SciMedCentral

JSM Gastroenterology and Hepatology

Research Article

Characterization of East Asian and African Genotypes among *Helicobacter pylori* Strains Isolated from Different Human Populations

Guillermo I. Perez-Perez[1]*, Malin Andersson[2], and Asalia Z. Olivares[1]

[1]*Department of Medicine and Microbiology, New York University Medical Center, USA*
[2]*Deparment of Science, University of Goterborg, Sweden*

***Corresponding author**
Guillermo I. Perez Perez, 6027W VAMC, 423 East 23d. Street, New York, NY 10010, USA, Tel: 212-263-4101; Fax: 212-263-4108; Email: perezg02@med.nyu.edu

Submitted: 06 February 2018

Accepted: 13 February 2018

Published: 17 February 2018

Keywords

- *H. pylori*
- HspA
- cagPAI
- Geographic and ethnic distribution

Abstract

Background: *Helicobacter pylori,* a genetically diverse bacterium has shown to have geographic specific genotypes. This may be relevant since there are also differences in clinical outcome among *H. pylori* infected patients from different geographical origin.

Aim: The aim of this study was to identify a true marker for African origin based on the sequence of a housekeeping gene (*hspA*) and if African origin could be defined, to establish possible differences in virulence attributes between African and non-African strains.

Material and methods: The *hspA* gene was analyzed by PCR amplification and direct sequencing of the PCR products in 163 strains from patients with diverse ethnic origin. To asses for virulence, the *cag* pathogenic island (*cag* PAI) was analyzed by three different kinds of PCR amplifications in the same 163 strains.

Results: The *hspA* gene showed to be very informative for differentiating strains of diverse ethnic origin. In particular, different phenotypes defining African and East Asian origin were found when amino acid sequences were compared. In addition, a marker for East Asian origin was found upstream of the *hspA* gene. Based on the presence of *cag* PAI, African American strains together with East Asian strains were shown to be the most virulent. East Asian strains amplified lower molecular weight PCR products than African American strains at the *cag* PAI level. Whether this is relevant for the functionality of the *cag* PAI should be further explored.

Conclusion: Highly specific markers that define *H. pylori* strains of African and East-Asian origins were identified. However, no differences in virulence attributes were observed between those strains.

INTRODUCTION

Background

Helicobacter pylori are a gram-negative bacterium that colonizes the gastric mucosa. It is considered to be one of the most widespread human infections worldwide [1]. Unless treated with antibiotics, the infection will persist for decades or even throughout the lifetime of the host [2,3]. Persistent infection is strongly associated with gastric adenocarcinoma and peptic ulcer disease [3-7].

There are differences in clinical outcome between patients from different geographical origin. The incidence of gastric adenocarcinoma is for example much higher in East Asia than in Central Africa even though the incidence of *H. pylori* infection is high in both areas. This paradox is named the African enigma. The phenomenon may be explained because of genetic differences among human populations, environmental factors and/or strain differences of *H. pylori. H. pylori* are a genetically diverse bacterium, but, geographic specific genotypes have been described. Therefore it may be possible that strain differences in the bacteria are involved in different clinical outcomes [8,9].

The genetic fingerprints of *H. pylori* isolated from unrelated individuals are different. However, bacteria are mainly transmitted between individuals in one family, isolates from patients within the same family often has similar genetic fingerprinting. The fact that *H. pylori* are mainly transmitted within a family suggests vertical transmission and probably the same strain persists throughout generations. This indicates that the evolution of *H. pylori* is linked to the social behavior of humans. Since humans have been living in isolated communities until very recently, genetic exchange has been limited and genetic traits became segregated. It is likely that a segregation of the genes of *H. pylori* happened along with this, hence the geographic specific genotypes of the bacteria [3]. The genotype differences

may explain differences in virulence of the bacteria and thereby the differences in clinical outcome [9].

Geographic specific genotypes and markers have been found in several sites in the *H. pylori* genome. One example is an insert of 180bp that is more common in strains of *H. pylori* of African origin than in strains of other origin [10]. Another example is the *vacA* gene, which codes for a cytotoxin that induces vacuolization in epithelial cells. The toxin gene contains three variable portions; the s-region, the i-region, and m-region. The s-region codes for a signal peptide and exist as s1 or s2 allelic types. The s1 type has three subtypes; s1a, s1b and s1c. These subtypes have been shown to be unevenly distributed throughout the world and have shown ethnic specificity [11,12].

When the sequence of fragments of *vacA* was analyzed together with fragments of seven housekeeping genes it was shown that *H. pylori* can be divided into seven populations and subpopulations with distinct geographical distribution on the basis of those sequences. The populations all arose from populations in Africa, East Asia and Central Asia [13].

A housekeeping gene that has shown signs of geographic specific variability is *hspA* gene [14-16]. It is present in all *H. pylori* strains and encodes the protein HspA, a homologue of the heat shock proteins of the GroES class in *Escherichia coli*. There is a high degree of similarity between HspA and GroES when looking at the first 90 amino acids (the A domain). However, it has a unique *C*-terminus of 27 amino acids (the B domain) that GroES proteins lack. The HspA B domain is the area in which *H. pylori* shows geographic specific variability in amino acid sequence between strains. [14-17]. Evidence of geographic specific genotypes within this gene is seen in phylogenetic trees based on the nucleotide sequence of strains originating from different parts of the world. Both African and East Asian branches have been found in such analyses [14-15].

CagA (cytotoxin-associated gene A), an approximately 128 kDa protein encoded by the *cagA* gene, is a virulence factor of *H. pylori* [18]. The *cagA* gene is located at the end of the *cag* pathogenicity island (*cag* PAI), which is a 40-kb DNA segment that contains a total of 31 genes. Some of these genes encode for proteins that are components of a bacterial type IV secretion system that delivers bacterial proteins into eukaryotic cells. CagA is delivered to the cells of the human gastric epithelia by this system [9,19-21].

Strains of *H. pylori* can be divided into *cagA* positive and *cagA* negative strains using *cagA* as a marker for the whole *cag* PAI. *cagA* positive strains are associated with an increased risk of developing gastric adenocarcinoma whereas infection with *cagA* negative strains not [22-25].

Thus, the aims of this study were to try to find a true marker for African origin and, if African origin can be defined, to establish difference in virulence attributes between African and non-African strains. Because of the variability and geographic specific genotypes in all *H. pylori* strains, the *hspA* gene was analyzed to find a true marker for African origin. In addition, because of its strong association with virulence, the *cag* PAI was analyzed in this study to establish difference in virulence attributes between African and non-African strains [26-29].

MATERIALS AND METHODS

Bacterial strains

A total of 192 isolates of *H. pylori* were studied. The isolates where obtained from patients with diverse ethnic or geographic origin that suffered from *H. pylori* associated diseases. The ethnic origins of the patients included East Asian, African-American, Caucasian, Hispanic, and Amerindian. After the exclusion of a total of 29 strains (see below) the final ethnic distribution is presented in Table 1. The patients defined themselves as belonging to the different ethnical groups. The East Asians were patients with origin from Japan, China, Vietnam, Hong Kong, Burma and Malaysia. The African-Americans were patients living in North America, South America or the Caribbean with sub-Sahara African roots. Caucasians were patients of European or Middle East origin. The Hispanics were individuals who speak Spanish as their native language. Amerindians are Native Americans in Northern, Central and South America who probably mix with Hispanics, Caucasians or African-Americans.

Analysis of the hspA gene and its encoded protein

***hspA* PCR and sequencing:** Purified DNA from the 197 strains that had previously been extracted was used to analyze the *hspA* nucleotide sequence. A 487-bp segment containing the 354-bp *hspA* gene was amplified by polymerase chain reaction (PCR) (Figure 1, panel A) (Table 2).

Gel electrophoresis on a 2% agarose gel was carried out for all samples to confirm that the PCR amplification was successful. Positive and negative controls were used. If the amplification was not successful the first time, one additional amplification was performed before the samples were considered defect and thus unusable. The PCR products from the strains for which the PCR amplification was successful (n=186) were purified with the Qiagen QIAquick DNA purification kit (Qiagen, Valencia, CA) and were submitted for direct sequencing (SeqWright DNA Technology Services).

The same primers as for the *hspA* PCR were used to sequence both strands of the PCR product.

Sequence analysis: The resulting nucleotide sequences from all the strains were edited using SEQUENCHER (version 4.7). The nucleotide sequences were aligned using CLUSTALW. A phylogenetic tree of the *hspA* gene of 354bp was created by the neighbor-joining method (28) with MEGA (version 4.1). The stability of the tree was tested by performing 500 bootstrap replicates. The nucleotide sequences were translated into amino acid sequences using GCG (Genetics computer group; Madison, WI). The amino acid sequences were aligned using CLUSTALW.

Analysis of the cag PAI

Empty site PCR: In order to determine the absence of the *cag* PAI a 550bp segment containing the empty site was amplified by PCR in the 163 of the 186 strains included in this study. The *cag* PAI is located between the two protein coding *H. pylori* genes Hp0519 and Hp0549, (Figure 1, panel B) (Table 2). Gel electrophoresis on a 2% agarose gel was carried out for all samples to control which samples contained the empty site. Positive and negative controls were used.

Table 1: Ethnicity distribution and age and gender of patients included in the study.

Ethnicity	n	%	Subjects with gender and age available	Mean age (years)	Ratio F:M
East Asian	24	14.3	19	54.5	10:09
Amerindian	28	16.7	26	40.3	16:10
Hispanic	67	39.9	65	54.2	40:25:00
Caucasian	17	10.1	15	54.5	6:11
African American	27	16.1	21	60.2	1:20

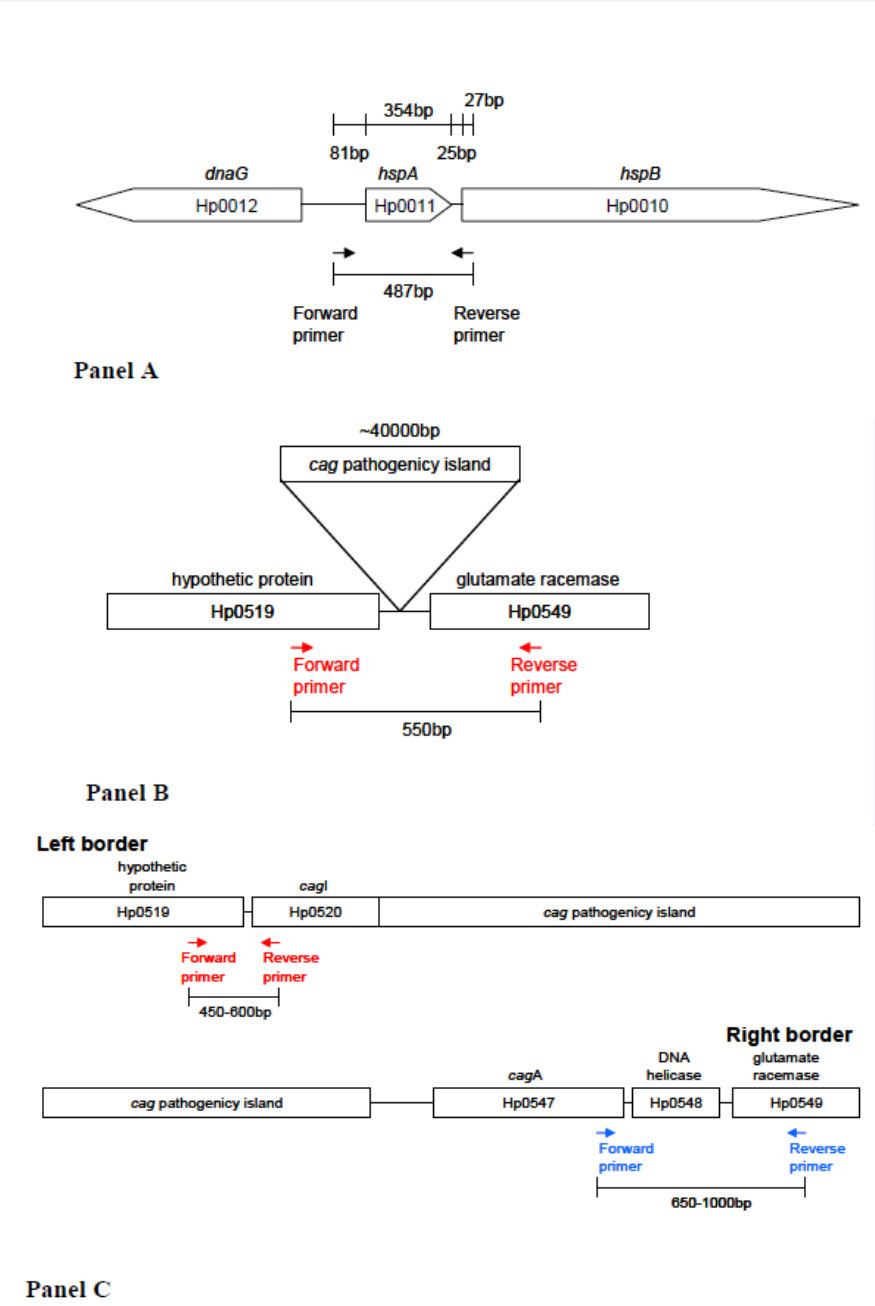

Figure 1 Schematic representation of the target genes selected for molecular analysis and/or sequencing. Panel A, schematic representation of loci *hspA*, based on *H. pylori* strain J-99. Panel B, Schematic representation of the loci used to analyse the empty site, based on *H. pylori* strain 26695. Panel C, Schematic representation of the loci used for analysing the left and right border of the *cag* pathogenicity island from *H. pylori* strain 26695 (Courtesy of Dr Steffen Backert).

Analysis of the left and right border of the *cag* PAI: Left and right border PCR can be used to determine which strains that are *cag* PAI positive. This means that only *cag* PAI positive strains will amplify a product with left and right border PCR. The approximate sizes of the left and right border can also be investigated with this technique (Figure 1, panel C) (Table 2).

Left border PCR: In the strains that did not yield the 550bp product that represents the empty site (n=122), the left border of *cag* PAI was further analyzed by left border PCR. Gel electrophoresis on a 2% agarose gel was carried out for all samples to control that the PCR amplification was successful and to see what approximate sizes the amplified segments had. Positive and negative controls were used.

Right border PCR: In the strains that did not yield the 550bp product that represents the empty site (n=122), the right side of *cag* PAI was further analysed by right border PCR.

Gel electrophoresis on a 2% agarose gel was carried out for all samples to control that the PCR amplification was successful and to see what approximate sizes the amplified segments had. Positive and negative controls were used.

Intergenic region between *jhp0153* and *jhp0152*

A PCR reaction amplified an intergenic region from the *H. pylori* genome as previously reported [10]. Based on the product size amplified, we could identify *H. pylori* strains of African origin by the distinctive 180bp insert that they carry in this intergenic region (Table 2) [10].

Statistics

The chi-square test was carried out with Epi-Info software (version 3.3.2), when appropriate. Differences were defined as being statistically significant if the *P* value was less than 0.05.

RESULTS

Analysis of the hspA gene and its encoded protein

***hspA* PCR:** Amplification was successful the first time for all but 14 (8.8%) samples and the second time for all but 6 samples, ending up with a total of 3.1% failed amplifications and 186 successfully amplified strains. All strains amplified a product of the same size (~487bp).

***hspA* nucleotide sequences:** Direct DNA sequencing was successful for 174 (93.5%) of the 186 strains. After the sequence analysis 11 strains were excluded, one strain was excluded for lack of ethnicity criteria, 9 were excluded for being replicates of the same patient, and one 1 was excluded for being the same strains previously included. The nucleotide sequence of the hspA gene was analyzed in 163 strains. The phylogenetic tree derived from the *hspA* gene of 354bp of the 163 strains suggested geographic segregation of *H. pylori* (Figure 2). Out of the 24 East Asian strains, 19 (83.3%) belonged to the same branch in the tree (East Asian branch). Out of the 27 African-American strains, 17 (62.9%) were part of another branch (African branch). However, the East Asian branch and the African branch were not supported by a significant bootstrap number (Figure 2).

Comparison of HspA amino acid sequences: Next, amino acid sequence of HspA was analyzed for 163 strains. The amino acid sequences showed high degree of conservation in domain A, except for amino acid 90.

Variability in this residue was seen in 36 strains. Besides the variation in residue 90, 40 strains had some other amino acid substitution in domain A. Forty-three strains exhibited a single amino acid substitution and two strains had two substitutions. The other 123 strains were exactly identical in the first 89 amino acids (Data not shown).

There was a greater variation in domain B, confined to amino acid residues 91 and 97-99. Three major phenotypes can be seen when analyzing residue 91-100 of domain B together with residue 90 of domain A (Table 3).

One group had Serine instead of Glycine in residue 90, and Asparagine/Serine or Aspartic acid/Serine instead of Glycine/Asparagine in residues 98-99 or both when compared to the reference phenotype (Table 3). Because these phenotypes exist in almost all strains of East Asian origin (91.7%) the group was designated as the East Asian group and the phenotypes are collectively referred to East Asian phenotypes. A significantly higher percentage of the East Asian strains have these phenotypes than strains from any other ethnic origin (Table 4). The East Asian phenotypes are present among strains from all ethnic groups except African-American, where none of the strains have these phenotypes.

The second group had either Alanine instead of Serine in residue 91, and/or Alanine/Asparagine/Serine instead of Threonine/Glycine/Asparagine in residues 97-99 or both when compared to the reference phenotype (Table 3). Because these phenotypes exist in most strains of African origin (74%) the group was designated the African group and the phenotypes are collectively referred as African phenotypes. A significantly higher percentage of the African-American strains have these phenotypes than strains from any other origin (Table 5). The African phenotypes are present among African-American, Hispanic, Caucasian and Amerindian strains but not in the East Asian strains.

The third group does not seem to have anything in common with the African or East Asian phenotypes and it is referred as the non-African/ non East Asian group. The reference phenotype is included in this group. The reference group it is the most abundant phenotype and it is similar in amino acid sequence in the 90 to 99 residues to the reference or prototype strains of previous studies (14-16).

Comparison of amino acid sequences upstream of the *hspA* gene: The amino acid sequence in the inter-genetic region between the *dnaG* gene and the *hspA* gene was also found to vary. Two phenotypes can be clearly identified. One group has Arginine/Isoleucine instead of Isoleucine/Threonine at residues 18-19 before the starting codon of the *hspA* gene, when compared to the reference phenotype (Table 6). Because this phenotype exists in almost all strains of East Asian origin (87.5%) the group was designated the East Asian group and the phenotypes are collectively referred to as East Asian phenotype. A significantly higher percentage of the East Asian strains have this phenotype than strains from any other ethnic origin (Data not shown). As

Figure 2 Phylogenetic tree of *hspA* nucleotide sequences of 163 strains of *H. pylori* of different self-defined ethnical origins. Strains marked with pink are from Amerindian origin, strains marked with green are from Hispanic origin, strains marked with red are African American origin, strains marked with yellow are East Asian origin and strains marked with blue are Caucasian origin.

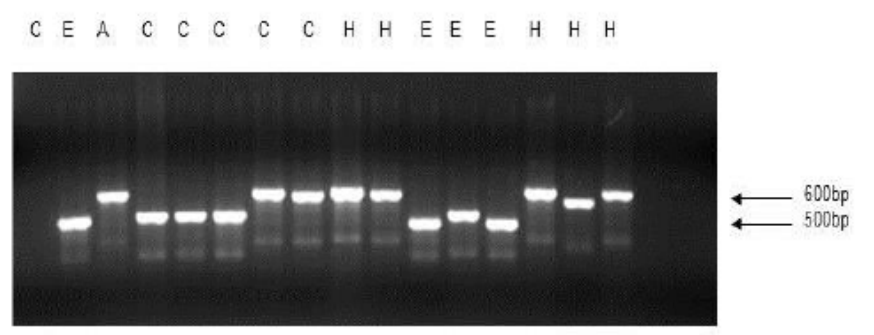

Figure 3 Visualization of the left border of the cag PAI by PCR products. 2% agarose gel electrophoresis Line C is the negative control H2O. Lanes E represented H. pylori strains of East Asian origin, lanes C represented Caucasian origin and H lanes represented Hispanic origin.

previously observed, the East Asian phenotype was present in strains from all ethnic groups except African-American, where none of the strains has this phenotype.

The second group was made out of the strains not having the Arginine/isoleucine substitution. It is designated the non-East Asian group. The reference phenotype is included in this group. An analysis of the strains that had both East Asian phenotypes upstream the *hspA* gene and between residues 90-100 in HspA was carried out. A significantly higher percentage of the East Asian strains (87.5%) have both East Asian phenotypes than strains of any other ethnic group (Table 7). All East-Asian strains had the East Asian phenotypes in the upstream region and in HspA. Very few of the strains of Amerindian, Hispanic and Caucasian origin had the East Asian phenotype in both places.

Analysis of the *cag* PAI

Empty site PCR: Amplification was successful in 41 (25.1%) of the 163 strains, meaning that those strains where empty site positive and thereby *cag* PAI negative. All successfully amplified strains amplified a product of the same size (~550bp) (Data not

Table 2: Primers used for PCR and qPCR.

Gene target	Primers	Temperature and step
hspA	5'GCTATCTGAAAATTTGATTTCTTTTGC-3'	94°C 30s, 52°C 30s, 72°C 2min.
	5'-TGCGCTATAGTTGTGTCGC-3'	
cagPAI	5'-CTCTTTTTGTGCCTTTGATTGAA-3'	94°C 30s, 54°C 45s, 72°C 1min.
Empty-site	5'-CCAAATACATTTTGGCTAAATAAAC-3'	
Cag PAI	5'-CCAAATACATTTTGGTAAATAAAC	94°C 1 min, 57°C 1 min, 72°C 1.5 min
Left border	5´-GCTTATCAGTCAAATTGTTTTTG	
Cag PAI	5´-GGCTCAAGCTCGTGAATGAT	94°C 1 min, 60°C 1 min, 72°C 1.5 min
Right border	5'-CTCTTTTTGTGCCTTTGATTGAA	
Intergenic region	5'-GTGGCGCGTTTCTTGCAATACC-3'	94°C 30s, 57°C 30s, 72°C 1min.
	5'-AACTCGCTCAAAAACTCGGC-3'	

Table 3: HspA geographic-specific phenotypes determined by amino acid sequence in 163 *H. pylori* strains isolated from patients with different ethnic origin.

		Distribution by sequence type					
Designation	Amino Acid sequence	Total	EA	AA	His	Cauc	Afri
		163	24	28	67	17	27
Reference[a]	^{90}GSGSCCHTGNH^{100}	63	1	9	35	11	7
1[b]	^{90}S - - - - - - - - NS - ^{100}	27	8	12	6	1	
2[b]	^{90} - - - - - - - - - NS - ^{100}	11	10	1			
3[b]	^{90}S - - - - - - - - N - - ^{100}	1	1				
4[b]	^{90}S - - - - - - - - DS - ^{100}	1	1				
5[b]	^{90} - - - - - - - - - DS - ^{100}	1	1				
6[b]	^{90} - - - - - - - - - D - - ^{100}	1	1				
7[b]	^{90}S - - - - - - - - - - - ^{100}	3			3		
8[c]	^{90}SA - - - - - - ANS-^{100}	1					1
9[c]	^{90} - - - - - - - - ANS - ^{100}	3			3		
10[c]	^{90} -A - - - - - -ANS- ^{100}	5		1	2		2
11[c]	^{90} - - - - - - - - A -S - ^{100}	1		1			
12[c]	^{90} - A- - - - - - - - - - ^{100}	37		4	16	4	13
13[c]	^{90} -A- - - - - - - -S - - ^{100}	1					1
14[c]	^{90} -A - - - - - - - D - ^{100}	2			1		1
15[a]	^{90} - - - - - - - - - - S- ^{100}	1				1	
16[a]	^{90} - A - - - - - - I - - - ^{100}	2			1		1
17[a]	^{90} - - - - - - - - - E - - ^{100}	1	1				
18[a]	^{90} DE - - - - - - - - - - ^{100}	1					1

[a]Non African/ non East Asian. *H. pylori* strains J99 and 26695 belongs to this group, having the reference phenotype.
[b]East Asian (EA)
[c]African (Afri)

Table 4: Comparative prevalence of East Asian phenotype between residues 90-100 in the HspA protein from *H. pylori* strains isolated from 163 patients, by ethnicity.

Ethnicity	n	% East Asian	OR	CI	P-value
East Asian	24	91.7	1.00		
Amerindians	28	46.4	0.08	0.01-0.46	0.002
Hispanic	67	13.4	0.01	0.00-0.08	<0.001
Caucasian	17	5.9	0.01	0.00-0.09	<0.001
African American	27	0.0	0.00	0.00-0.03	0.001

Table 5: Comparative prevalence of African phenotype between residues 90-100 in the HspA protein from *H. pylori* strains isolated from 163 patients, by ethnicity.

Ethnicity	n	%	OR	CI	P-value
African American	27	74.1	1.0		
Caucasian	17	23.5	0.2	0.03-0.72	0.013
Hispanic	67	32.8	0.2	0.08-0.69	0.006
Amerindians	28	21.4	0.1	0.03-0.52	0.002
East Asian	24	0.0	0.0	0.00-0.13	<0.001

Table 6: Geographic-specific phenotype determine by amino-acid residues left of the HspA protein in 163 *H. pylori* strains from patients with different ethnic origin.

Designation	Amino acid sequence	Distribution by sequence type					
		Total n=168	EA n=24	AA n=28	His n=67	Cauc n=17	Afri n=27
Reference[a]	LITISAKFLFYLSKLRRTEM	101	2	14	48	13	24
1c	- RI- - - - - - - - - - - - - - - -	29	8	4	16	1	
2c	- RI- - -N - - - - - - - - - - - -	4	2		2		
3c	- RI- - -N - - - - - - - - - - -N -	7	6	1			
4c	- RI- - - - - - - - - - - - - - - N -	4	2	1		1	
5c	- RI- -TN - - - - - - - - - - - - -	1		1			
6c	- RI- - -N - - - - - - - - - - I - -	1	1				
7c	- RI- - -I - - - - - - - - - - - N -	1				1	
8c	- RI- - - - - - - - - - - - - - - S -	1	1				
9c	- RI- - - - - P - - - - - - - - -N -	4		4			
10c	- RI- - T - - - - - - - - - - - -N -	1	1				
11b	- - - - - - - - - - - - - - - - - -N -	5	1	1	1	1	1
12b	- T - - - - - - - - - - - - - - - - -	1					1
13b	- R - - - - - - - - - - - - - - - - -	1					1
14b	- - - - - T- - - - - - - - - - - - -	2		2			

aNon East Asian. *H. pylori* strain J99 belongs to this group, having the reference phenotype.

Table 7: Comparative prevalence having both East Asian phenotypes in of *H. pylori* strains isolated from 163 patients by ethnic origin.

Ethnicity	n	% with both phenotypes	OR	CI 95%	P-value
East Asian	24	87.5	1.00		
Amerindian	28	10.7	0.02	0.00-0.12	<0.001
Hispanic	67	9.0	0.01	0.00-0.07	<0.001
Caucasian	17	5.9	0.01	0.00-0.11	<0.001
African American	27	0.0	Undef.	0.00-0.04	<0.001

Table 8: Comparative prevalence of cag PAI negative strains of *H. pylori* isolated from 163 patients, by ethnicity.

Ethnicity	n	% CagPAI negative	OR	CI 95%	P-value
Hispanic	67	37.3	1.00		
Caucasian	17	41.2	1.18	0.35-3.39	ns
Amerindians	28	21.4	0.46	0.14-1.41	ns
African American	27	11.1	0.21	0.05-0.84	0.024
East Asian	24	0.0	0.00	0.00-0.38	0.001

Table 9: Comparative prevalence of low molecular weight PCR products in cag PAI positive strains of *H. pylori*, by ethnicity.

	Left border (n=105)			Right border (n=111)		
Ethnicity	n	% with low mol. Weight*	P-value	n	% with low mol. Weight*	n
East Asian	18	94.4		22	90.9	0.001
Caucasian	9	33.3	0.002	10	30.0	<0.001
Hispanic	36	16.7	<0.001	38	5.3	<0.001
Amerindian	19	10.5	<0.001	19	10.5	<0.001
African-American	23	8.7	<0.001	22	9.1	<0.001

*Low molecular weight <500bp for the left and <700bp for the right

Table 10: Prevalence of East Asian markers in H. pylori strains from different ethnic groups.

Ethnicity		**% of strains with:**				
	n	East Asian HspA type	East Asian Upstream type	Absence of 180bp	Low MW *cag PAI* left side	Low MW *cag PAI* right side
East Asian	24	91.7	83.3	100.0	70.8	83.3
Amerindian	28	46.4	39.2	60.7	7.1	7.1
Hispanic	67	13.4	26.9	77.6	9.0	3.0
Caucasian	17	5.9	17.6	70.6	17.6	17.6
African American	27	0.0	0.0	48.1	7.4	7.4

Table 11: Number of East Asian characteristics in the different *H. pylori* strains, by ethnicity.

Ethnicity	**n**	**Number of strains with:**					
		5 types	4 types	3 types	2 types	1 types	0 types
East Asian	24	13	7	2	2	0	0
Amerindian	28	1	0	8	5	5	9
Hispanic	67	0	0	5	18	36	8
Caucasian	17	1	0	0	5	7	4
African American	27	0	0	2	0	11	14

shown). The Caucasian, Hispanic and Amerindian ethnic groups had the highest percentage of empty site presence, meaning that they had the highest percentage *cag* PAI negative strains (Table 8). East Asian and African American ethnic groups had the lowest percentage *cag* PAI negative strains. There was no significant difference between East Asian and African American strains in *cag* PAI negativity.

Left border PCR: Of the 122 strains that did not amplify an empty site PCR product, 105 (86.1%) were successfully amplified with the left border PCR. Products of four different sizes were amplified; 450bp, 500bp, 550bp and 600bp (Figure 3). Product less than or equal to 500bp in size was considered being of low molecular weight. The East Asian strains had the low molecular weight PCR product significantly more often than strains from any other ethnic group (Table 9).

Right border PCR: Of the 122 strains that did not amplify with empty site PCR, 111 (91.0%) were successfully amplified with the right border PCR. Of the 11 that did not amplify, 4 did not amplify with left border PCR either. Products of four different sizes amplified; 650bp, 700bp, 800bp and 1000bp (Data not shown). Products less than or equal to 700bp in size was considered being of low molecular weight. East Asian strains had low molecular weight PCR product significantly more often than strains from any other ethnic group (Table 9).

Correlation between East Asian origin and East Asian characteristics

As previously reported, an insert of 180bp is more often present in *H. pylori* strains of African origin than in strains from other ethnic origins [10]. Information about which strains have the insert was obtained. The characteristics that seem to be typical of an East Asian population are the presence of the East Asian phenotypes in HspA and upstream of the *hspA* gene, the lack of the 180bp insert and the presence of low molecular weight PCR products of the left and right border of the *cag* PAI. Prevalence of each characteristic in the different ethnic groups can be seen in Table 10. The concordance in all characteristics in the different ethnic groups can be seen in Table 11. Strains that shared four or more of the East Asian characteristics can be considered true East Asian. The strains of self-defined East Asian origin was significantly more often true East Asian than strains from any other ethnic origin (P<0.001).

DISCUSSION

H. pylori strains were examined and our findings have some

limitations which need to be pointed out. The strains were obtained from symptomatic patients and thus, the results are representative only for symptomatic patients and not the entire population with *H. pylori*. The samples were obtained from several laboratories around the world which makes it possible that differences in methods used affect the results. The fact that no reliable method was used to define the ethnic group in the patients studied and was based of self-defined ethnicity makes it possible that some of the patients were placed in the wrong group. The gender and age data was also limited. The data available showed that the mean age was relatively even throughout the groups although, based on the mean age, Amerindians was the youngest group and the African Americans the oldest group when compared with other groups. The proportion of females to males was relatively even in general. However African American males outnumber the females. This is because samples from these patients were mainly obtained at a hospital for US veterans in New York, where male patients greatly outnumber the female patients.

Strains that did not amplify a PCR product were run several times, thus it is unlikely that amplification failure is due to technical difficulties. Also, positive controls showed in all cases PCR products followed by the gel electrophoresis visualization. Therefore, poor quality of DNA from the failure samples could be the main reason for failure. At the end, the percentage of failed amplifications was small (3.1%).

It is important indicated that branches in the phylogenetic tree were not supported by a significant bootstrap number, suggesting that the tree has a low confidence level. However, East Asian and African branches have been seen before in phylogenetic trees of the *hspA* gene [14-15,30].

Conservation in the first 89 amino acids of HspA and the variability confined to residue 90 of domain A and 91 together with 97-99 of domain B is in line with previous findings [14-16]. Some of the phenotypes belonging to the East Asian group have previously been reported in strains of *H. pylori* isolated from patients in Hong Kong [14]. In another study, both the East Asian phenotypes and the African phenotypes have been reported based on the East Asian sequences and African sequences respectively [15]. The phenotypes upstream of the *hspA* gene is new and may be used as a marker for East Asian origin since it correlates strongly with East Asian origin and it is present significantly more often in East Asian strains than in strains form any other origin.

East Asian phenotypes both in HspA and upstream of the *hspA* gene combined, it seems to be an extremely good marker for East Asian origin of *H. pylori* strains.

The strong correlation between East Asian origin and the East Asian phenotypes and the fact that were present significantly more often in East Asian strains than in strains of other origin suggests that they can be used as markers for East Asian origin. The correlation between African-American origin and African phenotypes is weak but there is a significant difference between African-American origins suggesting that they can be used as markers for African origin.

The fact that Amerindian strains have the East Asian phenotypes relatively often may be explained by the fact that human migrations took place from East Asia region colonizing North and South America with this East Asian population of *H. pylori* until the arrival of the European colonization [12,13,30].

Hispanic and Caucasian strains also have the East Asian phenotypes indicating possible mix with East Asian population. Similar mixing may explain why in some cases Hispanic, Caucasian and Amerindian strains, carried the African phenotype. None of the East Asian strains have the African phenotypes and none of the African-American strains have the East Asian phenotypes, indicating that no mix has occurred between those groups.

The fact that Amerindian strains have the East Asian phenotype relatively often and that Hispanic and Caucasian strains may have the same explanations as above. Strains that did not amplify a product with empty site PCR but that did amplify with left and/or right border PCR were definitely *cag* PAI positive. The four samples that failed to amplify with empty site PCR, left border PCR and right border PCR were probably of poor quality, hence the failed amplifications.

Caucasian, Hispanic and Amerindian ethnic groups have the highest percentage of *cag* PAI negative strains. In contrast, African strains together with East Asian strains are the most virulent based on the highest percentage of *cag* PAI positive strains. No difference in *cag* PAI presence between East Asian and African American strains was observed. As previously mentioned differences in clinical outcome between the ethnic groups cannot be explained exclusively by differences in virulence due to *cag* PAI presence. A systematic review of prospective endoscopic studies in African populations reported that prevalence rate of peptic ulcer disease and gastric cancer among patients infected with *H. pylori* was similar to that of many developed countries [29].

An explanation to bacterial virulence factors causing different clinical outcome is host genetic factors which determine the immune and inflammatory responses to *H. pylori* infection. Polymorphisms in the interleukin-1 gene cluster and in the gene of tumour necrosis factor α, both encoding for pro-inflammatory cytokines, have been associated with *H. pylori*-related gastric cancer and its precursors [31,32].

The differences in size of the left and right border PCR products between East Asian and African American strains might indicate some difference within the *cag* PAI, but whether the low molecular weight PCR product characteristic of East Asian strains plays a role in increase virulence among the East Asian strains or affects the functionality of the *cag* PAI is unknown. There are reports involving the right border size variations, particularly the size of *cagA* which is a part of the right border PCR product, has been shown to vary due to a variable number of repeat sequences in the 3′ region of the gene that encoding for phosphorylation motifs [18]. *cagA* genotypes with the higher number of repeats were found disproportionally among *H. pylori* strains of non-East Asian origin [26,27,33,34].

Variations in the 3′ region of *cagA* may accounts for size variations in the right border and can explain why the low molecular weight PCR product of the right border are more prevalent in *H. pylori* strains of East Asia origin. However, CagA

alone may nothing to do with the pathogenesis of gastric cancer and it is the whole *cag* PAI that can be associated with disease [35].

The presence of the East Asian phenotypes in HspA and upstream of the *hspA* gene, the lack of the 180bp insert and the presence of low molecular weight PCR products of the left and right border of the *cag* PAI correlates strongly with East Asian origin. When four or more of the molecular markers are combined, we can determine a true East Asian strain. However, there is nothing that indicates that these molecular markers are correlated with higher virulence of the bacteria.

The markers for East Asian and African origin found in this study can be used to differentiate *H. pylori* strains of diverse ethnic or geographic origin. The possibility that colonization with *H. pylori* is not harmful but may also protect its host from diseases of the esophagus, it may not always wise to eradicate the bacteria [8]. The recommendations for the future are therefore to investigate if any of the characteristics in any way are correlated with the virulence of the bacteria. This is particularly interesting for the low molecular weight PCR products of the left and the right border of the *cag* PAI characteristic to East Asian origin since disease outcome is more severe in East Asia than in other parts of the world.

REFERENCES

1. World Health Organization.
2. Blaser MJ. The bacteria behind ulcers. Sci Am. 1996; 274: 104-107.
3. Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappuoli R. *Helicobacter pylori* virulence and genetic geography. Science. 1999; 284: 1328-1333.
4. Sung JJ, Chung SC, Ling TK, Yung MY, Leung VK, Ng EK, et al. Antibacterial treatment of gastric ulcers associated with *Helicobacter pylori*. N Engl J Med. 1995; 332: 139-142.
5. Nomura A, Stemmermann GN, Chyou P, Kato I, Perez-Perez GI, Blaser MJ. *Helicobacter pylori* infection and gastric carcinoma among Japanese-Americans in Hawaii. N Engl J Med. 1991; 325: 1132-1136.
6. Uemura N, Okamoto S, Yamamoto S. Matsumura N, Yamaguchi S, Yamakido M. *H. pylori* infection and the development of gastric cancer. N Engl J Med. 2001; 345: 784-789.
7. Hentschel E, Brandstatter G, Dragosics B, Hirschl AM, Nemec H, Schutze K, et al. Effect of ranitidine and amoxicillin plus metrodiazole on the eradication of *Helicobacter pylori* and the recurrence of duoldenal ulcers. N Engl J Med. 1993; 328: 308-312.
8. Blaser MJ. An endangered species in the stomach. Sci Am. 2005; 292: 38-45.
9. Hatakeyama M. Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. Nat Rev Cancer. 2004; 4: 688-694.
10. McNulty SL, Mole BM, Dailidiene D, Segal I, Ally R, Mistry R, et al. Novel 180-and 480-Base-pair insertion in African and African-American Strains of *Helicobacter pylori*. J Clin Microbiol. 2004; 42: 5658-5663.
11. Van Doorn LJ, Figueiredo C, Mégraud F, Pena S, Milodo P, de Magalhães Queiroz DM, et al. Geographic distribution of *vacA* allelic types of *Helicobacter pylori*. Gastroenterol. 1999; 116: 823-830.
12. Chung C, Olivares A, Torres E, Yilmaz O, Cohen H, Perez-Perez GI. Diversity of VacA intermediate region among *Helicobacter pylori* strains from several regions of the world. J Clin Microbiol. 2010; 48: 690-696.
13. Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, et al. Traces of human migrations in *Helicobacter Pylori* populations. Science. 2003; 299: 1582-1585.
14. Enders K, NG W, Thompson SA, Perez-Perez GI, Kansau I, van der Ende A, et al. *Helicobacter pylori* heat shock protein A: serologic responses and genetic diversity. Clin Diagn Lab Immunol. 1999; 6: 377-382.
15. Raymond J, Thiberge JM, Chevalier C, Kalach N, Bergert M, Labigne A, et al. Genetic and transmission analysis of *Helicobacter pylori* strains within a family. Emerg Infect Dis. 2004; 10: 1816-1821.
16. Kansau I, Guillain F, Thiberge JM, Labigne A. Nickel binding and immunological properties of the C-terminal domain of the *Helicobacter pylori* GroES homologue (HspA). Mol Microbiol. 1996; 22: 1013-1023.
17. Suerbaum S, Thiberge JM, Kansau I, Ferrero RL, Labigne A. *Helicobacter pylori hspA-hspB* heat-shock gene cluster: nucleotide sequence, expression, putative function and immunogenicity. Mol Microbiol. 1994; 14: 959-974.
18. Covacci A, Censini S, Bugnoli M, Petracca R, Burroni D, Macchia G, et al. Molecular characterization of the 128-kDa immunodominant antigen of Helicobacter pylori associated with cytotoxicity and duodenal ulcer. Proc Natl Acad Sci. 1993; 90: 5791-5795.
19. Censini S, Lange C, Xiang Z, Crabtree JE, Ghiara P, Borodovsky M, et al. 1996. *cag*, a pathogenicy island of *Helicobacter pylori,* encodes type I-specific and disease-associated virulence factors. Proc Natl Acad Sci. 1996; 93: 14648-14653.
20. Akopyants NS, Clifton SW, Kersulyte D, Crabtree JE, Youree BE, Reece CA, et al. Analyses of the *cag* pathogenicy island of *Helicobacter pylori*. Mol Microbiol. 1998; 28: 37-53.
21. Rhode M, Püls J, Buhrdorf R, Fischer W, Haas R. A novel sheathed surface organelle of the *Helicobacter pylori cag* type IV secretion system. Mol Microbiol. 2003; 49: 219-234.
22. Parsonnet J, Friedman GD, Orentreich N, Vogelman H. Risk for gastric cancer in people with CagA positive or CagA negative Helicobacter pylori infection. Gut. 1997; 40: 297-301.
23. Kuipers EJ, Perez-Perez GI, Meuwissen SG, Blaser MJ. 1995. *Helicobacter pylori* and atrophic gastritis: importance of the cagA status. J Natl Cancer Inst. 1995; 87: 1777-1780.
24. Blaser MJ, Perez-Perez GI, Cover TL, Peek RM, Chyou PH, Stemmermann GN, et al. Infection with Helicobacter pylori strains possessing cagA is associated with an increased risk of developing adenocarcinoma in the stomach. Cancer Res. 1995; 55: 2111-2115.
25. Nomura AM, Lee J, Stemmermann GN, Nomura RY, Perez-Perez GI, Blaser MJ. Helicobacter pylori CagA seropositivity and gastric carcinoma risk in a Japanese American population. J Infect Dis. 2002; 186: 1138-1144.
26. Yamaoka Y, Kodama T, Kashima K, Graham DY, Sepulveda AR. Variants of the 3' region of the *cagA* gene in *Helicobacter pylori* isolates from patients with different *H. pylori* associated disease. J Clin Microbiol. 1998; 36: 2258-2263.
27. Azuma T, Yamakawa A, Yamazaki S, Fukuta K, Ohtani M, Ito Y, et al. Correlation between variation of the 3´region of the *cagA* gene in *Helicobacter pylori* and disease outcome in Japan. J Infect Dis. 2002; 186: 1621-30.
28. Li WH. Molecular evolution. Sunderland (MA): Sinauer Associates, Inc. 1997.
29. Agha A, Graham DY. Evidence-based examination of the African Enigma. Scand J Gastroenterol. 2005; 40: 523-529.
30. Camorlinga-Ponce M, Perez-Perez G, Gonzalez-Valencia G, Mendoza I, Peñaloza-Espinosa R, Ramos I, et al. *Helicobacter pylori* Genotyping

from American Indigenous Groups Shows Novel Amerindian vacA and cagA Alleles and Asian, African and European Admixture. PLoS One. 2011; 6: 27212.

31. El-Omar EM, Rabkin CS, Gammon MD, Vaughan TL, Risch HA, Shoenberg JB, et al. Increased risk of noncardia gastric cancer associated with proinflammatory cytokine gene polymorphisms. Gastroenterol. 2003; 124: 1193-1201.

32. El-Omar EM, Carrington M, Chow W-H, McColl KEL, Bream JH, Young HA, et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature. 2002; 404: 398-402.

33. Zhang XS, Tegtmeyer N, Traube L, Jindal S, Perez-Perez G, Sticht H, et al. A specific A/T polymorphisms in Western tyrosine phosphorylation B-motifs regulates *Helicobacter pylori* CagA epithelial cell interactions. PLoS Pathog. 2015; 11.

34. Yamazaki S, Yamakawa A, Okuda T, Ohtani M, Suto H, Ito Y, et al. Distinct diversity of vacA, cagA, and cagE genes of *Helicobacter pylori* associated with peptic ulcer in Japan. J Clin Microbiol. 2005; 43: 3906-3916.

35. Lu H, Yamaoka Y, Graham DY. *Helicobacter pylori* virulence factors: facts and fantasies. Curr Opin Gastroenterol. 2005; 21: 653-659.

Cite this article

Perez-Perez GI, Andersson M, Olivares AZ (2018) Characterization of East Asian and African Genotypes among Helicobacter pylori Strains Isolated from Different Human Populations. JSM Gastroenterol Hepatol 6(1): 1091.