

# In silico Analysis of Virulence, Resistance Genes and Phylogeny of *Helicobacter pylori* Strains from Different Continents

Farklı Kıtalarındaki *Helicobacter pylori* Suşlarının Virülans, Direnç Genleri ve Filogenisinin *in silico* Analizi

Mehmet Demirci<sup>1</sup> , Özge Ünlü<sup>2</sup> , Bekir S. Kocazeybek<sup>3</sup> 

<sup>1</sup>Department of Medical Microbiology, Faculty of Medicine, Kırklareli University, Kırklareli, Turkey

<sup>2</sup>Department of Medical Microbiology, Faculty of Medicine, Beykent University, Istanbul, Turkey

<sup>3</sup>Department of Medical Microbiology, Cerrahpaşa School of Medicine, Istanbul University-Cerrahpaşa, Istanbul, Turkey

ORCID ID: M.D. 0000-0001-9670-2426; Ö.Ü. 0000-0002-5411-5925; B.S.K. 0000-0003-1072-3846

**Cite this article as:** Demirci M, Ünlü Ö, Kocazeybek BS. *In silico* analysis of virulence, resistance genes and phylogeny of *Helicobacter pylori* strains from different continents. Experimed 2021; 11(3): 170-8.

## ABSTRACT

**Objective:** *Helicobacter pylori* (*H. pylori*) is a bacterium that infects the gastric mucosa of 50% of the world population. It is known that different regional treatment practices used against the infections of *H. pylori* affect both the expression of virulence and antimicrobial resistance genes, giving the bacteria geographic differentiation. The aim of this study was to perform *in silico* analysis of virulence, resistance genes and phylogeny of *H. pylori* strains obtained from people living in different continents.

**Material and Method:** Complete gene sequences of 18 *H. pylori* strains from six continents were downloaded from the National Center for Biotechnology Information (NCBI) database. The phylogeny of the strains, resistance and virulence genes were analyzed by CSI phylogeny, CARD and VFAnalyzer, respectively.

**Results:** African strains were the most distant identity to European strains. A2147G single nucleotide polymorphism associated with clarithromycin resistance was detected in South American and Asian origin. It was determined that strains were differentiated by a total of 95 related virulence genes under eight headings. The *cagA*, *cagE*, *cagL* and *vacA* genes were found in all strains in Asia.

**Conclusion:** In conclusion, our study demonstrated that *H. pylori* strains, whose data were collected in different continents, differ from each other in terms of similarities and there is a serious difference especially in terms of virulence genes.

**Keywords:** *Helicobacter pylori*, virulence genes, *in silico* analysis, geographic phylogeny

## ÖZ

**Amaç:** *Helicobacter pylori* (*H. pylori*), tüm dünya popülasyonunun %50'sinin mide mukozasını enfekte eden bir bakteridir ve bölgesel farklı tedavi uygulamaları, hem virülans genleri, hemde antimikrobiyal direnç genlerini etkileyerek, bakteriyi coğrafik olarak farklılaşma kazandırdığı bilinmektedir. Çalışmamızda, dünyanın farklı kıtalarında yaşayan insanlardan elde edilen *H. pylori* kökenlerinin filogeni, virülans ve antimikrobiyal direnç genleri açısından *in silico* analizinin yapılması amaçlanmıştır.

**Gereç ve Yöntem:** Altı kıtadan, toplam 18 *H. pylori* kökenine ait tüm genom dizileri NCBI veritabanından indirilerek çalışmamıza dahil edildi. Kökenlerin evrimsel yakınlıkları, direnç gen belirteçleri ve virülans genleri, sırasıyla CSI filogeni, CARD ve VFAnalyzer online yazılımları ile gerçekleştirildi.

**Bulgular:** Avrupa kökenine göre en uzak benzerlik Afrika kökenliydi. Klaritromisin direnci ile ilişkili A2147G tek nokta polimorfizmi Güney Amerika ve Asya kökeninde saptandı. Suşların 8 başlık altında toplam 95 ilişkili virülans geni taşıdığı belirlendi. Asya'daki tüm suşlarda *cagA*, *cagE*, *cagL* ve *vacA* genleri bulundu.

**Sonuç:** Sonuç olarak, çalışmamızın verileri, farklı kıtalarda tespit edilen *H. pylori* kökenlerinin birbirinden farklılıklar gösterdiği ve özellikle virülans genleri açısından ciddi farklılık içerdiğini ortaya koymuştur.

**Anahtar Kelimeler:** *Helicobacter pylori*, virülans genleri, *in silico* analiz, coğrafik filogeni

**Corresponding Author/Sorumlu Yazar:** Mehmet Demirci **E-mail:** demircimehmet@hotmail.com

**Submitted/Başvuru:** 12.07.2021 **Revision Requested/Revizyon Talebi:** 22.08.2021

**Last Revision Received/Son Revizyon:** 02.09.2021 **Accepted/Kabul:** 04.09.2021



Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

## INTRODUCTION

Until Warren and Marshall discovered and identified *Helicobacter pylori* (*H. pylori*) infection in the gastric mucosa in 1983, it was believed that the stomach was sterile due to its highly acidic content (1). However, it is now known that *H. pylori*, which is a spiral-shaped, gram-negative, microaerophilic bacterium, infects the gastric mucosa of 50% of the world population and may lead to gastritis, ulcer, or gastric cancer (2). *H. pylori* infections, which have the ability to colonize the human gastric mucosa, are usually acquired in the early stages of life and can survive for a lifetime (3). The manifestation of *H. pylori* infections in various clinical presentations is due to bacterial virulence factors (e.g. *cagA*, *vacA*, *babA*), host genetic characteristics (e.g. age, immune system) and environmental factors (e.g. nutrition, geographical region, living and socio-economic conditions). It has been shown that *H. pylori* may fail to colonize the gastric mucosa by silencing genes that express virulence factors such as flagella, urease production or chemotaxis (4,5). It is understood that different regional treatment algorithms used against *H. pylori* give the bacteria geographic differentiation by affecting both virulence genes and antimicrobial resistance genes (3,4,6,7). The high throughput data obtained by molecular-based systems such as the next generation sequencing systems that have been developed in recent years are import-

ant in terms of revealing this geographic differentiation. In addition, by sequencing the whole genome of microbial agents using these systems, virulence factors and resistance genes can be detected quickly and accurately (8). Thus, in our study, we aimed to perform *in silico* analysis of virulence, resistance genes and phylogeny of *H. pylori* strains obtained from people living in different continents.

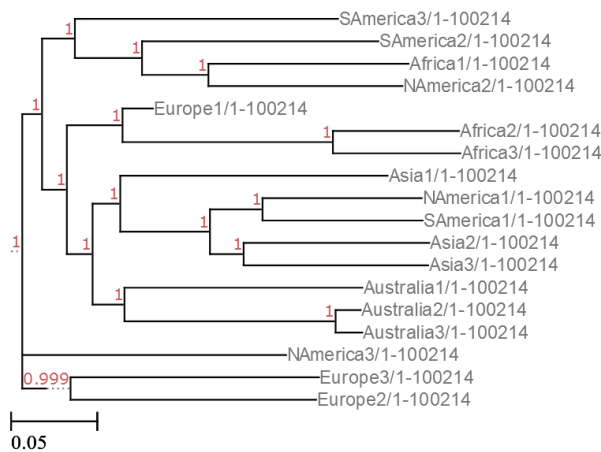
## MATERIAL AND METHOD

Genomic data belonging to a total of 18 *H. pylori* strains from six different continents whose gene sequences uploaded to open databases were downloaded from the National Center for Biotechnology Information (NCBI, <https://www.ncbi.nlm.nih.gov/>) database and included in our study. NCBI accession numbers and information of these strains obtained from patients after endoscopy in Africa, South America, North America, Asia, Europe, and Australia, three from each continent, are presented in Table 1. The strains included in this study, constitute the most selected reference strains in many publications. In addition, although the NCBI database contains data on more than 2000 strains, the whole genome sequencing of all these strains has not yet been completed, but the whole genome sequencing of all the strains we have included in our study has been completed.

**Table 1.** NCBI accession numbers and information of all strains included in this study.

| Continent          | Country      | Name of the Strain | NCBI Accession number |
|--------------------|--------------|--------------------|-----------------------|
| <b>Africa1</b>     | Gambia       | Gambia94/24        | NC_017371.1           |
| <b>Africa2</b>     | South Africa | South Africa7      | NC_017361.1           |
| <b>Africa3</b>     | South Africa | South Africa20     | CP006691.1            |
| <b>S.America1</b>  | Venezuela    | v225d              | NC_017355.1           |
| <b>S. America2</b> | Peru         | PeCan18            | NC_017742.1           |
| <b>S. America3</b> | El Salvador  | ELS37              | NC_017063.1           |
| <b>N. America1</b> | Canada       | Aklavik117         | NC_019560.1           |
| <b>N. America2</b> | USA          | J99                | NZ_CP011330.1         |
| <b>N. America3</b> | Mexico       | 29CaP              | NZ_CP012907.1         |
| <b>Asia1</b>       | India        | India7             | NC_017372.1           |
| <b>Asia2</b>       | Taiwan       | ML3                | NZ_AP014712.1         |
| <b>Asia3</b>       | China        | XZ274              | NC_017926.1           |
| <b>Europe1</b>     | England      | ATCC 26695         | NC_000915.1           |
| <b>Europe2</b>     | France       | B38                | NC_012973.1           |
| <b>Europe3</b>     | Germany      | P12                | NC_011498.1           |
| <b>Australia1</b>  | Australia    | ATCC 43504         | NZ_LS483488.1         |
| <b>Australia2</b>  | Australia    | BM012B             | NZ_CP007605.1         |
| <b>Australia3</b>  | Australia    | BM012S             | NC_022911.1           |

S. America: South America, N. America: North America



**Figure 1.** Phylogeny analysis of genomic data of the strains included in the study. S. America: South America, N. America: North America

The evolutionary relatedness of these strains with each other was revealed with the CSI phylogeny software (<https://www.genomicepidemiology.org/>) (9). The presence of antimicrobial resistance markers in the strains was detected by Comprehensive Antibiotic Resistance Database (CARD - <https://card.mcmaster.ca/home>) online software (10). Comparative genomic analysis of virulence genes was performed with the VF analyzer software (<http://www.mgc.ac.cn/VFs/>) (11).

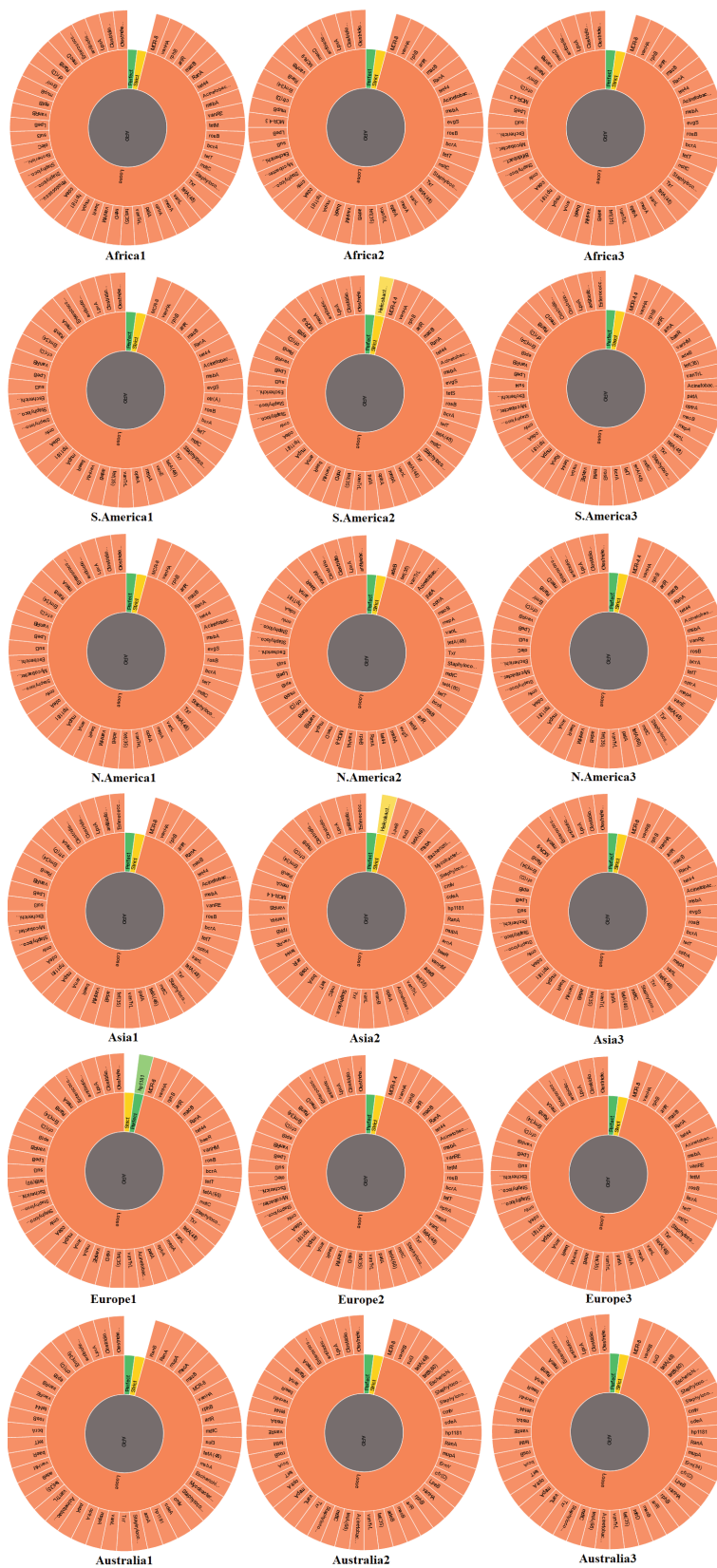
## RESULTS

When the identities of 18 *H. pylori* strains from six continents in our study were evaluated *in silico* phylogenetically, the most distant identity to the European1 strain was the strain of Africa2, North America1 and Africa3, and the identity rates were 86.5%, 86.86% and 86.90%, respectively. The strains with the closest identity to the European1 strain were Europe3 with 94.04% and Australia3 with 92.74% identity rates (Table 2).

**Table 2:** Identity rates of all strains examined in this study and the distribution of antimicrobials affected according to resistance gene analysis.

| Continent   | Country      | Strain name    | NCBI Accession number | Identity rate | Affected antimicrobials              | Resistance gene and SNP |
|-------------|--------------|----------------|-----------------------|---------------|--------------------------------------|-------------------------|
| Africa1     | Gambia       | Gambia94/24    | NC_017371.1           | 92.26%        |                                      |                         |
| Africa2     | South Africa | South Africa7  | NC_017361.1           | 86.50%        |                                      |                         |
| Africa3     | South Africa | South Africa20 | CP006691.1            | 86.90%        |                                      |                         |
| S. America1 | Venezuela    | v225d          | NC_017355.1           | 89.61%        |                                      |                         |
| S. America2 | Peru         | PeCan18        | NC_017742.1           | 91.51%        | Clarithromycin                       | 23S rRNA -A2147G        |
| S. America3 | El Salvador  | ELS37          | NC_017063.1           | 91.48%        |                                      |                         |
| N. America1 | Canada       | Aklavik117     | NC_019560.1           | 86.86%        |                                      |                         |
| N. America2 | USA          | J99            | NZ_CP011330.1         | 90.92%        |                                      |                         |
| N. America3 | Mexico       | 29CaP          | NZ_CP012907.1         | 90.51%        |                                      |                         |
| Asia1       | India        | India7         | NC_017372.1           | 91.18%        |                                      |                         |
| Asia2       | Taiwan       | ML3            | NZ_AP014712.1         | 90.24%        | Clarithromycin                       | 23S rRNA -A2147G        |
| Asia3       | China        | XZ274          | NC_017926.1           | 91.12%        |                                      |                         |
| Europe1     | England      | ATCC 26695     | NC_000915.1           | 100%          | Quinolone, tetracycline, nitrosamide | HP1181                  |
| Europe2     | France       | B38            | NC_012973.1           | 90.06%        |                                      |                         |
| Europe3     | Germany      | P12            | NC_011498.1           | 94.04%        |                                      |                         |
| Australia1  | Australia    | ATCC 43504     | NZ_LS483488.1         | 92.35%        |                                      |                         |
| Australia2  | Australia    | BM012B         | NZ_CP007605.1         | 92.55%        |                                      |                         |
| Australia3  | Australia    | BM012S         | NC_022911.1           | 92.74%        |                                      |                         |

S. America: South America, N. America: North America



**Figure 2.** Distribution of resistance gene identifiers of strains included in this study (Green: Perfect Sequence hit, Red: Strict Sequence Hit and Yellow: Loose Sequence Hit). S. America: South America, N. America: North America

*In silico* analysis of resistance gene identifiers of the genomic data examined in our study is presented in Table 1 and Figure 2. While resistance to clarithromycin was found strict hit to South America<sup>2</sup> and Asia<sup>2</sup> strains, A2147G single nucleotide polymorphism (SNP) was detected in both strains related to clarithromycin resistance in 23S rRNA gene region. In Europe<sup>1</sup>, the HP1181 gene, responsible for the expression of resistance to quinolone, tetracycline, and nitrosamine demonstrated a perfect hit. In the genomic data of all strains, loose binding was found against different antibiotic groups at similar levels (Table 2).

According to the *in silico* comparative virulence analysis of the genomic data of the strains examined in our study, it was determined that the virulence of the strains were directed by a total of 95 related virulence genes under eight headings, including acid resistance, adherence, immune evasion, immune modulator, motility, secretion system, toxin and other factors. In the strains belonging to the Australian continent, *H. pylori* strains carried 90.1% of these 95 genes containing open reading frames (ORF) for the virulence genes. In the strains belonging to the North American continent, this rate was the lowest with 68.6% compared to the belonging to other continents. The samples taken from all continents, virulence genes belonging to seven ORF appeared active for acid resistance genes. Moreover, among the adhesins, differences were detected in the number of ORF that would carry the virulence genes of sialic acid-binding adhesins *sabA* / *hopP*, *sabB* / *hopT* and the adhesins that would bind blood group antigens, *babA* / *hopS* and *babB* / *hopT*. Strains carrying ORF for the sialic acid binding adhesin virulence factors, especially in Africa, Asia and North America were found to contain lower ORF counts than the other continents. When the ORF that would carry *futA*, *futB* and *futC* virulence genes from Lipopolysaccharide Lewis antigens to immune evasion were analyzed, it was found that their numbers were low in African and European strains. The immunomodulator and virulence factors involved in motility were similar between the strains from different continents. A similarity was found in the strains between the continents in the *napA* gene, which forms the neutrophil-activating protein with an immune modulator effect. In addition to this, the plasticity region was determined as the other virulence factor with the greatest difference in terms of the number of open reading frames in the strains obtained from different continents. The numbers of these regions were the highest in Australian and African strains. It was observed that virulence genes belonging to the *cagPAI* type secretion system contained less ORF in strains belonging to Africa and North America while it was the highest in the strains of Australia. It was observed that genes belonging to the *cagA*, *cagE* and *cagL* ORF, which are virulence genes of the secretion system, were found in all Australian, South American and Asian strains. When the ORF in association with toxin production were examined, it was found that vacuolated cytotoxin-generating ORF were lowest in Australian strains and in all strains of Europe, America, and Asia carried the *vacA* gene (Table 3). When the geographical distribution

of virulence genes, which are constantly examined in previous studies such as *cagA*, *cagE*, *cagL* and *vacA*, and it was found that the strains in Australia, Asia and South America often carried ORFs for *cagA*, *cagE* and *cagL* genes. However, the lowest level of virulence genes associated with *vacA* toxin production gene was detected in strains belonging to Australia (Table 4).

## DISCUSSION

Although half of the world population is colonized by *H. pylori* in their gastric mucosa, it is known that only 15% of them encounter infections, among which only 1% have severe conditions such as gastric cancer. The reason for this situation is explained as a result of the multifactorial nature of the infection (12). Compared to the low prevalence rate of *H. pylori* in geographical regions such as North America and Australia, this rate is higher in Africa and Asia. However, despite the high prevalence of *H. pylori* in Africa and Asia, the rates of gastric cancer formation do not correlate with these results. This remarkable situation has been defined as the "Asian and African enigma". The host's genetic and immune responses, virulence factors of different *H. pylori* strains, and environmental factors are used to explain this enigma (13).

Phuc et al., reported that Asian strains were mostly similar to European strains in their study conducted in Vietnam in 2021 (14). Delahay et al., in their phylogeny geographic study on *H. pylori* strains in 2018, reported that North American strains were distant from Asian strains in terms of similarity and formed a separate cluster, also they added that European and African strains formed a separate cluster from Asian strains (12). In our study, similar to the data of these two studies, we found that European strains were similar to Australian and Asian strains, while they had the lowest similarity to the strains of Africa. It has been concluded that the variations might differ depending on the genes for which the similarities of the strains are examined. Kumar et al., reported that the strains originating from Bangladesh showed 90%-92% similarity with the European strains, in their study conducted in 2021 (15). In our study, we also found that the Asian1 strain which is originating from India was similar to European strains.

Qumar et al. compared the virulence genes of strains obtained from different continents in 2021 and they reported that *cagPAI* virulence genes could be encoded in the genome in 18 of 20 strains originating from Bangladesh. Moreover, 10 of these 18 strains were found similar to Asian sources while eight of them resembled European sources. They reported that 90% of genes that cause virulence factors were found in their strains originating from Bangladesh (16). Similarly, in the virulence analysis of the genomic data of our isolates originating from Asia, we found this rate at the level of 89%.

Saribasak et al. stated that Asian strains have certain types of *cagA* in their study (16). In our study, we found that the presence of the *cagA* gene appears to be low in the strains belonging to North America, Africa and Europe, that supports this data.

**Table 3.** Distribution of comparative virulence genes analysis of all strains included in the study.

| Virulence Factor                             | Related gene number | Africa | S. America | N. America | Asia | Europe | Australia |
|--|---------------------|--------|------------|------------|------|--------|-----------|
| <b>Acid Resistance (1 item)</b>              |                     |        |            |            |      |        |           |
| Urease                                       | 7                   | 7      | 7          | 7          | 7    | 7      | 7         |
| <b>Adherence (8 Items)</b>                   |                     |        |            |            |      |        |           |
| Blood group antigen binding adhesins         | 2                   | 2      | 2          | 1.6        | 2    | 1.6    | 2         |
| Sialic acid binding adhesins                 | 2                   | 1      | 1.6        | 1          | 1    | 1.3    | 2         |
| HopZ   | 1                   | 1      | 1          | 1          | 1    | 1      | 1         |
| adherence-associated lipoprotein AlpA (hopC) | 1                   | 1      | 1          | 1          | 1    | 1      | 1         |
| AlpB (hopB)                                  | 1                   | 1      | 1          | 1          | 1    | 1      | 1         |
| H. pylori adhesin A                          | 1                   | 1      | 1          | 1          | 1    | 1      | 1         |
| HorB   | 1                   | 1      | 1          | 1          | 1    | 1      | 1         |
| PEB1   | 1                   | -      | -          | -          | -    | -      | -         |
| <b>Immune evasion (1 Item)</b>               |                     |        |            |            |      |        |           |
| Lipopolysaccharide Lewis antigens            | 3                   | 2.3    | 2.6        | 2.6        | 2.6  | 2.3    | 0.3       |
| <b>Immune modulator (2 Items)</b>            |                     |        |            |            |      |        |           |
| Neutrophil-activating protein (HP-NAP)       | 1                   | 1      | 1          | 1          | 1    | 1      | 1         |
| Outer inflammatory protein                   | 1                   | 1      | 1          | 0.6        | 1    | 1      | 1         |
| <b>Motility (1 Item)</b>                     |                     |        |            |            |      |        |           |
| Flagella                                     | 38                  | 36.6   | 37         | 37         | 36.6 | 36.6   | 37        |
| <b>Others (2 Items)</b>                      |                     |        |            |            |      |        |           |
| DupA (duodenal ulcer promoting)              | 1                   | -      | -          | -          | -    | -      | -         |
| Plasticity region                            | 3                   | 2      | 0.3        | 1          | 1.6  | 1      | .2.6      |
| <b>Secretion system (2 Items)</b>            |                     |        |            |            |      |        |           |
| Cag PAI type IV secretion system             | 26                  | 8      | 23.6       | 8          | 23.6 | 16.3   | 24        |
| T4SS effectors cytotoxin-associated gene A   | 1                   | 0.3    | 1          | 0.3        | 1    | 0.6    | 1         |
| <b>Toxin (2 Items)</b>                       |                     |        |            |            |      |        |           |
| Vacuolating cytotoxin                        | 1                   | 0.6    | 1          | 1          | 1    | 1      | 0.3       |
| Cytolethal distending toxin                  | 3                   | -      | -          | -          | -    | -      | -         |
| <b>Total ORF gene counts (Mean)</b>          | 95                  | 65.9   | 82.8       | 65.2       | 83.4 | 74.1   | 85.6      |
| <b>Total gene percentage (%)</b>             |                     | 69.3   | 87.1       | 68.6       | 89.8 | 78     | 90.1      |

S. America: South America, N. America: North America

Yamaoka et al. reported that geographic differences in gastric cancer cases depend not only on the differences of *cagA* and *vacA*, but also on the differences in other virulence factors such as *oipA* and *babA* (17). In our study, we determined that there

are differences in virulence genes such as *oipA* and *babA*, as well as differences in *cagA* and *vacA*. Erzin et al., in their study on Turkish patients in 2006, reported that *cagE* was an independent variable for duodenal ulcer and gastric cancer. They

**Table 4.** The status of the frequently studied virulence genes of the strains included in the study.

| Related virulence factor                | related gene      | Africa | S. America | N. America | Asia | Europe | Australia |
|---|-------------------|--------|------------|------------|------|--------|-----------|
| <b>Adherence</b>                        | <i>babA/hopS</i>  | 1.33   | 1.00       | 1.33       | 1.00 | 1.67   | 2.33      |
| <b>Adherence</b>                        | <i>babB/hopT</i>  | 1.00   | 1.00       | 0.67       | 1.00 | 1.67   | 1.00      |
| <b>Adherence</b>                        | <i>sabA/hopP</i>  | 1.33   | 1.00       | 0.67       | 1.33 | 0.67   | 1.00      |
| <b>Adherence</b>                        | <i>sabB/hopO</i>  | 0.00   | 0.67       | 0.33       | 0.00 | 1.00   | 1.00      |
| <b>Immune evasion</b>                   | <i>futA</i>       | 0.33   | 0.67       | 0.33       | 0.33 | 0.33   | 1.00      |
| <b>Immune evasion</b>                   | <i>futB</i>       | 1.33   | 1.00       | 1.67       | 1.33 | 1.67   | 1.00      |
| <b>Immune modulator</b>                 | <i>oipA/hopH</i>  | 1.00   | 1.33       | 0.67       | 1.67 | 1.00   | 1.00      |
| <b>Other</b>                            | Plasticity region | 2.00   | 0.33       | 1.00       | 1.67 | 1.00   | 2.67      |
| <b>Cag PAI type IV secretion system</b> | <i>cagA</i>       | 0.33   | 1.67       | 0.33       | 1.00 | 0.67   | 1.00      |
| <b>Cag PAI type IV secretion system</b> | <i>cagE</i>       | 0.33   | 1.00       | 0.33       | 1.00 | 0.67   | 1.00      |
| <b>Cag PAI type IV secretion system</b> | <i>cagL</i>       | 0.33   | 1.00       | 0.33       | 1.00 | 0.67   | 1.00      |
| <b>Toxin</b>                            | <i>vacA</i>       | 0.67   | 1.00       | 1.00       | 1.00 | 1.00   | 0.33      |

S. America: South America, N. America: North America

found *cagE* and *vacA* as biomarkers in duodenal ulcer patients and *cagE* and *babA2* as biomarkers in gastric cancer patients (18). Erzin et al, in 2008, examined host factors and bacterial virulence factors in their study, emphasized the importance of *babA2* in terms of cancer development and also determined the protective role of IL-1B 31TT genotype in the host (19).

While our *cagE* data may explain the low incidence of gastric cancer in Africa, similar to the results of this study, it contradicts the situation in Asia. Hence the *cagE* or *babA2* alone might not be the solution to the African-Asian enigma. Demiryas et al., in their study conducted in 2020, reported that when they compared the virulence genes of *H. pylori* strains found in gastric cancer, duodenal ulcer and non-ulcer dyspepsia patient groups, *cagL* was significantly different between the groups (20). The inclusion of different virulence factors in all these studies indicates that our knowledge about *H. pylori*'s direction of cancer development is limited. Kocazeybek et al., in 2015, showed that not only the presence of *cagA* but also special motifs or patterns such as EPIYA seen in the *cagA* region could cause geographical differences in this carcinogenesis process (21). In a study conducted in 2020, Saribas et al., showed that gastrointestinal diseases can be explained not only by EPIYA patterns but also by including host genetic factors such as HLAs (22). Kocak et al., reported that in patients with *cagL* and *cagA* positive *H. pylori*, bacterial virulence factors, as well as host genetic factors such as HLAs, are also involved in the gastrointestinal disease process (23). Like Erzin et al., (19), Saribaş et al., (22) and Kocak et al., (23) emphasized the importance of host factors in the process. It has strengthened our belief that our knowledge of the mechanism of cancer development of *H. pylori* is lacking. On the other hand, Sun et al., in their study

in 2020, regarding the virB11 protein produced by the *virB11* gene located in the *H. pylori* plasticity region, found that this protein plays a role in the ATP formation mechanism of *H. pylori* and is important for providing energy (24). Yamaoka showed the protective effect of the plasticity region against gastric cancer (25). Besides Africa and Asia strains, as well as Australia strains, the amount of ORF plasticity region carriage was found high in our study, so this virulence gene alone cannot explain the Africa-Asian enigma.

In the study Mwangi et al., conducted in 2020, they stated that the *H. pylori* strain they isolated from a patient in Kenya was similar to Asian stains. Although the main virulence factors they detected were found in other African strains, it was reported that this strain was not similar to other African strains in terms of phylogeny (26). In our study, we found that our results were similar in terms of virulence factors, ORF and phylogenies, but that African strains we studied generally differed from European and Asian strains in terms of both virulence gene prevalence and phylogeny. Lamichhane et al., reported that they found subgroups of Australian strains belonging to people coming from Europe 200 years ago, and the strains they examined were similar to Australian strains (27). Among the strains we examined in our study, we found that Australian strains have more similarities to European strains than other continents.

As Suzuki et al., pointed out, the geographic differences of *H. pylori* strains may reveal human migrations. They also noted that the genetic diversity within *H. pylori* strains was much greater than in other bacterial strains and 50 times greater than in human populations. In addition, they reported that with the molecular epidemiological studies carried out with new molecular sequencing techniques, we can have more information about

both virulence genes and resistance genes, and the mechanisms of gastric cancer formation can thus be understood (28).

Boyanova et al. reported that more than 20% of clarithromycin resistance was seen in Asia, Europe and South America, and above 10% of quinolone resistance was seen in Asia (29). In our study, we detected clarithromycin resistance and A2147G SNP in 23S rRNA gene causing this resistance in an Asian and a South American strain. However, we found a gene associated with quinolone resistance only in our European strain, this could be related to the limited number of strains that we included in this study. Kocazeybek and Tokman reported that primary antibiotic resistance detected in *H. pylori* strains were affected by geographic differences and crowded population (3). This is consistent with our resistance data.

The limitations of our study were both not using the genomic data of all strains in the NCBI database and not having our own strains in the study. The strains we included in our study appear to be mostly selected reference strains in many publications. In addition, although the NCBI database contains data on more than 2000 strains, the whole genome sequencing of all these strains has not yet been completed, but the whole genome sequencing of all the strains we have included in our study has been completed.

## CONCLUSIONS

In conclusion, our study has revealed that *H. pylori* strains, whose data were reported from different continents, vary from each other phylogenetically and especially in terms of virulence genes. We believe that it is important to reveal the virulence genes of these strains by using new molecular sequencing techniques, in order to reveal the situation described as both the African-Asian enigma and the pathogenesis mechanisms used by *H. pylori* during gastric cancer.

**Ethics Committee Approval:** Ethics committee approval is not required because of no material or experimental animal that would require permission.

**Peer-review:** Externally peer-reviewed.

**Author Contributions:** Conception/Design of Study - M.D., Ö.Ü., B.S.K.; Materials - M.D., Ö.Ü., B.S.K.; Data Collection - M.D., Ö.Ü., B.S.K.; Analysis and/or Interpretation - M.D., Ö.Ü., B.S.K.; Drafting Manuscript - M.D., Ö.Ü., B.S.K.; Final Approval and Accountability - M.D., Ö.Ü., B.S.K.

**Conflict of Interest:** The authors have no conflict of interest to declare.

**Financial Disclosure:** The authors declared that this study has received no financial support.

**Etik Komite Onayı:** Bu çalışmada, etik komite iznine gerek duyulacak bir materyal ya da deney hayvanı kullanılmamıştır.

**Hakem Değerlendirmesi:** Dış bağımsız.

**Yazar Katkıları:** Çalışma Konsepti/Tasarımı - M.D., Ö.Ü., B.S.K.; Materyal: M.D., Ö.Ü., B.S.K.; Veri Toplama - M.D., Ö.Ü., B.S.K.; Veri Analizi/Yorumla-

ma - M.D., Ö.Ü., B.S.K.; Yazma - M.D., Ö.Ü., B.S.K.; Son Onay ve Sorumluluk - M.D., Ö.Ü., B.S.K.

**Çıkar Çatışması:** Yazarlar çıkar çatışması bildirmemişlerdir.

**Finansal Destek:** Yazarlar bu çalışmada finansal destek almadıklarını beyan etmişlerdir.

## REFERENCES

1. Marshall BJ, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1984; 1(8390): 1311-5. [CrossRef]
2. Keikha M, Karbalaeei M. Correlation between the geographical origin of *Helicobacter pylori* homB-positive strains and their clinical outcomes: a systematic review and meta-analysis. *BMC Gastroenterol* 2021; 21: 181. [CrossRef]
3. Kocazeybek B, Tokman HB. Prevalence of Primary Antimicrobial Resistance of *H. pylori* in Turkey: A Systematic Review. *Helicobacter* 2016; 21: 251-60. [CrossRef]
4. Chang WL, Yeh YC, Sheu BS. The impacts of *H. pylori* virulence factors on the development of gastroduodenal diseases. *J Biomed Sci* 2018; 25: 68. [CrossRef]
5. de Brito BB, da Silva FAF, Soares AS, Pereira VA, Santos MLC, Sampaio MM, et al. Pathogenesis and clinical management of *Helicobacter pylori* gastric infection. *World J Gastroenterol* 2019; 25: 5578-89. [CrossRef]
6. Thyagarajan SP, Ray P, Das BK, Ayyagari A, Khan AA, Dharmalingam S, et al. Geographical difference in antimicrobial resistance pattern of *Helicobacter pylori* clinical isolates from Indian patients: Multicentric study. *J Gastroenterol Hepatol* 2003; 18: 1373-8. [CrossRef]
7. Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappuoli R. *Helicobacter pylori* virulence and genetic geography. *Science* 1999; 284(5418): 1328-33. [CrossRef]
8. Gilchrist CA, Turner SD, Riley MF, Petri WA Jr, Hewlett EL. Whole-genome sequencing in outbreak analysis. *Clin Microbiol Rev*. 2015; 28: 541-63. [CrossRef]
9. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One* 2014; 9: e104984. [CrossRef]
10. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2020; 48: D517-D525. [CrossRef]
11. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res* 2019; 47(D1): D687-D692. [CrossRef]
12. Delahay RM, Croxall NJ, Stephens AD. Phylogeographic diversity and mosaicism of the *Helicobacter pylori* tfs integrative and conjugative elements. *Mob DNA* 2018; 9: 5. [CrossRef]
13. Ofori EG, Adinortey CA, Bockarie AS, Kyei F, Tagoe EA, Adinortey MB. *Helicobacter pylori* Infection, Virulence Genes' Distribution and Accompanying Clinical Outcomes: The West Africa Situation. *Biomed Res Int* 2019; 2019: 7312908. [CrossRef]
14. Phuc BH, Tuan VP, Dung HDQ, Binh TT, Tung PH, Tri TD, et al. *Helicobacter pylori* type 4 secretion systems as gastroduodenal disease markers. *Sci Rep* 2021; 11: 4584. [CrossRef]
15. Kumar S, Nguyen TH, Nahar S, Sarker N, Baker S, Bulach D, et al. A comparative whole genome analysis of *Helicobacter pylori* from a human dense South Asian setting. *Helicobacter* 2021; 26: e12766. [CrossRef]



16. Saribasak H, Salih BA, Yamaoka Y, Sander E. Analysis of Helicobacter pylori genotypes and correlation with clinical outcome in Turkey. *J Clin Microbiol* 2004; 42(4): 1648-51. [\[CrossRef\]](#)
17. Yamaoka Y, Kato M, Asaka M. Geographic differences in gastric cancer incidence can be explained by differences between Helicobacter pylori strains. *Intern Med* 2008; 47: 1077-1083. [\[CrossRef\]](#)
18. Erzin Y, Koksall V, Altun S, Dobrucali A, Aslan M, Erdamar S, et al. Prevalence of Helicobacter pylori vacA, cagA, cagE, iceA, babA2 genotypes and correlation with clinical outcome in Turkish patients with dyspepsia. *Helicobacter* 2006; 11: 574-80. [\[CrossRef\]](#)
19. Erzin Y, Koksall V, Altun S, Dobrucali A, Aslan M, Erdamar S, et al. Role of host interleukin 1beta gene (IL-1B) and interleukin 1 receptor antagonist gene (IL-1RN) polymorphisms in clinical outcomes in Helicobacter pylori-positive Turkish patients with dyspepsia. *J Gastroenterol* 2008; 43(9): 705-10. [\[CrossRef\]](#)
20. Demiryas S, Caliskan R, Saribas S, Akkus S, Gareayaghi N, Kirmusaoglu S, et al. The association between cagL and cagA, vacAs-m, babA genes in patients with gastric cancer, duodenal ulcer, and non-ulcer dyspepsia related to Helicobacter pylori. *Acta Gastroenterol Belg* 2020; 83(3): 385-92.
21. Kocazeybek BS, Caliskan R, Erdamar Cetin S, Ergin S, Kuskucu M, Kepil N, et al. Patterns of EPIYA motifs among cagA-positive Helicobacter pylori strains: a case-control study in a Turkish population with Eurasian geographical features. *J Med Microbiol* 2015; 64(10): 1117-23. [\[CrossRef\]](#)
22. Saribas S, Demiryas S, Yilmaz E, Uysal O, Kepil N, Demirci M, et al. Association between human leukocyte antigen gene polymorphisms and multiple EPIYA-C repeats in gastrointestinal disorders. *World J Gastroenterol* 2020; 26: 4817-32. [\[CrossRef\]](#)
23. Kocak BT, Saribas S, Demiryas S, Yilmaz E, Uysal O, Kepil N, et al. Association between polymorphisms in HLA-A, HLA-B, HLA-DR, and DQ genes from gastric cancer and duodenal ulcer patients and cagL among cagA-positive Helicobacter pylori strains: The first study in a Turkish population. *Infect Genet Evol* 2020; 82: 104288. [\[CrossRef\]](#)
24. Sun DL, Gong ZH, Shao SL, Shi XL, Yuan XY, Luo H, et al. virB11 gene potentially involves in ATP metabolism to provide energy in H. pylori infection. *Microb Pathog* 2020; 142: 104067. [\[CrossRef\]](#)
25. Yamaoka Y. Roles of the plasticity regions of Helicobacter pylori in gastroduodenal pathogenesis. *J Med Microbiol* 2008; 57: 545-53. [\[CrossRef\]](#)
26. Mwangi C, Njoroge S, Tshibangu-Kabamba E, Moloo Z, Rajula A, Devani S, et al. Whole Genome Sequencing Reveals Virulence Potentials of Helicobacter pylori Strain KE21 Isolated from a Kenyan Patient with Gastric Signet Ring Cell Carcinoma. *Toxins (Basel)* 2020; 12: 556. [\[CrossRef\]](#)
27. Lamichhane B, Wise MJ, Chua EG, Marshall BJ, Tay CY. A novel taxon selection method, aimed at minimizing recombination, clarifies the discovery of a new sub-population of Helicobacter pylori from Australia. *Evol Appl* 2019; 13: 278-89. [\[CrossRef\]](#)
28. Suzuki R, Shiota S, Yamaoka Y. Molecular epidemiology, population genetics, and pathogenic role of Helicobacter pylori. *Infect Genet Evol* 2012; 12(2): 203-13. [\[CrossRef\]](#)
29. Boyanova L, Mitov I. Geographic map and evolution of primary Helicobacter pylori resistance to antibacterial agents. *Expert Rev Anti Infect Ther* 2010; 8(1): 59-70. [\[CrossRef\]](#)