Molecular Phylogenetic Analysis of 16S rRNA Sequences Identified Two lineages of
H. pylori Strains Detected from Different Regions in Sudan Suggestive of
Differential Evolution

Authors

Abeer Babiker Idris,1# Hadeel Gassim Hassan,1 Maryam Atif Salaheldin Ali,1* Sulafa Mohamed Eltaher,2* Leena Babiker Idris,3# Hisham N. Altayb,4 Amin Mohamed Abass,5 Mustafa Mohammed Ahmed Ibrahim,5 El-Amin Mohamed Ibrahim,1 Mohamed A. Hassan.6,7

*These authors contributed equally to this work
#Corresponding authors: abeer.babiker89@gmail.com; dr.leena.babiker@gmail.com

1Department of Medical Microbiology, Faculty of Medical Laboratory Sciences, University of Khartoum, Khartoum, Sudan
2The Academy of Health Sciences, The Republic of Sudan Federal Ministry of Health, Khartoum, Sudan
3Faculty of Medicine, The International University of Africa, Khartoum, Sudan
4Department of Microbiology, Faculty of Medical Laboratory Science, Sudan University of Science and Technology, Khartoum, Sudan
5Department of internal Medicine, University of Bahri, Khartoum-Sudan
6Consultant Physician and Gastroenterology, Khartoum-Sudan
6Department of Bioinformatics, Africa city of technology, Khartoum-Sudan
7Department of Bioinformatics, DETAGEN Genetic Diagnostics Center, Kayseri, Turkey

Abstract

Background: H. pylori is ubiquitous among humans, and one of the best studied examples of an intimate association between bacteria and humans. Under several diverse socio-demographic factors in Sudan, a continuous increase in the prevalence rate of H. pylori infection has been noticed which represents a major public health challenge. In this study, we analyzed the molecular evolution of H. pylori Strains detected from different ethnic and regions of Sudan using 16S rRNA gene and phylogenetic approach.

Materials and methods: A total of 75 gastric biopsies taken from patients who had been referred for endoscopy from different regions of Sudan. The DNA extraction was done by using the guanidine chloride method. Two sets of primers (universal and specific for H. pylori) were used to amplify the 16S ribosomal gene. Sanger sequencing was performed; then Blast these sequences with those available in the NCBI nucleotide database. The evolutionary aspects were analyzed using a MEGA7 software.

Result: Molecular detection of H. pylori has shown that 28 (37.33%) of patients were positive for H. pylori. Bivariate analysis has found no significant differences exhibited across sociodemographic, endoscopy series and H. pylori infection. Nucleotide variations were found at five nucleotide positions (positions 219, 305, 578, 741 and 763-764) and one insertion mutation (750_InsC_751) was present in sixty-seven percent (7/12) of our strains. The phylogenetic tree diverged into two lineages.

Conclusion: The phylogenetic analysis of 16S rRNA sequences identified two lineages of H. pylori strains detected from different regions in Sudan. Sex mutations were detected in regions of the 16S rRNA not closely associated with either tetracycline or tRNA binding sites. 66.67% of them were located in the central domain of 16S rRNA. Studying the effect of these mutations on the functions of 16S rRNA molecules in protein synthesis and antibiotic resistance is of great importance.

Keywords: H. pylori, 16S rRNA gene, evolutionary tree, Sudan
1. Introduction

*H. pylori* is ubiquitous among humans,\(^{(1)}\) and one of the best studied examples of an intimate association between bacteria and humans.\(^{(2)}\) It has infected human around 100,000 years ago (range 88,000-116,000).\(^{(1, 3)}\) and has co-evolved with human ancestors for approximately 58000 ± 3500 years, during their first migrations from east Africa.\(^{(1, 4, 5)}\) And it largely escaped notice until it was cultured by Marshall and Warren.\(^{(6)}\) In fact, *H. pylori* possesses several properties that helping this bacterium to persist for several decades, transmit from generation to generation and make an intimate association with its human host.\(^{(2, 7)}\) *H. pylori* infection is predominantly transmitted vertically from parent to child and between individuals in close contact such as in a family.\(^{(2, 8)}\) This close transmission pattern has resulted in a clear phylogeographic differentiation within these bacteria because of the local dispersion of single nucleotide polymorphisms by high rates of homologous recombination.\(^{(9, 10)}\) However, under poverty and inappropriate conditions, especially in the developing countries, *H. pylori* can be transmitted horizontally which makes infection with multiple strains probably more common than in the developed world.\(^{(11, 12)}\)

In Sudan, the population is culturally, linguistically and ethnically diverse with more than 597 tribes who speak more than 400 dialects and languages.\(^{(13, 14)}\) The majority of the Sudanese population is rural with an urban population of just 33.2%, most of them are in Khartoum.\(^{(14, 15)}\) Sudan has severely suffered war, famine and flood in recent decades and has a large population of IDPs.\(^{(16, 17)}\) Although Sudan is rich in terms of natural and human resources, the effects of the civil conflict on health, nutrition, population, the economic and social development have undoubtedly been significant. These several diverse socio-demographic factors lead to continually increasing in the prevalence rate of *H. pylori* infection ranging from 48% to 65.8% which represents a major public health challenge.\(^{(18, 19)}\)

However, studying the genetics of the *H. pylori* population has been of great interest due to its clinical and phylogeographic significance.\(^{(4)}\) A number of studies have discovered the importance of evolutionary history to the clinical outcome of *H. pylori* infection; and indicated that the scramble of the relationship between bacterial and human ancestries at the individual level due to a consequence of migrations, invasions or racial admixture may moderate adverse outcomes for the host and disrupting the selection for a reduced virulence; and this may give some explanation to the continental enigmas of *H. pylori*.\(^{(12, 20-23)}\) Also, Phylogeny and Phylogeography of *H. pylori* strains are known to mirror human migration patterns\(^{(24-27)}\) and reflect significant demographic events in human prehistory.\(^{(25, 28)}\) Therefore, the phylogeographic pattern of *H. pylori* is a discriminatory tool to investigate human evolution and migration in addition to the traditional human genetic tools, e.g. mitochondrial DNA (mtDNA) and languages.\(^{(1)}\)
16S rRNA gene amplification and sequencing have been extensively used for bacterial phylogeny and taxonomy; and eventually, the establishment of large public-domain databases. (29-32) Several properties of the 16S rRNA gene make it the “Ultimate molecular chronometer”, (33) and the most common housekeeping genetic marker; and hence a useful target for clinical identification and phylogeny. (34, 35) These properties include: First, its present in all bacteria, often existing as a multigene family or operons, thus it is a universal target for bacterial identification. (32, 35, 36) Second, the function of 16S rRNA has not changed over a long period, so random sequence changes are more likely to reflect the microbial evolutionary change (phylogeny) than selected changes which may alter the molecule’s function. (33) Finally, the 16S rRNA gene is large enough, approximately 1,500bp, for informatics purposes. (34, 35, 37) Most importantly that the 16S rRNA gene consists of approximately 50 functional domains and any introduction of selected changes in one domain does not greatly affect sequences in other domains i.e. less impact selected changes have on phylogenetic relationships. (35)

Here in this study, two sets of primers (universal and specific for H. pylori) were used to amplify the 16S ribosomal gene directly from gastric endoscopic biopsy samples collected from dyspeptic patients who had been referred for endoscopy. Sanger sequencing was performed and by matching these sequences with those available in the National Center for Biotechnology Information (NCBI) nucleotide database, we analyzed the evolutionary aspects of Sudanese H. pylori strains using a phylogenetic approach. The novelty of our study resides in being the first study to characterize H. pylori Sudanese strains detected from different regions and ethnic of Sudan using the 16S rRNA gene.

2. Material and methods

2.1 Study design and Study settings

A prospective cross-sectional hospital-based study was conducted in public and private hospitals in Khartoum state from June 2018 to September 2019. These hospitals receive patients from all over the regions of Sudan. Currently, there are 16 states in Sudan which divided between four geographic regions: Eastern, Western, Northern and Central Sudan which includes the largest metropolitan area, Khartoum (that includes Khartoum, Khartoum North and Omdurman), (38) see a map in Figure 3(a) for more illustration. The capital Khartoum is quickly growing and ranges between 6 and 7 million of populations, which includes approximately 2 million internally displaced persons (IDPs) from the southern war zone and the drought-affected areas in the west and east. (14, 17)

The hospitals include Ibin Sina specialized hospital, Soba teaching hospital, Modern Medical Centre and Al Faisal Specialized Hospital. All laboratory processes were
performed in the Molecular biological lab at the Faculty of Medical Laboratory Sciences at the University of Khartoum.

2.2 Study population

The study population composed of 75 patients who had been referred for endoscopy and most of them because of dyspepsia. Structured questionnaires were provided for participants to obtain information about their socio-demographic characteristics. Patients were selected from those who were not taking antibiotics or NSAIDS. The patients gave written informed consent before they enrolled in the study. The diagnosis of gastroduodenal diseases was based on the investigation of an experienced gastroenterologist during the upper GI endoscopy procedure. While gastric cancer was diagnosed based on, in addition to gastroscopy, histology.

2.3 Sample collection

For DNA extraction purposes, gastric biopsies were collected in 1.5µl Eppendorf tubes with 400µl phosphate buffer saline (PBS), while for histological examination, the biopsies were transported in formalin. Then the biopsies were labeled and transported immediately to the laboratory for further processes.

2.4 DNA extraction

The DNA extraction was done by using the manual guanidine chloride method, as described by Alsadig et al. Briefly, biopsies were washed with 400µl PBS and centrifuged for 5min at 100 rpm after each wash. Then, the samples were subjected to digestion by adding 400µl of WBCs lysis buffer, 200µl of 6M Guanidine chloride, 50µl of 7.5M Ammonium Acetate and 5µl of 20mg/µl proteinase K; and incubated at 37°C overnight. After that, On the following day samples were cooled down to room temperature then added 400µl of cooled pre-chilled Chloroform and centrifuged at 1000rpm for 5min. Then three layers were separated and the supernatant was collected to a new labeled Eppendorf. After that, 1ml of cooled pre-chilled absolute Ethanol was added and mixed gently back and forth quickly. Then samples were put in -20°C freezer overnight. After that, samples were subjected to quick vortex for one minute then centrifuged for 5min at 1000rpm; and the supernatants were discarded. Then washing with 70% ethanol was performed and the supernatant was drained with much care to avoid losing the DNA pellet at the bottom of the Eppendorf. Then the Eppendorf was inverted upside down of a tissue paper leaving the pellet to dry from alcohol for at least 2hours. Finally, the DNA pellet was re-suspended in 35µl of deionized water and was put into -20°C until use.

2.5 16S rRNA gene amplification

Extracted DNA was amplified for the universal 16S rRNA gene using the following primers: (primers: F:5’-AGAGTTTGATCCTGGCTCAG-3’) (R:5’-
CTACGGCTACCTTGTTACG-3'). (40) PCR amplification was carried out with Maxime PCR PreMix Kit (i-Taq) (iNtRON BIOTECHNOLOGY, Seongnam, Korea) and a PCR thermocycler (SensoQuest, Germany). The PCR reaction mixtures contained 2.5U of i-Taq TM DNA polymerase (5U/µl), 2.5mM of each deoxynucleoside tri- phosphates (dNTPs), 1X of PCR reaction buffer (10X),1X of gel loading buffer and 1µl of DNA template. The temperature cycle for the PCR was carried out using a method described previously. (41, 42)

To detect the DNA, 3µl of each PCR products was loaded onto 2% agarose gels stained with 3µl ethidium bromide (10mg/ml) and subjected to electrophoresis in 1x Tris EDTA Buffer (TEB buffer) (89mM of Tris base, 89mM Boric acid and 2mM EDTA dissolved in 1Litter H2O) for 30 min at 120V and 50mA. The gel was visualized under UV light illumination. A 100MW DNA ladder (iNtRON BIOTECHNOLOGY, Seongnam, Korea) was used in each gel as a molecular size standard. The amplified product for the rrs gene is 1500bp.

The DNA amplification for the specific 16S rRNA gene was performed to confirm the infection of H. pylori. (primers: F:5'-GCGCAATCAGCGTCAGGTAATG-3') (R:5'-GCTAAGAGAGCAGCCTATGTCC-3'). A PCR reaction was carried out using a previously described method by Abeer et al. (43)

2.6 Sequencing of H. pylori 16S rRNA gene

The amplified 16S rRNA gene (for the universal and specific) was purified and sequenced, using Sanger dideoxy sequencing method, commercially by Macrogen Inc, Korea.

2.7 Bioinformatics analysis

2.7.1 Sequence analysis

The nucleotide sequence was visualized and analyzed by using the Finch TV program version 1.4.0. (44) The nucleotide Basic Local Alignment Search Tool (BLASTn; https://blast.ncbi.nlm.nih.gov/) was used for searching about the similarity with other sequences deposited in GenBank. (45) The 16S rRNA sequence was submitted in the GenBank nucleotide database under the following accession numbers: from MN845181 to MN845190 and from MN845952 to MN845954.

2.7.2 Molecular Phylogenetic Analysis

Highly similar sequences were retrieved from NCBI GenBank and subjected to multiple sequence alignment (MSA) using Clustal W2 (46)-BioEdit software. (47) Gblocks was used to eliminate poorly aligned positions and divergent regions of aligned sequences so the
alignment becomes more suitable for phylogenetic analysis. The Neighbor-Joining phylogenetic tree of our 16S rRNA sequences with those obtained from the database was constructed using Jukes-Cantor (JC) model from Substitution (ML) model. The tree was replicated 1000 replicates in which the association with taxa clustered together in the bootstrap test. Molecular Evolutionary Genetics Analysis Version 7.0 (MEGA7) was used to conduct evolutionary analyses.

2.8 Statistical analysis

Data were analyzed using GraphPad Prism 5. Regarding the prevalence of *H. pylori* infection, differences in frequency distribution by age were examined by the Mann-Whitney test. While bivariate analysis with a categorical variable was assessed by the \( \chi^2 \) test or Fisher's test. The statistical significance level was determined at \( P<0.05 \).

3. Result

3.1 Characteristics of the study population

A total of seventy-five patients were included. Forty-one patients (54.67%) were male and thirty-four (45.33%) female. Forty-two patients (56%) were urban, and thirty-three (44%) were rural. The patients’ age ranged from 15 to 85 years, with a mean age of 45.11±17.45 years. Most of the participants came from northern and central Sudan. And their ethnicities were distributed as follows: Shagia (10, 13.33%), Jalyeen (9, 12%), Mahas (8, 10.67%), Rezaigat (4, 5.33%), Zaghawa (4, 5.33%), Kawahla (4, 5.33%), Masalamyia (3, 4%) and Other (33, 44%). Regarding clinical symptoms of gastrointestinal disturbances of the participants, abdominal pain was the major symptom (22, 29.33%), followed by nausea (16, 21.33%). Molecular detection of *H. pylori* has shown that 28 (37.33%) of patients were positive for *H. pylori*. (Figure1(b)) Patients from western Sudan were more prone to *H. pylori* infection 50% (6/12). Bivariate analysis has found no significant differences exhibited across sociodemographic, endoscopy series and *H. pylori* infection, as illustrated in Table1 and Table2.
### Table 1. Socio-demographic characteristic of patients

| Variables                        | Total n=75 | H. pylori (+ve) n=28 | H. pylori (-ve) n=47 | P-value |
|----------------------------------|------------|----------------------|----------------------|---------|
| **Mean age (years) ± Std. Deviation (range)** |            |                      |                      |         |
|                                  | 45.11±17.45 (15-85) | 41.39±17.57 (15-85) | 47.32±17.18 (21-80) | 0.1170  |
| **Gender**                       |            |                      |                      |         |
| Male                             | 41 (54.67%) | 15 (53.57%)          | 26 (55.32%)          | 1.0000  |
| Female                           | 34 (45.33%) | 13 (46.43%)          | 21 (44.68%)          |         |
| **Residence**                    |            |                      |                      | 0.8122  |
| Urban                            | 42 (56%)   | 15 (53.57%)          | 27 (57.45%)          |         |
| Rural                            | 33 (44%)   | 13 (46.43%)          | 20 (42.55%)          |         |
| **Hospital**                     |            |                      |                      | 0.7986  |
| Public                           | 51 (68%)   | 20 (71.43%)          | 31 (65.96%)          |         |
| Private                          | 24 (32%)   | 8 (28.57%)           | 16 (34.04%)          |         |
| **History of previous infection**|            |                      |                      | 0.8122  |
| Yes                              | 42 (56%)   | 15 (53.57%)          | 27 (57.45%)          |         |
| No                               | 33 (44%)   | 13 (46.43%)          | 20 (42.55%)          |         |
| **The Frequency of recurrence (If Yes n=42)** | |                      |                      |         |
| One time                         | 32 (76.19%)| 10 (66.67%)          | 22 (81.49%)          | 0.4508  |
| More than one time               | 10 (23.81%)| 5 (33.3%)            | 5 (18.52%)           |         |
| **Family history of infection**  |            |                      |                      | 0.4833  |
| Yes                              | 36 (48%)   | 15 (53.58%)          | 21 (44.68%)          |         |
| No                               | 39 (52%)   | 13 (46.43%)          | 26 (55.32%)          |         |
| **Smoking**                      |            |                      |                      | 0.7718  |
| Yes                              | 16 (21.33%)| 5 (17.86%)           | 11 (23.4%)           |         |
| No                               | 59 (78.67%)| 23 (82.14%)          | 36 (76.6%)           |         |
| **Tribe**                        |            |                      |                      | 0.5733  |
| Shagia                           | 10 (13.33%)| 4 (14.29%)           | 6 (12.77%)           |         |
| Jalyeen                          | 9 (12%)    | 4 (14.29%)           | 5 (10.64%)           |         |
| Mahas                            | 8 (10.67%) | 3 (10.71%)           | 5 (10.64%)           |         |
| Rezaigat                         | 4 (5.33%)  | 1 (3.57%)            | 3 (6.38%)            |         |
| Zaghawa                          | 4 (5.33%)  | 3 (10.71%)           | 1 (2.13%)            |         |
| Kawahla                          | 4 (5.33%)  | 2 (7.14%)            | 2 (4.26%)            |         |
| Masalamyia                       | 3 (4%)     | 2 (7.14%)            | 1 (2.13%)            |         |
| Other                            | 33 (44%)   | 9 (32.14%)           | 24 (51.06%)          |         |
| **Region**                       |            |                      |                      | 0.2697  |
| Northern Sudan                   | 15 (20%)   | 6 (21.43%)           | 9 (19.15%)           |         |
| Central Sudan                    | 31 (41.33%)| 12 (42.86%)          | 19 (40.43%)          |         |
| Eastern Sudan                    | 3 (4%)     | 2 (7.14%)            | 1 (2.13%)            |         |
| Western Sudan                    | 12 (16%)   | 6 (21.43%)           | 6 (12.77%)           |         |
| Southern Sudan                   | 14 (18.67%)| 2 (7.14%)            | 12 (25.53%)          |         |
Table 2. Distribution of participants according to endoscopy series and status of *H. pylori* infection

| Endoscopy results          | No. (%) | *H. pylori* (+ve) | *H. pylori* (-ve) | P-value   |
|----------------------------|---------|-------------------|-------------------|-----------|
| Normal gastric findings    | 9 (12%) | 4 (14.29%)        | 5 (10.64%)        | 0.5259    |
| Esophagitis                | 4 (5.33%) | 3 (10.71%)   | 1 (2.13%)         |           |
| Esophageal varices         | 6 (8%)  | 1 (3.57%)        | 5 (10.64%)        |           |
| Gastritis                  | 26 (34.67%) | 8 (28.57%) | 18 (38.3%)        |           |
| Duodenitis                 | 6 (8%)  | 3 (10.71%)        | 3 (6.38%)         |           |
| Peptic ulcer               | 6 (8%)  | 3 (10.71%)        | 3 (6.38%)         |           |
| Cancer                     | 18 (24%) | 6 (21.43%)        | 12 (25.53%)       |           |
| Total                      | 75 (100%) | 28 (37.33%) | 47 (62.67%)       |           |

3.2 Analysis of 16S rRNA sequences

12 sequences of 16S rRNA of *H. pylori* from Sudanese patients were analyzed for mutations and their conservative nature, and findings revealed diversity with few differences. The description of Sudanese *H. pylori* strains is shown in Table 3. The amplicon of universal 16S rRNA for patients 22 resulted in *Acinetobacter radioresistens* (Figure 1(c)). The patient 22 was male with 33 years and he is Mahasi from northern Sudan, and he was diagnosed with antral gastritis.

Regarding mutations, nucleotide variations were found at five nucleotide positions (positions 219, 305, 578, 741 and 763-764) and one insertion mutation (750_InsC_751) was present in sixty-seven percent (7/12) of our strains, the numbering was according to the *E.coli* sequence). Three strains (17, 21 and 39) revealed a novel transition C→T at position 219. A triple base pair substitution (GCC763-764CAA) was detected in strain 17, 23 and 39. However, a cancerous strain 101 had two G → A transitions at positions 741 and 765. Strain 44, which was detected in a patient with normal gastric finding and suspected celiac disease and lives in Khartoum, had also a novel C→T transversion at nucleoside position 578. Also strain 39, which was found in a patient from west of Sudan (Eastern Darfur) and was diagnosed with antral gastritis, revealed a novel G→A transition at positions 305. See Figure 2 for more illustration.
Table 3. Description of Sudanese *H. pylori* strains.

| Strain | Age of patient | Gender of patient | Clinical origin                  | Ethnic origin | Province         | Region             | Accession No. |
|--------|----------------|-------------------|----------------------------------|---------------|------------------|--------------------|---------------|
| 12     | 28             | Female            | Antral gastritis                 | Rezaigat      | Eastern Darfur   | West of Sudan      | MN845953      |
| 17     | 23             | Female            | Distal esophagitis and antral gastritis | Halawyeen   | Khartoum         | Middle of Sudan    | MN845187      |
| 21     | 32             | Male              | Antral gastritis                 | Kawahla       | Sinnar           | Southeast of Sudan | MN845181      |
| 23     | 15             | Female            | Distal esophagitis and antral gastritis | Manaseer   | Kartoum          | Middle of Sudan    | MN845188      |
| 39     | 58             | Male              | Antral gastritis                 | Masalamyia   | Eastern Darfur   | West of Sudan      | MN845189      |
| 40     | 68             | Male              | Duodenitis                       | -             | Khartoum         | Middle of Sudan    | MN845182      |
| 44     | 35             | Female            | Normal gastric finding           | Jalyeen       | Khartoum         | Middle of Sudan    | MN845183      |
| 46     | 52             | Male              | Duodenitis                       | Shagia        | Northern State   | North of Sudan     | MN845184      |
| 47     | 35             | Male              | GRED                             | Shagia        | Northern State   | North of Sudan     | MN845954      |
| 59     | 36             | Male              | Normal gastric finding           | Shagia        | Northern State   | North of Sudan     | MN845190      |
| 101    | 60             | Female            | Gastric cancer                   | Zaghawa       | North Kordofan   | South of Sudan     | MN845185      |
| 102    | 56             | Male              | Gastric cancer                   | Zaghawa       | Northern Darfur  | East of Sudan      | MN845186      |

3.4 Phylogenetic analysis

The phylogenetic tree diverged into two lineages. In the lineage one, all the Sudanese *H. pylori* strain clustered with strains from different countries. The *16S rRNA* sequence of strain 44, although clustered with other global strains in one clade, had a novel C→T transversion at nucleotide position 578. However, cancerous strain 101 shared a common ancestor with strains from the USA (AF535198) and Venezuela (DQ829805); and represented with them a separated clade with a bootstrap value of 57%. A novel single nucleotide variation (G→A) at position 305 of the *16S rRNA* of strain 23 made it outgroup as a kind of strain evolution with a bootstrap value of 61%.

In the second lineage, strains 17, 23 and 39 were closely related to strains from Hungary (KF297893 and KF297892). Strain 17 and Strain 39, from Khartoum and Eastern Darfur, respectively, were sisters with a bootstrap value of 75%, as presented in Figure 3.
4. Discussion

In this study, the phylogenetic analysis of 16S rRNA sequences identified two lineages of H. pylori strains detected from different regions in Sudan which suggested differential evolution. This finding is in agreement with a number of studies that found an unusually high degree of genetic diversity in genomic sequence analyses of H. pylori strains in connection with their house-keeping and virulence associated genes. Strain17, Strain21 and Strain39 were derived from one lineage along with two strains from Hungary, with a bootstrap value of 61%. They shared a double base pair substitution (GC763-764CA) and one insertion mutation (751_InsC_752). However, both mutations were located in the central domain of the 16S rRNA gene. Despite Sudanese strains were from different ethnicity and regions of Sudan but they all caused antral gastritis, while the Hungarian strains (KF297892 and KF297893) were isolated from Demodex mites and caused rosacea.  

Laboratory diagnosis of H. pylori by conventional cultural methodology and phenotypic identification tests is still less specific, time consuming, increased capital costs, needs of highly skilled personnel and also it difficult to diagnose re-infections. However, molecular tests and PCR technology have been raised as a useful alternative means of bacterial identification, which may circumvent some of these difficulties but has limited mainly by contamination and inadequate sensitivity issues. In this study, molecular detection of H. pylori has shown that 37.33% of patients were positive for H. pylori in comparison to other studies conducted in different regions of Sudan, 22.2%, 59%, 65.8%, 40.1%, 48%, 21.8% by Mona et. al., Yousra et. al., Tajeldin et. al., Karimeldin et. al., Abdalsadeg et. al., and Mohammed et. al., respectively using different detection methods. 

Tetracycline is one of the 30S-targeting antibiotics that inhibit translation elongation by sterically interfering with the binding of aminoacyl-tRNA to the A-site of the ribosome. Therefore, mutations located in two domains (III and IV) in 16S rRNA: helix 34 and the loop next to helix 31 can affect the conformation of the tetracycline-binding site, leading to high-level resistance. In this study, six mutations were detected in two domains (I and II) which are in regions of the 16S rRNA not closely associated with either tetracycline or tRNA binding sites. This finding is partially in agreement with a study conducted by Catharine et. al., that found six nucleotide changes in two tetracycline-resistance strains. Two of these changes were located in domain I and domain II, G360A and deletion of G771, respectively. However, we observed that 66.67% of nucleotide variations were located in the central domain of 16S rRNA (nucleotides 567–915) with five associated ribosomal proteins (S15, S6 and S18) folds into the platform of the small subunit. Moreover, deleterious mutations (C18G, A55G, A161G, A373G, G521A, C614A, A622G, and deletion of one A in a triple-A cluster 607–609) in domains I and II which were suggested by Aymen et. al., in purposes of revealing covered putative
functional regions of the ribosome and aiding in the development of new antibiotics,\(^{(73)}\) were found conserved in our strains, as illustrated in Figure 2(b).

Interestingly, the insertion mutation (InsC), that was detected in sixty-seven percent (7/12) of our strains, between two nucleotides (750 and 751) located at the lower portion of helix 22 (H22), which is one of the lower three-helix junctions (3HJ) of 16S rRNA that bind with S15.\(^{(74, 75)}\) Also, the mutation in cancerous strain 101 (G → A transition at positions 741) was located in the upper portion of helix 22 (H22). The molecular dynamic (MD) simulations of Wen et. al., showed that 16S rRNA and S15 bind across the major groove of H22 via electrostatic interactions, i.e. the negatively charged phosphate groups of G658, U740, G741 and G742 bind to the positively charged S15 residues Lys7, Arg34 and Arg37.\(^{(74)}\) However, studying the effect of these mutations on the functions of 16S rRNA molecules in protein synthesis and antibiotic resistance is of great importance especially for essential regions like the central domain.\(^{(76)}\) For example, Prescott et. al., in 1990 found that the presence of substitution from a C to G at position 726 induces the synthesis of heat shock protein and affects the expression levels of various proteins.\(^{(77)}\) Also, mutations in the sequence 911-915 were conferred Streptomycin-resistance by impaired the binding of this antibiotic.\(^{(78-80)}\)

We acknowledge that limitations of the study are the small sample size and the phylogenetic tree was built based on the 16S rRNA gene only which is unable to cover more complex evolutionary events and distinguish between closely related strains or species.\(^{(14)}\) Therefore, further studies with a large sample size and building a phylogeny by increasing the number of genes analyzed like MLST.

**In conclusion**

The phylogenetic analysis of 16S rRNA sequences identified two lineages of \(H.\ pylori\) strains detected from different regions in Sudan. Sex mutations were detected in regions of the 16S rRNA not closely associated with either tetracycline or tRNA binding sites. 66.67% of them were located in the central domain of 16S rRNA. Studying the effect of these mutations on the functions of 16S rRNA molecules in protein synthesis and antibiotic resistance is of great importance.
Declarations

- **Ethical approval and consent to participate**

Approval to conduct this study was obtained from the Khartoum ministry of the health research department, University of Khartoum and Research Ethics Committees of hospitals. The patients gave written informed consent before they enrolled in the study.

- **Consent for publication**

All authors consent for publication

- **Availability of data and materials**

All data generated or analyzed during this study are included in this published article

- **Conflict of interest:**

The authors declare that there are no conflicts of interest.

- **Funding**

The authors received no specific funding for this work.

- **Authors' contributions**

MAH and EMI supervised the methodology and revised the manuscript. ABI, HGH, LBI, AMA and MMAI collected the samples. ABI, MASA and SME extracted the DNA. ABI, MASA and HNA amplified the 16S rRNA gene. ABI analyzed the data and wrote the manuscript. MAH edited and revised the final manuscript.

- **Acknowledgments**

We would like to thanks patients who participated in the study and the staff of the gastroscopic unit in Ibin Sina specialized hospital, Soba teaching hospital, Modern Medical Centre, and Al Faisal Specialized Hospital.
References

1. Linz B, Balloux F, Moodley Y, Manica A, Liu H, Roumagnac P, et al. An African origin for the intimate association between humans and Helicobacter pylori. Nature. 2007 February 22;445(7130): 915–8.
2. Thorell K, Yahara K. Rapid evolution of distinct Helicobacter pylori subpopulations in the Americas. 2017;13(2):e1006546.
3. Moodley Y, Linz B, Bond RP, Nieuwoudt M, Soodyall H, Schlebusch CM, et al. Age of the association between Helicobacter pylori and man. PLoS pathogens. 2012;8(5):e1002693-e.
4. Munoz-Ramirez ZY, Mendez-Tenorio A, Kato I, Bravo MM, Rizzato C, Thorell K, et al. Whole Genome Sequence and Phylogenetic Analysis Show Helicobacter pylori Strains from Latin America Have Followed a Unique Evolution Pathway. Frontiers in cellular and infection microbiology. 2017;7:50.
5. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003;164(4):1567-87.
6. Suzuki R, Shiota S, Yamaoka Y. Molecular epidemiology, population genetics, and pathogenic role of Helicobacter pylori. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2012;12(2):203-13.
7. Testerman TL, Morris J. Beyond the stomach: an updated view of Helicobacter pylori pathogenesis, diagnosis, and treatment. World J Gastroenterol. 2014;20(36):12781-808.
8. Schwarz S, Morelli G, Kusecek B, Manica A, Balloux F, Owen RJ, et al. Horizontal versus familial transmission of Helicobacter pylori. PLoS pathogens. 2008;4(10):e1000180-e.
9. Morelli G, Didelot X, Kusecek B, Schwarz S, Bahlawane C, Falush D, et al. Microevolution of Helicobacter pylori during prolonged infection of single hosts and within families. PLoS genetics. 2010;6(7):e1001036-e.
10. Achtman M, Azuma T, Berg DE, Ito Y, Morelli G, Pan ZJ, et al. Recombination and clonal groupings within Helicobacter pylori from different geographical regions. Molecular microbiology. 1999;32(3):459-70.
11. Ghose C, Perez-Perez GI, van Doorn LJ, Domínguez-Bello MG, Blaser MJ. High frequency of gastric colonization with multiple Helicobacter pylori strains in Venezuelan subjects. Journal of clinical microbiology. 2005;43(6):2635-41.
12. Kodaman N, Pazos A, Schneider BG, Piazuelo MB, Mera R, Sobota RS, et al. Human and Helicobacter pylori coevolution shapes the risk of gastric disease. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(4):1455-60.
13. MP L. Ethnologue: Languages of the World. Dallas, Texas SIL International; 2009.
14. Sudan Population (2019-05-11) [cited 2019 20/5]. Available from: http://worldpopulationreview.com/countries/sudan/
15. Factbook TW. Sudan 2019 [cited 2019 20/5]. Available from: https://www.cia.gov/library/publications/the-world-factbook/geos/su.html
16. Sharaf Eldin GS, Fadl-Elmula I, Ali MS, Ali AB, Salih AL, Mallard K, et al. Tuberculosis in Sudan: a study of Mycobacterium tuberculosis strain genotype and susceptibility to anti-tuberculosis drugs. BMC Infect Dis. 2011;11:219.
17. El-Sony AI KA, Enarson DA, Baraka O, Mustafa SA, Bjune G. Treatment results of DOTS in 1797 Sudanese tuberculosis patients with or without HIV co-infection. Int J Tuberc Lung Dis. 2002;6(12):1058-66.
18. Abdalsadeg N, Adam A, Abdul-Aziz H, Omer W, Osman H, Bolad A. Comparison of different diagnostic methodsof Helicobacter pylori infection in Sudanese patients. Al Neelain Med J. 2012;2(4):27-34.
19. Abdallah TM, Mohammed HB, Mohammed MH, Ali AAA. Sero-prevalence and factors associated with Helicobacter pylori infection in Eastern Sudan. Asian Pac J Trop Dis. 2014;4(2):115-9.
20. Sitaraman R. Allergies, Helicobacter pylori and the continental enigmas. Frontiers in Microbiology. 2015;6(578).
21. Galvani AP. Epidemiology meets evolutionary ecology. Trends in Ecology & Evolution. 2003;18(3):132-9.
22. Herre EA. Population structure and the evolution of virulence in nematode parasites of fig wasps. Science (New York, NY). 1993;259(5100):1442-5.
23. Stewart AD, Logsdon JM, Jr., Kelley SE. An empirical study of the evolution of virulence under both horizontal and vertical transmission. Evolution; international journal of organic evolution. 2005;59(4):730-9.
24. Latifi-Navid S, Ghorashi SA, Siavoshi F, Linz B, Massarrat S, Khegay T, et al. Ethnic and geographic differentiation of Helicobacter pylori within Iran. PloS one. 2010;5(3):e9645-e.
25. Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, et al. Traces of human migrations in Helicobacter pylori populations. Science (New York, NY). 2003;299(5612):1582-5.
26. Wirth T, Wang X, Linz B, Novick RP, Lum JK, Blaser M, et al. Distinguishing human ethnic groups by means of sequences from Helicobacter pylori: lessons from Ladakh. Proc Natl Acad Sci U S A. 2004;101(14):4746-51.
27. Moodley Y, Linz B, Yamaoka Y, Windsor HM, Breurec S, Wu JY, et al. The peopling of the Pacific from a bacterial perspective. Science (New York, NY). 2009;323(5913):527-30.
28. Moodley Y, Linz B. Helicobacter pylori Sequences Reflect Past Human Migrations. Genome dynamics. 2009;6:62-74.
29. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990;87(12):4576-9.
30. Van de Peer Y, Jansen J, De Rijk P, De Wachter R. Database on the structure of small ribosomal subunit RNA. Nucleic acids research. 1997;25(1):111-6.
31. Maidak BL, Olsen GJ, Larsen N, Overbeek R, McCaughey MJ, Woese CR. The Ribosomal Database Project (RDP). Nucleic Acids Res. 1996;24(1):82-5.
32. Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. Journal of clinical microbiology. 2000;38(10):3623-30.
33. Woese CR. Bacterial evolution. Microbiol Rev. 1987;51(2):221-71.
34. Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J Clin Microbiol. 2007;45(9):2761-4.
35. Patel JB. 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. Molecular diagnosis : a journal devoted to the understanding of human disease through the clinical application of molecular biology. 2001;6(4):313-21.
36. Tang YW, Ellis NM, Hopkins MK, Smith DH, Dodge DE, Persing DH. Comparison of phenotypic and genotypic techniques for identification of unusual aerobic pathogenic gram-negative bacilli. Journal of clinical microbiology. 1998;36(12):3674-9.
37. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. Journal of bacteriology. 1991;173(2):697-703.
38. EMRO. Health Systems Profile- Sudan. Cairo-Egypt: World Health Organization, 2006.
39. Gassoum A, A.Arbab M, Aldeaf SAH, Elhassan LA, Elshibli E, Elhassan AM. Allele Frequency Of P53 Gene Arg72Pro In Sudanese Meningioma Patients And Controls. IJSTR. 2014;3(6):243-8.
40. Liu W, Sun Z, Zhang J, Gao W, Wang W, Wu L, et al. Analysis of microbial composition in acid whey for dairy fan making in Yunnan by conventional method and 16S rRNA sequencing. Current microbiology. 2009;59(2):199-205.
41. Alfatih YM, Idris AB, Hassan HG, Nour EOM, Elhaj NMA, El-Zaki S-E, et al. Detection of a novel mutation G511T in the 530 loop in 16S rRNA in multi drugs resistant Mycobacterium tuberculosis isolated from Sudanese patients. bioRxiv. 2018:497628.
42. Liu W, Bao Q, Jirimutu, Qing M, Siriguleng, Chen X, et al. Isolation and identification of lactic acid bacteria from Tarag in Eastern Inner Mongolia of China by 16S rRNA sequences and DGGE analysis. Microbiol Res. 2012;167(2):110-5.
43. Idris AB, Ataelmanan AE, Eltaher SM, Idris EB, Osman Arbab BM, Ibn Idris A, et al. Independently Carriage of IL-1RN*2 Allele Associated with Increased Risk of Gastric Cancer in The Sudanese Population. medRxiv. 2019:19013573.
44. FinchTV. 1.4.0 ed. USA: Geospiza, Inc.; Seattle, WA,; 2012.
45. Altschul SF, Madden TL, Schäffer AA, J Zhang ZZ, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST. A new generation of protein database search programmes. Nucleic Acids Res. 1997;25(17):3389–402.
46. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, et al. Analysis Tool Web Services from the EMBL-EBI. Nucleic Acids Res. 2013;41(Web Server issue):W597-600.

47. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999;41:95-8.

48. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular biology and evolution. 2000;17(4):540-52.

49. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic biology. 2007;56(4):564-77.

50. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution. 1987;4(4):406-25.

51. Jukes TH, Cantor CR. Evolution of Protein Molecules. 1969:21-132.

52. Nei M, Kumar S. Molecular Evolution and Phylogenetics. USA: Oxford University Press; 2000.

53. Felsenstein J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. Evolution; international journal of organic evolution. 1985;39(4):783-91.

54. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular biology and evolution. 2016;33(7):1870-4.

55. Brosius J, Palmer ML, Kennedy PJ, Noller HF. Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America. 1978;75(10):4801-5.

56. Laith, Hassan A, Al-Obaidi L, M Jabbar A. Sequencing and Phylogenetic Analysis of Helicobacter pylori through 16S rRNA Gene Isolated from Gastritis Biopsies 2017.

57. Taylor DE, Eaton M, Chang N, Salama SM. Construction of a Helicobacter pylori genome map and demonstration of diversity at the genome level. Journal of bacteriology. 1992;174(21):6800-6.

58. Jarmuda S, O'Reilly N, Zaba R, Jakubowicz O, Szkaradkiewicz A, Kavanagh K. Potential role of Demodex mites and bacteria in the induction of rosacea. Journal of medical microbiology. 2012;61(Pt 11):1504-10.

59. Banerjee HN, Gramby M, Hawkins Z. Molecular Diagnosis of Helicobacter Pylori Strain by 16S rDNA PCR Amplification and Direct Sequencing. Journal of bioprocessing & biotechniques. 2011;1:1105e.

60. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proceedings of the National Academy of Sciences of the United States of America. 1985;82(20):6955-9.

61. Petti CA. Detection and identification of microorganisms by gene amplification and sequencing. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2007;44(8):1108-14.
62. Salih KMA, Elfaki OA, Hamid YHM, Eldouch WMA, Diab M, Abdelgadir SO. Prevalence of Helicobacter Pylori among Sudanese children admitted to a specialized children hospital. Sudan J Paediatr. 2017;17(1):14-8.
63. Abbas M, Sharif FA, Osman SM, Osman AM, El Sanousi SM, Magzoub M, et al. Prevalence and Associated Symptoms of Helicobacter pylori Infection among Schoolchildren in Kassala State, East of Sudan. Interdisciplinary Perspectives on Infectious Diseases. 2018;2018:5.
64. Mirghani YAA, Ahmed S, Ahmed M, Ismail MO, Fedail SS, Kamel M, et al. Detection of Helicobacter Pylori in Endoscopic Biopsies in Sudan. Tropical Doctor. 1994;24(4):161-3.
65. Mamoun M, M ES, Enan KA, Abdo AE, Hassan MA. Molecular Identification of 16s Ribosomal RNA Gene of Helicobacter pylori Isolated from Gastric Biopsies in Sudan. American Journal of Microbiological Research. 2015;3(2):50-4.
66. Wilson DN. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nature Reviews Microbiology. 2014;12(1):35-48.
67. Dailidiene D, Bertoli MT, Miciuleviciene J, Mukhopadhyay AK, Dailide G, Pascasio MA, et al. Emergence of Tetracycline Resistance in Helicobacter pylori: Multiple Mutational Changes in 16S Ribosomal DNA and Other Genetic Loci. Antimicrobial agents and chemotherapy. 2002;46(12):3940.
68. Trieber CA, Taylor DE. Mutations in the 16S rRNA genes of Helicobacter pylori mediate resistance to tetracycline. Journal of bacteriology. 2002;184(8):2131-40.
69. Xu Z, Culver GM. Differential assembly of 16S rRNA domains during 30S subunit formation. RNA. 2010;16(10):1990-2001.
70. Batey RT, Williamson JR. Interaction of the Bacillus stearothermophilus ribosomal protein S15 with 16S rRNA: I. Defining the minimal RNA site. Journal of molecular biology. 1996;261(4):536-49.
71. Agalarov SC, Williamson JR. A hierarchy of RNA subdomains in assembly of the central domain of the 30S ribosomal subunit. RNA. 2000;6(3):402-8.
72. Agalarov SC, Zheleznyakova EN, Selivanova OM, Zheleznaya LA, Matvienko NI, Vasiliev VD, et al. In vitro assembly of a ribonucleoprotein particle corresponding to the platform domain of the 30S ribosomal subunit. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(3):999-1003.
73. Yassin A, Fredrick K, Mankin AS. Deleterious mutations in small subunit ribosomal RNA identify functional sites and potential targets for antibiotics. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(46):16620-5.
74. Li W, Ma B, Shapiro B. Binding interactions between the core central domain of 16S rRNA and the ribosomal protein S15 determined by molecular dynamics simulations. Nucleic acids research. 2003;31:629-38.
75. Agalarov SC, Sridhar Prasad G, Funke PM, Stout CD, Williamson JR. Structure of the S15,S6,S18-rRNA complex: assembly of the 30S ribosome central domain. Science (New York, NY). 2000;288(5463):107-13.
76. Brink MF, Nels RN, Verbeet MP, de Boer HA. Specialized ribosomes allow for the study of mutations in functionally important regions in 16 S rRNA, without affecting cell growth. The identification of functional regions in the central domain of 16S rRNA. Journal of molecular biology. 1994;237(4):368-77.

77. Prescott CD, Dahlberg AE. A single base change at 726 in 16S rRNA radically alters the pattern of proteins synthesized in vivo. EMBO J. 1990;9(1):289-94.

78. Montandon PE, Wagner R, Stutz E. E. coli ribosomes with a C912 to U base change in the 16S rRNA are streptomycin resistant. EMBO J. 1986;5(13):3705-8.

79. Leclerc D, Melançon P, Brakier-Gingras L. Mutations in the 915 region of Escherichia coli 16S ribosomal RNA reduce the binding of streptomycin to the ribosome. Nucleic acids research. 1991;19(14):3973-7.

80. Frattali AL, Flynn MK, De Stasio EA, Dahlberg AE. Effects of mutagenesis of C912 in the streptomycin binding region of Escherichia coli 16S ribosomal RNA. Biochimica et biophysica acta. 1990;1050(1-3):27-33.
Figure 1. 1a) and b). PCR amplification results of universal and specific H. pylori 16S rRNA gene, respectively examined on 2% agarose gel electrophoresis. 1c). Sequencing result of Acinetobacter radioresistens chromatogram using Finch TV software. The nucleotide sequence was deposited in the GenBank database under the accession number: MN845952.
Figure 2. 2a) and 2c) Sequencing results of chromatograms using Finch TV software show nucleotide variations in the 16S rRNA gene of H. pylori which illustrated by squares. 2b) Multiple Sequence Alignment (MSA) of 16S rRNA sequences of 12 Sudanese H. pylori strains compared with the rRNA gene of H. pylori strain 26695 (NC_000915) and other selected strains obtained from GenBank databases using Clustal W2.
Figure 3. Distribution and evolutionary relationships of Sudanese H. pylori strains. 3a) Map of regional origin of strains in Sudan. 3b) The Neighbor-Joining Phylogenetic tree. The percentage of replicate trees (1000 replicates) are shown next to the branches. The evolutionary distance was computed using (JC) method and are in the units of the number of base substitution per site. Evolutionary analyses were conducted using MEGA7.