquinolone | under\_development | No resistance | No resistance |
| Quinupristin | streptogramin b | No resistance | No resistance |
| Fosfomycin | fosfomycin | No resistance | No resistance |
| Cephalothin | beta-lactam | No resistance | No resistance |
| Hygromycin | aminoglycoside | No resistance | No resistance |
| Spectinomycin | aminocyclitol | No resistance | No resistance |
| Piperacillin+clavulanic acid | beta-lactam | No resistance | No resistance |
| Paromomycin | aminoglycoside | No resistance | No resistance |
| Clindamycin | lincosamide | No resistance | No resistance |
| Amoxicillin+clavulanic acid | beta-lactam | No resistance | No resistance |
| Gentamicin | aminoglycoside | No resistance | No resistance |
| Piperacillin+tazobactam | beta-lactam | No resistance | No resistance |
| Tiamulin | pleuromutilin | No resistance | No resistance |
| Ribostamycin | aminoglycoside | No resistance | No resistance |
| Erythromycin | macrolide | No resistance | No resistance |
| Kanamycin | aminoglycoside | No resistance | No resistance |
| Teicoplanin | glycopeptide | No resistance | No resistance |
| Amikacin | aminoglycoside | No resistance | No resistance |
| Tigecycline | tetracycline | No resistance | No resistance |
| Ticarcillin+clavulanic acid | beta-lactam | No resistance | No resistance |
| Cephalotin | beta-lactam | No resistance | No resistance |
| Cefoxitin | beta-lactam | No resistance | No resistance |
| Penicillin | beta-lactam | No resistance | No resistance |
| Neomycin | aminoglycoside | No resistance | No resistance |
| Dalfopristin | streptogramin a | No resistance | No resistance |
| Piperacillin | beta-lactam | No resistance | No resistance |
| Colistin | polymyxin | No resistance | No resistance |
| Telithromycin | macrolide | No resistance | No resistance |
| Amoxicillin | beta-lactam | No resistance | No resistance |
| Meropenem | beta-lactam | No resistance | No resistance |
| Ethidium bromide | quaternary ammonium compounds | No resistance | No resistance |
| Sisomicin | aminoglycoside | No resistance | No resistance |
| Unknown aminoglycoside | aminoglycoside | No resistance | No resistance |
| Cefepime | beta-lactam | No resistance | No resistance |
| Butiromycin | aminoglycoside | No resistance | No resistance |
| Doxycycline | tetracycline | No resistance | No resistance |
| Unknown fluoroquinolone | fluoroquinolone | No resistance | No resistance |
| Fusidic acid | steroid antibacterial | No resistance | No resistance |
| Ciprofloxacin | fluoroquinolone | No resistance | No resistance |
| Minocycline | tetracycline | No resistance | No resistance |
| Temperature | heat | No resistance | No resistance |
| Imipenem | beta-lactam | No resistance | No resistance |
| Arbekacin | aminoglycoside | No resistance | No resistance |
| Nalidixic acid | fluoroquinolone | No resistance | No resistance |
| Metronidazole | nitroimidazole | No resistance | No resistance |
| Cefixime | beta-lactam | No resistance | No resistance |
| Unknown phenicol | phenicol | No resistance | No resistance |
| Pristinamycin ia | streptogramin b | No resistance | No resistance |
| Formaldehyde | aldehydes | No resistance | No resistance |
| Tylosin | macrolide | No resistance | No resistance |
| Benzylkonium chloride | quaternary ammonium compounds | No resistance | No resistance |
| Cefotaxime+clavulanic acid | beta-lactam | No resistance | No resistance |
| Rifampicin | rifamycin | No resistance | No resistance |
| Ceftriaxone | beta-lactam | No resistance | No resistance |
| Ceftazidime | beta-lactam | No resistance | No resistance |
| Fortimicin | aminoglycoside | No resistance | No resistance |
| Carbomycin | macrolide | No resistance | No resistance |
| Ticarcillin | beta-lactam | No resistance | No resistance |
| Azithromycin | macrolide | No resistance | No resistance |
| Chlorhexidine | quaternary ammonium compounds | No resistance | No resistance |
| Kasugamycin | aminoglycoside | No resistance | No resistance |
| Chloramphenicol | phenicol | No resistance | No resistance |
| Cetylpyridinium chloride | quaternary ammonium compounds | No resistance | No resistance |
| Ampicillin+clavulanic acid | beta-lactam | No resistance | No resistance |
| Cefotaxime | beta-lactam | No resistance | No resistance |
| Unknown tetracycline | tetracycline | No resistance | No resistance |
| Quinupristin+dalfopristin | streptogramin a | No resistance | No resistance |
| Ceftazidime+avibactam | beta-lactam | No resistance | No resistance |
| Apramycin | aminoglycoside | No resistance | No resistance |
| Spiramycin | macrolide | No resistance | No resistance |
| Unknown rifamycin | rifamycin | No resistance | No resistance |
| Dibekacin | aminoglycoside | No resistance | No resistance |
| Ertapenem | beta-lactam | No resistance | No resistance |
| Tetracycline | tetracycline | No resistance | No resistance |
| Linezolid | oxazolidinone | No resistance | No resistance |
| Netilmicin | aminoglycoside | No resistance | No resistance |
| Lincomycin | lincosamide | No resistance | No resistance |
| Aztreonam | beta-lactam | No resistance | No resistance |
| Unknown beta-lactam | beta-lactam | No resistance | No resistance |
| Pristinamycin iia | streptogramin a | No resistance | No resistance |
| Streptomycin | aminoglycoside | No resistance | No resistance |

Table S4 List of virulence genes annotated in sample M1

|  |  |  |
| --- | --- | --- |
| Gene | Product | Classification |
| tlpB | Methyl-accepting chemotaxis protein |  |
| HP1092 | Flagellar basal-body rod protein FlgF |  |
| cheV3 | methyl-accepting chemotaxis protein (MCP) |  |
| HP0017 | ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex, VirB4 |  |
| hopZ | Outer membrane protein | Adherence |
| flgD | Flagellar basal-body rod modification protein FlgD |  |
| gluE | UDP-glucose 4-epimerase (EC 5.1.3.2) | Endotoxin, Low toxicity, Adherence |
| pseC | UDP-4-amino-4,6-dideoxy-N-acetyl-beta-L-altrosamine transaminase (EC 2.6.1.92) |  |
| HP0758 | Histidine permease YuiF |  |
| cagH | cag island protein | Secretion system, Type IV secretion system, Proinflammatory effect |
| fliY | Flagellar motor switch protein FliN |  |
| HP0254 | Putative outer membrane protein |  |
| flgE\_1 | Paralog of flagellar hook protein FlgE | Motility |
| HP0245 | hypothetical protein |  |
| flaA | Flagellin |  |
| cagG | cag island protein | Secretion system, Type IV secretion system, Proinflammatory effect |
| fliD | Flagellar cap protein FliD |  |
| HP0392 | Signal transduction histidine kinase CheA |  |
| virB11 | ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex, VirB11 | Secretion system, Type IV secretion system,Proinflammatory effect |
| cheA | Signal transduction histidine kinase CheA |  |
| futA | Alpha-(1,3/1,4)-L-fucosyltransferase | Molecular mimicry |
| HP1206 | Multi-drug resistance efflux ABC transporter, permease/ATP-binding protein HetA |  |
| flgE | Paralog of flagellar hook protein FlgE |  |
| flgR | Signal-transduction regulatory protein FlgR |  |
| flaA | Flagellin | Motility |
| ureH | Urease accessory protein UreD | Enzyme, Acid resistance, Colonization |
| fliH | Flagellar assembly protein FliH |  |
| horB | Putative outer membrane protein |  |
| hopH | Outer membrane protein |  |
| HP0232 | secreted protein involved in flagellar motility |  |
| flaB | Flagellin | Motility |
| pflA | Paralysed flagella protein PflA |  |
| fliI | Flagellum-specific ATP synthase FliI |  |
| fliD | Flagellar cap protein FliD |  |
| cagU | cag pathogenicity island protein (cag11) | Secretion system, Type IV secretion system, Proinflammatory effect |
| ureF | Urease accessory protein UreF | Enzyme, Acid resistance, Colonization |
| fliG | Flagellar motor switch protein FliG |  |
| HP0302 | Dipeptide ABC transporter, ATP-binding protein DppF (TC 3.A.1.5.2) |  |
| fliS | Flagellar biosynthesis protein FliS |  |
| flgS | Flagellar sensory histidine kinase FlgS |  |
| alpA | Outer membrane protein/porin |  |
| HP0071 | Urea channel UreI |  |
| cds6 | Putative periplasmic protein |  |
| HP1082 | Efflux ABC transporter, permease/ATP-binding protein |  |
| HP0788 | Uncharacterized outer membrane protein HP0788 |  |
| flgI | Flagellar P-ring protein FlgI |  |
| flhB | Flagellar biosynthesis protein FlhB |  |
| cagQ | cag pathogenicity island protein (cag14) | Secretion system,Type IV secretion system,Proinflammatory effect |
| hemC | Porphobilinogen deaminase (EC 2.5.1.61) |  |
| neuA/flmD | Pseudaminic acid cytidylyltransferase (EC 2.7.7.81) / UDP-2,4-diacetamido-2,4,6-trideoxy-beta-L-altropyranose hydrolase (EC 3.6.1.57) | Molecular mimicry |
| rfaC | Lipopolysaccharide core heptosyltransferase I | Endotoxin,Low toxicity,Adherence |
| babB/hopT | outer membrane protein – adhesin | Adherence |
| pdxA | 4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262) |  |
| wbpB | UDP-2-acetamido-2-deoxy-D-glucuronic acid dehydrogenase (NAD+) (EC 1.1.1.335) | Endotoxin,Low toxicity,Adherence |
| fliL | Flagellar basal body-associated protein FliL |  |
| HP0714 | RNA polymerase sigma-54 factor RpoN |  |
| motA | Flagellar motor rotation protein MotA |  |
| virB2/cagC | cag pathogenicity island protein C | Secretion system,Type IV secretion system,Proinflammatory effect |
| HP0350 | hypothetical protein |  |
| fliA | RNA polymerase sigma factor for flagellar operon |  |
| cheY | Chemotaxis regulator - transmits chemoreceptor signals to flagellar motor components CheY |  |
| ureG | Urease accessory protein UreG | Enzyme,Acid resistance,Colonization |
| flgL | Flagellar hook-associated protein FlgL |  |
| flgG | Flagellar basal-body rod protein FlgF |  |
| cagN | cag pathogenicity island protein (cag17) | Secretion system,Type IV secretion system,Proinflammatory effect |
| surE | 5'-nucleotidase SurE (EC 3.1.3.5) |  |
| fliN | Flagellar motor switch protein FliN |  |
| virB4/cagE | CAG pathogenicity island protein 23 (Protein picB) | Secretion system,Type IV secretion system,Proinflammatory effect |
| HP0973 | hypothetical protein |  |
| hopH | Outer membrane protein | Proinflammatory effect |
| fliF | Flagellar M-ring protein FliF |  |
| virD4 | Coupling protein VirD4, ATPase required for T-DNA transfer | Secretion system,Type IV secretion system,Proinflammatory effect |
| virB10/cagC | cag pathogenicity island protein (cag7) | Secretion system,Type IV secretion system,Proinflammatory effect |
| HP0486 | Uncharacterized outer membrane protein HP0486 |  |
| sabA/hopP | Outer membrane protein | Adherence |
| tlyA | RNA binding methyltransferase FtsJ like |  |
| flhF | Flagellar biosynthesis protein FlhF |  |
| rfbD | GDP-mannose 4,6-dehydratase (EC 4.2.1.47) | Endotoxin,Low toxicity,Adherence |
| virB7/cagT | cag pathogenicity island protein (cag12) | Secretion system,Type IV secretion system,Proinflammatory effect |
| HP0397 | D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) |  |
| cagE | CAG pathogenicity island protein 23 (Protein picB) |  |
| HP0067 | Urease accessory protein UreD |  |
| flgH | Flagellar L-ring protein FlgH |  |
| cheV1 | Chemotaxis protein CheV (EC 2.7.3.-) |  |
| flgC | Flagellar basal-body rod protein FlgC |  |
| virB9/cagX | cag pathogenicity island protein (cag8) | Secretion system,Type IV secretion system,Proinflammatory effect |
| HP0256 | hypothetical protein |  |
| HP1486 | Efflux ABC transporter, permease protein HP\_1486 |  |
| HP0055 | Proline/sodium symporter PutP (TC 2.A.21.2.1) @ Propionate/sodium symporter |  |
| fliF | Flagellar M-ring protein FliF |  |
| virB5/cagL | cag pathogenicity island protein (cag18) | Secretion system,Type IV secretion system,Proinflammatory effect |
| fliE | Flagellar hook-basal body complex protein FliE |  |
| fliS | Flagellar biosynthesis protein FliS |  |
| ureI | Urea channel UreI | Enzyme,Acid resistance,Colonization |
| virB6/cagC | cag pathogenicity island protein (cag9) | Secretion system,Type IV secretion system,Proinflammatory effect |
| ureG | Urease accessory protein UreG |  |
| virB8/cagV | cag pathogenicity island protein (cag10) / Inner membrane protein forms channel for type IV secretion of T-DNA complex, VirB8 | Secretion system,Type IV secretion system,Proinflammatory effect |
| tlpA | methyl-accepting chemotaxis protein (tlpA) |  |
| HP0797 | flagellar sheath adhesin hpaA |  |
| pdxJ | Pyridoxine 5'-phosphate synthase (EC 2.6.99.2) |  |
| vacA | vacuolating cytotoxin | Toxin,Pore-forming,A-B type |
| flaG | Flagellar protein FlaG |  |
| hpaA2 | putative neuraminyllactose-binding hemagglutinin HpaA |  |
| flgG | Flagellar basal-body rod protein FlgG |  |
| napA | DNA protection during starvation protein | Proinflammatory effect |
| flgC | Flagellar basal-body rod protein FlgC |  |
| flgB | Flagellar basal-body rod protein FlgB |  |
| gluP | Homolog of fucose/glucose/galactose permeases | Endotoxin,Low toxicity,Adherence |
| flgD | Flagellar basal-body rod modification protein FlgD |  |
| homB | Putative outer membrane protein |  |
| fliP | Flagellar biosynthesis protein FliP |  |
| ureA | Urease gamma subunit (EC 3.5.1.5) / Urease beta subunit (EC 3.5.1.5) | Enzyme,Acid resistance,Colonization |
| HP1506 | Sodium/glutamate symporter |  |
| cagF | cag pathogenicity island protein | Secretion system,Type IV secretion system,Proinflammatory effect |
| HP0360 | UDP-glucose 4-epimerase (EC 5.1.3.2) |  |
| ureC | Urease alpha subunit (EC 3.5.1.5) |  |
| virB1 | cag island protein | Secretion system,Type IV secretion system,Proinflammatory effect |
| flhA | Flagellar biosynthesis protein FlhA |  |
| cheV2 | Chemotaxis protein CheV (EC 2.7.3.-) |  |
| sabA | Outer membrane protein |  |
| flgH | Flagellar L-ring protein FlgH |  |
| cagS | cag pathogenicity island protein (cag13) | Secretion system,Type IV secretion system,Proinflammatory effect |
| HP1091 | Alpha-ketoglutarate permease |  |
| fliR | Flagellar biosynthesis protein FliR |  |
| flgL | Flagellar hook-associated protein FlgL |  |
| HP0288 | Uncharacterized protein HP0288 |  |
| fliM | Flagellar motor switch protein FliM |  |
| rfaJ | Lipopolysaccharide biosynthesis protein | Endotoxin,Low toxicity,Adherence |
| motB | Flagellar motor rotation protein MotB |  |
| cagD | cag island protein | Secretion system,Type IV secretion system,Proinflammatory effect |
| wbcJ | GDP-L-fucose synthetase (EC 1.1.1.271) | Endotoxin,Low toxicity,Adherence |
| HP0859 | ADP-L-glycero-D-manno-heptose-6-epimerase (EC 5.1.3.20) |  |
| cagZ | cag pathogenicity island protein Z | Secretion system,Type IV secretion system,Proinflammatory effect |
| kdtB | Phosphopantetheine adenylyltransferase (EC 2.7.7.3) | Endotoxin,Low toxicity,Adherence |
| sabB/hopO | Outer membrane protein |  |
| cagM | cag pathogenicity island protein (cag16) | Secretion system,Type IV secretion system,Proinflammatory effect |
| ureB | Urease alpha subunit (EC 3.5.1.5) | Enzyme,Acid resistance,Colonization |
| rfaJ | putative lipopolysaccharide biosynthesis protein | Endotoxin,Low toxicity,Adherence |
| tlpC | Methyl-accepting chemotaxis signal transduction protein |  |
| fliI | Flagellum-specific ATP synthase FliI |  |
| cheW | Positive regulator of CheA protein activity (CheW) |  |
| futB | Alpha-(1,3/1,4)-L-fucosyltransferase | Molecular mimicry |
| futC | putative ALPHA(1,2)FUCOSYLTRANSFERASE |  |
| fliQ | Flagellar biosynthesis protein FliQ |  |
| ureE | Urease accessory protein UreE | Enzyme,Acid resistance,Colonization |
| flgA | Flagellar basal-body P-ring formation protein FlgA |  |
| flgK | Flagellar hook-associated protein FlgK |  |
| HP1421 | ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex, VirB11 |  |
| HP0169 | Uncharacterized protease HP\_0169 |  |
| fliM | Flagellar motor switch protein FliM |  |
| pseB | UDP-N-acetylglucosamine 4,6-dehydratase (inverting) (EC 4.2.1.115) |  |
| HP0366 | UDP-4-amino-4,6-dideoxy-N-acetyl-beta-L-altrosamine transaminase (EC 2.6.1.92) |  |
| HP0073 | Urease gamma subunit (EC 3.5.1.5) / Urease beta subunit (EC 3.5.1.5) |  |
| fliQ | Flagellar biosynthesis protein FliQ |  |
| alpB | Outer membrane protein/porin |  |
| flaB | Flagellin |  |
| HP1525 | Outer membrane protein |  |
| cagA | cag island protein, CYTOTOXICITY ASSOCIATED IMMUNODOMINANT ANTIGEN | Type IV secretory protein |
| flhB2 | Uncharacterized homolog of the cytoplasmic domain of flagellar protein FhlB |  |
| lpxB | Lipid-A-disaccharide synthase (EC 2.4.1.182) | Endotoxin,Low toxicity,Adherence |
| HP0393 | Chemotaxis protein CheV (EC 2.7.3.-) |  |
| cagI | cag pathogenicity island protein (cag19) | Secretion system,Type IV secretion system,Proinflammatory effect |
| ylxH | Flagellar synthesis regulator FleN |  |
| HP0289 | toxin-like outer membrane protein |  |
| cag1 | cag pathogenicity island protein (cag1) | Secretion system,Type IV secretion system,Proinflammatory effect |
| flgE | Flagellar hook protein FlgE |  |
| rfbM | Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) / Mannose-6-phosphate isomerase (EC 5.3.1.8) | Endotoxin,Low toxicity,Adherence |
| napA | DNA protection during starvation protein |  |
| cag3 | cag pathogenicity island protein (cag3) | Secretion system,Type IV secretion system,Proinflammatory effect |
| flgM | FlgM protein |  |

Table S5 List of virulence genes annotated in sample M2

|  |  |  |
| --- | --- | --- |
| Gene | Product | Classification |
| HP0288 | Uncharacterized protein HP0288 |  |
| lpxB | Lipid-A-disaccharide synthase (EC 2.4.1.182) | Endotoxin,Low toxicity,Adherence |
| HP1082 | Efflux ABC transporter, permease/ATP-binding protein |  |
| hemC | Porphobilinogen deaminase (EC 2.5.1.61) |  |
| cheW | Positive regulator of CheA protein activity (CheW) |  |
| fliI | Flagellum-specific ATP synthase FliI |  |
| flgE | Paralog of flagellar hook protein FlgE |  |
| ylxH | Flagellar synthesis regulator FleN |  |
| flgE | Paralog of flagellar hook protein FlgE |  |
| sabA/hopP | Putative outer membrane protein | Adherence |
| hopZ | Putative outer membrane protein | Adherence |
| flhB2 | Uncharacterized homolog of the cytoplasmic domain of flagellar protein FhlB |  |
| ylxH | Flagellar synthesis regulator FleN |  |
| neuA/flmD | UDP-2,4-diacetamido-2,4,6-trideoxy-beta-L-altropyranose hydrolase (EC 3.6.1.57) | Molecular mimicry |
| flgK | Flagellar hook-associated protein FlgK |  |
| fliN | Flagellar motor switch protein FliN |  |
| HP0256 | hypothetical protein |  |
| kdtB | Phosphopantetheine adenylyltransferase (EC 2.7.7.3) | Endotoxin,Low toxicity,Adherence |
| fliS | Flagellar biosynthesis protein FliS |  |
| cheV3 | methyl-accepting chemotaxis protein (MCP) |  |
| HP0071 | Urea channel UreI |  |
| rfaJ | putative lipopolysaccharide biosynthesis protein | Endotoxin,Low toxicity,Adherence |
| HP1092 | Flagellar basal-body rod protein FlgF |  |
| flhA | Flagellar biosynthesis protein FlhA |  |
| ureH | Urease accessory protein UreD | Enzyme,Acid resistance,Colonization |
| flgE\_1 | Paralog of flagellar hook protein FlgE | Motility |
| cheA | Signal transduction histidine kinase CheA |  |
| flgI | Flagellar P-ring protein FlgI |  |
| tlpB | Methyl-accepting chemotaxis protein |  |
| HP1525 | Outer membrane protein |  |
| ureG | Urease accessory protein UreG | Enzyme,Acid resistance,Colonization |
| babA | outer membrane protein – adhesin |  |
| HP0397 | D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) |  |
| vacA | vacuolating cytotoxin |  |
| vacA | vacuolating cytotoxin | Toxin,Pore-forming,A-B type |
| ureG | Urease accessory protein UreG |  |
| fliI | Flagellum-specific ATP synthase FliI |  |
| ureC | Urease alpha subunit (EC 3.5.1.5) |  |
| pdxA | 4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262) |  |
| fliY | Flagellar motor switch protein FliN |  |
| cheV2 | Chemotaxis protein CheV (EC 2.7.3.-) |  |
| HP1506 | Sodium/glutamate symporter |  |
| pseC | UDP-4-amino-4,6-dideoxy-N-acetyl-beta-L-altrosamine transaminase (EC 2.6.1.92) |  |
| napA | DNA protection during starvation protein | Proinflammatory effect |
| fliE | Flagellar hook-basal body complex protein FliE |  |
| fliL | Flagellar basal body-associated protein FliL |  |
| HP0392 | Signal transduction histidine kinase CheA |  |
| ureB | Urease alpha subunit (EC 3.5.1.5) | Enzyme,Acid resistance,Colonization |
| futB | Alpha-(1,3/1,4)-L-fucosyltransferase | Molecular mimicry |
| babB/hopT | outer membrane protein – adhesin | Adherence |
| HP0393 | Chemotaxis protein CheV (EC 2.7.3.-) |  |
| flaG | Flagellar protein FlaG |  |
| hemC | Porphobilinogen deaminase (EC 2.5.1.61) |  |
| flgM | FlgM protein |  |
| fliI | Flagellum-specific ATP synthase FliI |  |
| horB | Putative outer membrane protein |  |
| fliP | Flagellar biosynthesis protein FliP |  |
| surE | 5'-nucleotidase SurE (EC 3.1.3.5) |  |
| rfbD | GDP-mannose 4,6-dehydratase (EC 4.2.1.47) | Endotoxin,Low toxicity,Adherence |
| ureE | Urease accessory protein UreE | Enzyme,Acid resistance,Colonization |
| HP1486 | Efflux ABC transporter, permease protein HP\_1486 |  |
| hopZ | Putative outer membrane protein |  |
| cds6 | Putative periplasmic protein |  |
| tlpC | Methyl-accepting chemotaxis signal transduction protein |  |
| pdxJ | Pyridoxine 5'-phosphate synthase (EC 2.6.99.2) |  |
| motA | Flagellar motor rotation protein MotA |  |
| HP0067 | Urease accessory protein UreD |  |
| ureI | Urea channel UreI | Enzyme,Acid resistance,Colonization |
| flhF | Flagellar biosynthesis protein FlhF |  |
| cheY | Chemotaxis regulator - transmits chemoreceptor signals to flagellar motor components CheY |  |
| napA | DNA protection during starvation protein |  |
| fliS | Flagellar biosynthesis protein FliS |  |
| homB | Putative outer membrane protein |  |
| flgE\_1 | Paralog of flagellar hook protein FlgE | Motility |
| HP0254 | Putative outer membrane protein |  |
| sabA | Putative outer membrane protein |  |
| flgC | Flagellar basal-body rod protein FlgC |  |
| fliI | Flagellum-specific ATP synthase FliI |  |
| HP0797 | flagellar sheath adhesin hpaA |  |
| flgG | Flagellar basal-body rod protein FlgF |  |
| ylxH | Flagellar synthesis regulator FleN |  |
| Tufa | Translation elongation factor Tu |  |
| futC | putative ALPHA(1,2)FUCOSYLTRANSFERASE |  |
| ureF | Urease accessory protein UreF | Enzyme,Acid resistance,Colonization |
| flgC | Flagellar basal-body rod protein FlgC |  |
| pflA | Paralysed flagella protein PflA |  |
| HP0169 | Uncharacterized protease HP\_0169 |  |
| flgG | Flagellar basal-body rod protein FlgG |  |
| HP1421 | ATPase required for both assembly of type IV secretion complex and secretion of T-DNA complex, VirB11 |  |
| HP0366 | UDP-4-amino-4,6-dideoxy-N-acetyl-beta-L-altrosamine transaminase (EC 2.6.1.92) |  |

1. Dabiri H, Maleknejad P Fau - Yamaoka Y, Yamaoka Y Fau - Feizabadi MM, et al. Distribution of Helicobacter pylori cagA, cagE, oipA and vacA in different major ethnic groups in Tehran, Iran. 2009(1440-1746 (Electronic))

2. Abu-Taleb AMF, Abdelattef RS, Abdel-Hady AA, et al. Prevalence of Helicobacter pylori cagA and iceA Genes and Their Association with Gastrointestinal Diseases. Int J Microbiol 2018;**2018**:4809093 doi: 10.1155/2018/4809093[published Online First: Epub Date]|.

3. Atherton JC, Cao P, Peek RM, Jr., Tummuru MK, Blaser MJ, Cover TL. Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori. Association of specific vacA types with cytotoxin production and peptic ulceration. The Journal of biological chemistry 1995;**270**(30):17771-7 doi: 10.1074/jbc.270.30.17771[published Online First: Epub Date]|.