

RESEARCH ARTICLE

Open Access



# Comparative study between *Helicobacter pylori* and host human genetics in the Dominican Republic

Takaaki Ono<sup>1,2</sup>, Modesto Cruz<sup>3,4</sup>, José A. Jiménez Abreu<sup>5</sup>, Hiroyuki Nagashima<sup>1,6</sup>, Phawinee Subsomsong<sup>1</sup>, Celso Hosking<sup>3</sup>, Seiji Shiota<sup>1</sup>, Rumiko Suzuki<sup>1</sup> and Yoshio Yamaoka<sup>1,7\*</sup> 

## Abstract

**Background:** *Helicobacter pylori*, a bacterium that infects the human stomach, has high genetic diversity. Because its evolution is parallel to human, *H. pylori* is used as a tool to trace human migration. However, there are few studies about the relationship between phylogeography of *H. pylori* and its host human.

**Methods:** We examined both *H. pylori* DNA and the host mitochondrial DNA and Y-chromosome DNA obtained from a total 119 patients in the Dominican Republic, where human demography consists of various ancestries. DNA extracted from cultured *H. pylori* were analyzed by multi locus sequence typing. Mitochondrial DNA and Y-chromosome DNA were evaluated by haplogroup analyses.

**Results:** *H. pylori* strains were divided into 2 populations; 68 strains with African group (hpAfrica1) and 51 strains with European group (hpEurope). In Y-chromosomal haplogroup, European origin was dominant, whereas African origin was dominant both in *H. pylori* and in mtDNA haplogroup. These results supported the hypothesis that mother-to-child infection is predominant in *H. pylori* infection. The Amerindian type of mtDNA haplogroup was observed in 11.8% of the patients; however, Amerindian type (hspAmerind) of *H. pylori* was not observed. Although subpopulation type of most hpAfrica1 strains in Central America and South America were hybrid (hspWAfrica/hpEurope), most Dominican Republic hpAfrica1 strains were similar to those of African continent.

**Conclusions:** Genetic features of *H. pylori*, mtDNA, and Y haplogroups reflect the history of colonial migration and slave trade in the Dominican Republic. Discrepancy between *H. pylori* and the host human genotypes support the hypothesis that adaptability of hspAmerind *H. pylori* strains are weaker than hpEurope strains. *H. pylori* strains in the Dominican Republic seem to contain larger proportion of African ancestry compared to other American continent strains.

**Keywords:** *Helicobacter pylori*, Population structure, Dominican Republic, Phylogeography, Genetic diversity, Human mitochondrial DNA, Human Y-chromosome DNA

## Background

*Helicobacter pylori* (*H. pylori*) is a spiral, gram-negative human pathogen and is a major cause of peptic ulcer disease and gastric cancer. We previously reported that *H. pylori* had already coexisted in the human stomach about 60 thousand years ago when the anatomically

modern human radiated from the African continent [1, 2]. Currently about 50% of world population is infected with this bacterium [3].

Since *H. pylori* frequently undergoes recombination among unrelated strains, its genetic diversity is higher compared with other bacteria, and is about 50 times higher than that of human [4–6]. This large genetic diversity of *H. pylori* helps us to predict host human migration across the world [7, 8]. In recent years, geographical information of *H. pylori* has been attempted to apply to forensic science for detailed survey

\* Correspondence: [yamaoka@oita-u.ac.jp](mailto:yamaoka@oita-u.ac.jp)

<sup>1</sup>Department of Environmental and Preventive Medicine, Faculty of Medicine, Oita University, Yufu, Oita, Japan

<sup>7</sup>Department of Medicine, Gastroenterology and Hepatology Section, Baylor College of Medicine, Houston, TX, USA

Full list of author information is available at the end of the article



of unidentified corpses [9, 10]. Furthermore, a comparative study of *H. pylori* and human DNA in Ladakh of India demonstrated that genetic diversity of *H. pylori* was more informative than human mitochondrial DNA (mtDNA) [11].

*H. pylori* infection occurs mainly vertically, as well as horizontally [12, 13]. In horizontal infection, the risk factor is sanitary conditions such as undeveloped water supply and sewerage and residential environment. Most of the infection is established during infancy until 2 years old when immunity is not sufficiently developed [14]. Vertical infection occurs during this period between intimate family members like a mother and the child [14, 15]. Because vertical infection is dominant, the evolution of *H. pylori* is basically parallel to the host human [16]. As for the Dominican Republic, the World Health Organization and UNICEF have reported that the percentage of improved sanitation facilities in the Dominican Republic in 2015 was 84% [17]. Because the sanitary conditions are not completely improved, *H. pylori* infection in the Dominican Republic can be both vertical and horizontal. However, when multiple strains of different lineage were infected in a patient, the more adaptive strain may overcome the existent strain or hybridize with each other to form a mosaic strain [18–21].

A previous study hypothesized that adaptability of *H. pylori* strains which had been originally infected to American aborigines might be lower than European strains brought by colonial settlers after Columbus's landing [22–25].

A conventional phylogenetic and population analysis of *H. pylori* are based on Multi Locus Sequence Typing (MLST), which sequences seven housekeeping genes (*atpA*, *efp*, *mutY*, *ppa*, *trpC*, *ureI*, *yphC*) for typing [8]. According to population analyses based on MLST, *H. pylori* in the world are classified into 7 population groups: hpEurope, hpEastAsia, hpAsia2, hpAfrica1, hpAfrica2, hpNEAfrica, and hpSahul [1, 16, 21, 26]. The hpEastAsia is subdivided into three subpopulations: hspMaori, hspAmerind, and hspEAsia. The hspMaori is commonly observed in Polynesia, Melanesia and Taiwanese aborigines, hspAmerind is observed in American aborigines, and hspEAsia is commonly observed in East Asia.

Human population was conventionally studied using mtDNA haplogroup and Y-chromosomal haplogroup (Y-haplogroup) as lineage markers [27–30]. The mtDNA is 16,569 bp long and contains two highly polymorphic segments (hypervariable regions, HV1 and HV2) that are located in control region (D-loop region), which does not encode genes. The mtDNA is inherited maternally without recombination and it has been studied for long time. Substantial databases of mtDNA are developed around the world. Y-chromosome DNA is inherited

paternally and undergoes recombination only in a limited region such as pseudoautosomal region (PAR).

Although the adverse effect of *H. pylori* on the digestive system of human hosts is mainly caused by pathogenic genes such as *cagA*, it is hypothesized that the population types to which the infected bacterium belongs may also affect clinical symptoms [31, 32]. Recent studies in Colombia show a mismatch of host (human) ancestry versus bacterial (*H. pylori*) ancestry can lead more severe gastric mucosal damages, thus coevolution likely modulated disease risk [33].

The Dominican Republic is located on the east side of Hispaniola Island in the Caribbean. This island accommodated the first permanent European settlement founded by Christopher Columbus. Our recent study showed the overall prevalence of *H. pylori* infection was 58.9% [34]. The ethnics in the Dominican Republic consists of 16% of European, 11% of African, and 73% of mixed race [35]. It is estimated that 2 to 7 million indigenous people had lived in the Caribbean before the Columbus's landing in 1492 [36]. Although pure descendants of American aborigines in the Dominican Republic were lost after Columbus and successive settlement of African slaves, the Dominican Republic people are expected to have three ancestral components. In this study, we evaluated correspondence between *H. pylori* and human genetics in the Dominican Republic population.

## Methods

### Study population and DNA extraction

Biopsy specimens from gastric mucosa were taken from 258 dyspeptic patients (158 in 2011 and 100 in 2016; 86 males and 172 females; age range, 17–91 years; mean age,  $46.2 \pm 15.8$  years) who underwent endoscopy examination at the Digestive Disease Center, Dr. Luis E. Aybar Health and Hygiene City, Santo Domingo, Dominican Republic. Of these patients, 219 had chronic gastritis, 38 had peptic ulcer diseases, and one had gastric cancer. For *H. pylori* culture, antral biopsy specimens were homogenized and inoculated onto antibiotic selection plates, and then subcultured on Mueller Hinton II Agar medium (Becton Dickinson, Sparks, MD) supplemented with 7% horse blood without antibiotics. The plates were incubated up to 10 days at 37 °C under microaerophilic conditions (10% O<sub>2</sub>, 5% CO<sub>2</sub>, and 85% N<sub>2</sub>). *H. pylori* isolates were identified based on colony morphology; Gram staining results; and oxidase, catalase, and urease reactions. Isolated strains were stored at –80 °C in *Brucella* broth (Becton Dickinson, Sparks, MD) containing 10% dimethyl sulfoxide and 10% horse serum. Bacterial DNA was extracted using a commercially available kit (QIAGEN Inc., Valencia, CA, USA). Eventually, 64 strains were cultured from 158 patients in 2011, and 56 strains were cultured from 100 patients in

2016, thus in total 120 *H. pylori* strains could be obtained. Human DNA was also extracted from biopsies by the same QIAGEN kit. Human DNA of one sample in 2011 had already used up in our previous study [37], therefore we excluded a *H. pylori* strain that does not have corresponding human DNA and used the rest of 119 strains of *H. pylori* and host human DNA in this study. The ethnicity of the 119 patients based on self-assessment at the time of medical examination was 113 multiracial (31 males, 82 females) and 6 African (3 males, 3 females). Written informed consent was obtained from the all participants, and the protocol was approved by the ethics committees of Dr. Luis E. Aybar Health and Hygiene City, the Institute of Microbiology and Parasitology, IMPA, Autonomous University of Santo Domingo, UASD, in the Dominican Republic and Oita University Faculty of Medicine, Japan.

#### Analysis of the host human DNA

The control region of mtDNA sequencing was performed by PCR-based sequencing as described previously [38]. Primers for PCR amplification and direct sequencing are shown in Additional file 1: Table S1. The sequences were aligned and compared with the revised Cambridge Reference Sequence (rCRS) [39, 40] using MEGA software (version 7.0) [41]. After the consensus sequence of each individual was obtained, polymorphisms were examined. The mtDNA haplogroup was determined by Haplogrep software (version 2.0, <https://haplogrep.uibk.ac.at/>) [42] or by direct comparison with PhyloTree 17 data [43]. For individuals that were difficult to judge mtDNA haplogroup by the control region, single nucleotide polymorphism (SNP) of the coding region were also examined as described previously [44].

On 34 male samples, 17 loci of Y-chromosomal short tandem repeats (Y-STR) were determined by AmpFISTR Yfiler Kit (Applied Biosystems, Foster city, CA) according to the manufacturer's instructions. Y-haplogroup was predicted by Haplogroup Predictor (<http://www.hprg.com/hapest5/>). Based on the results of Y-haplogroup estimated from Y-STR, 12 types of Y-SNP markers (M168, M145, P170, M201, M170, P209, M213, M9, M45, M207, M198, M343) were selected referring to the previous method, and all SNPs were determined by SNaPshot method as described previously [45]. Primers are shown in Additional files 2 and 3: Tables S2 and S3.

#### Analysis of *H. pylori* population structure

Seven housekeeping genes (*atpA*, *efp*, *mutY*, *ppa*, *trpC*, *ureI*, *yphC*) of MLST were determined by PCR-based sequencing as described previously [5]. Primers for PCR amplification and direct sequencing are shown in Additional file 4: Table S4. For construction of phylogenetic

tree, 1293 MLST sequences of global strains were obtained from the previous studies: 229, 67, 113, 544, 92, 50, 128, 39, and 31 strains from hpAfrica1, hpAfrica2, hpNEAfrica, hpEurope, hpAsia2, hpSahul, hspEAsia, hspMaori, and hspAmerind, respectively (Additional file 5: Table S5) [1, 2, 11, 16, 18, 21, 2231, 46–53]. We used MLST sequences of our Dominican Republic strains and these global strains for the successive analyses. Neighbor-joining trees were constructed by MEGA software (version 7.0) using Kimura-2 parameters model [41, 54, 55]. To investigate the population structure of the Dominican Republic strains, we executed Bayesian population assignments by STRUCTURE software (version 2.3.3), using “no-admixture model” and “linkage model” [56], on totally 1295 strains: the 119 Dominican Republic strains and 1176 published global reference sequences (excluding hpAfrica2 and hpSahul from the above 1293 reference strains because their groups had no relevance with the Dominican Republic strains in the phylogenetic tree). To determine the number of bacterial populations (K) within the dataset, STRUCTURE was executed by setting K from 4 to 7 (10 runs for each K) with 30,000 iterations following a burn-in period of 20,000 iterations. Further STRUCTURE analyses were performed to detect subpopulations for each hpEurope and hpAfrica1 group by setting K from 2 to 5 (10 runs for each K). The STRUCTURE runs with the highest posterior probability among 10 runs were used for subsequent analyses such as statistical test.

To investigate which strain of African continent is close to the hpAfrica1 strains in the Dominican Republic, a phylogenetic tree was constructed incorporating only strains that belonged to hpAfrica1 with probability of 100% in no-admixture model of the STRUCTURE (K = 5, 1295 strains) as pure hpAfrica1 strains (163 reference and 45 Dominican Republic). Neighbor-joining trees were constructed by MEGA software (version 7.0) using Kimura-2 parameters.

Estimates of the number of polymorphic sites, haplotype diversity and nucleotide diversity were calculated by the Arlequin software (version 3.5) [57]. Pairwise genetic distances were calculated by MEGA software (version 7.0) using Kimura-2 parameters. Statistical significance was tested using Kruskal-Wallis test, Fisher's exact test, and Wilcoxon rank sum test implemented in the R package (version 3.4.0). A *p*-value of < 0.05 was accepted as statistically significant.

## Results

#### Host human DNA

In mtDNA analysis, a total of 715 bp nucleotide sequences of HV1 (16024–16,400) and HV2 (73–407) were determined from all the 119 human DNA. There were 94 different mtDNA haplotypes and sequences of

82 individuals were all different from each other. The haplotype diversity and the nucleotide diversity of the 715 bp mtDNA control region of the 119 individuals were  $0.9930 \pm 0.0026$  and  $0.021 \pm 0.01$ , respectively. Thus, 119 patients were not thought to have biased kinship. Of the 12 haplotypes shared by multiple individuals, 4 were shared by 2 individuals, 5 were shared by 3 individuals, 2 were shared by 4 individuals and 1 was shared by 6 individuals. These 119 individuals were divided into 3 geographical classifications by mtDNA haplogroups. 96 belonged to African type: 9 were haplogroup L0a, 12 were L1 (8 and 4 of L1b and L1c, respectively), 30 were L2 (12, 6, 9, and 3 of L2a, L2b, and L2c, respectively), 43 were L3 (11, 11, 15 and 6 of L3b, L3d, L3e, and L3f, respectively), one was L4, one was U6. 14 belonged to Amerindian type: 5 were haplogroup A2, 3 were B2, 5 were C1, one was D1. 9 belonged to European type: 2 were haplogroup H, one was HV, 3 were J1, 3 were U5 (Additional files 6 and 7: Tables S6 and S7). In Y-STR analysis, 33 haplotypes were observed among 34 males, of which 1 haplotype was shared by 2 individuals. The haplotype diversity of Y-STR was

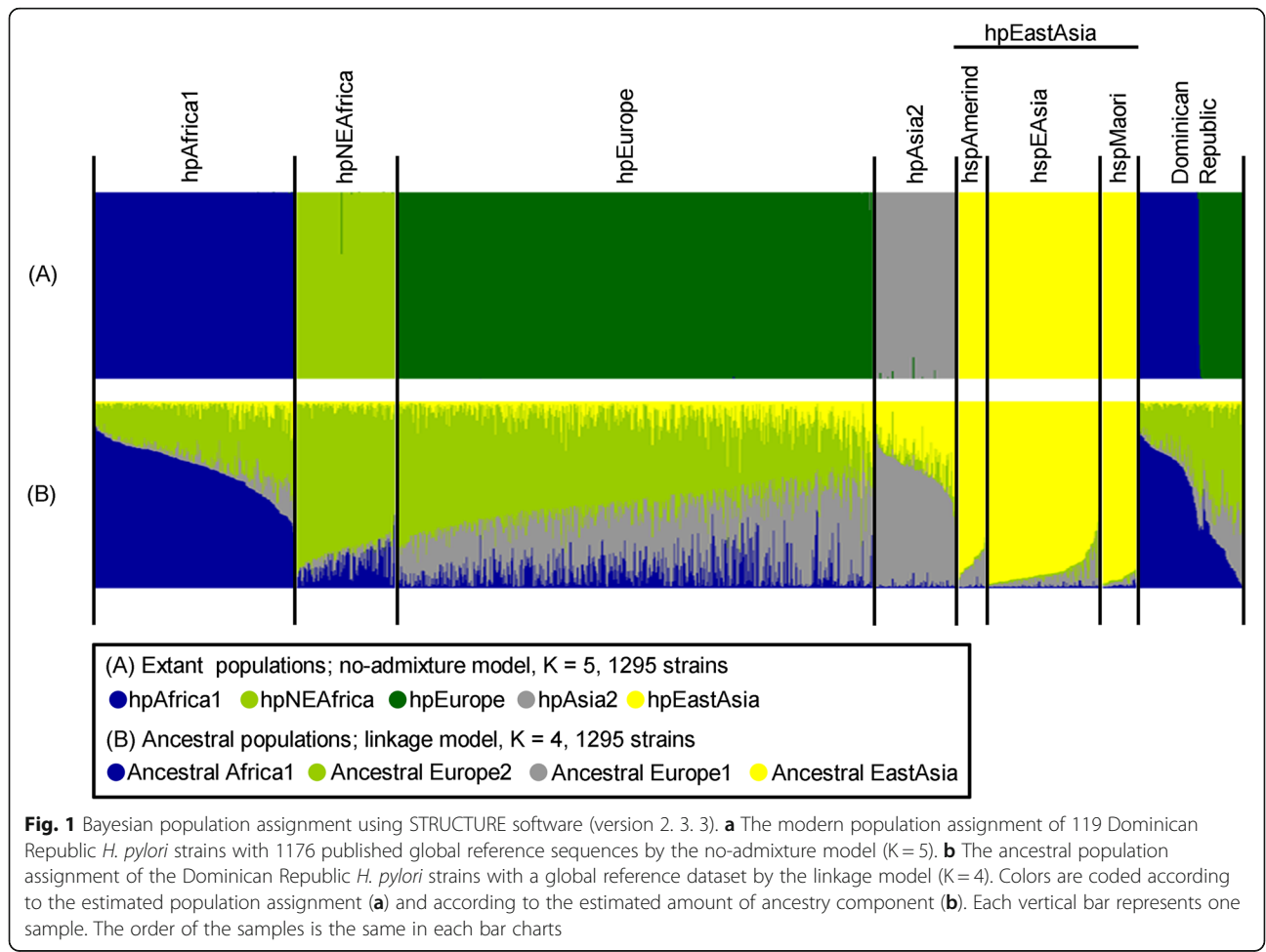
$0.9982 \pm 0.0077$ . These 34 males were divided into 2 geographical classifications by Y-STR and Y-SNP markers: 23 were European type (4 were haplogroup I, 7 were J, 12 were R1b) and 11 were African type (all were haplogroup E) (Additional files 7 and 8: Tables S7 and S8).

**H. pylori population structure**

MLST sequences of 7 housekeeping genes (in total 3406 bp) were obtained from the all 119 *H. pylori* strains. MLST sequences of the 119 strains were all different from each other and contained 701 polymorphic sites. No deletion or insertion was observed. The nucleotide diversity of the MLST sequences was  $0.037 \pm 0.017$ .

We constructed a phylogenetic tree based on MLST sequences of the 119 Dominican Republic strains and 1293 reference strains. The Dominican Republic strains were located in either hpEurope or hpAfrica1 subbranches (Additional file 9: Figure S1).

To investigate the population structure of the Dominican Republic *H. pylori* strains, we performed population analysis using the STRUCTURE software. Figure 1a shows the result of STRUCTURE run of no-admixture





model at  $K = 5$ . Under this condition, most of the global reference strains were clearly classified into the major five populations reported in the previous studies [1, 16, 26]. In the Dominican Republic, 119 strains were assigned to either hpAfrica1 (57.1%, 68/119) or hpEurope (42.9%, 51/119) (Additional file 7: Table S7). No Dominican Republic strain was classified to hpNEAfrica, hpAsia2 or hpEastAsia, and no novel population was observed. The result of STRUCTURE run of linkage model at  $K = 4$  is very well described in Fig. 1b. Under this condition, the major four ancestral components reported in the previous studies [1, 16, 26] were observed: ancestral Africa1 (AA1), ancestral Europe 1 (AE1), ancestral Europe 2 (AE2) and ancestral EastAsia (AEA). hpEurope is known as an admixture population between AE1 and AE2 that originated from Central Asia and Northeast Africa, respectively. In the Dominican Republic, 50/51 of hpEurope strains had the higher proportion of AE2 than AE1.

Next, we executed STRUCTURE no-admixture model on 51 Dominican Republic strains assigned to hpEurope and 544 global hpEurope reference strains. At  $K = 2$ , hspEuropeN and hspEuropeS were identified according to the origin of 544 reference hpEurope strains reported in the previous study [2, 16, 21] (Fig. 2). In the Dominican Republic, 4 strains were assigned to hspEuropeN (7.8%, 4/51) and 47 strains to hspEuropeS (92.2%, 47/51). In addition, we compared the ratio of ancestral component in the hspEuropeS strains dividing by geographic regions: Iberian Peninsula ( $n = 85$ ), Central America ( $n = 53$ ), South America ( $n = 93$ ), Dominican Republic ( $n = 47$ ). Statistical test showed that the Dominican Republic group had a significantly higher AA1 component than Iberian Peninsula group (Kruskal-Wallis test followed by Steel-Dwass post-hoc test,  $P < 0.001$ ; Additional file 10: Figure S2). Likewise, the Dominican Republic group had a significantly lower AEA component than Central America group and South America group (Kruskal-Wallis test followed by Steel-Dwass post-hoc test,  $P < 0.001$ ; Additional file 11: Figure S3). In contrast, AEA component of Iberian Peninsula group and the Dominican Republic group had no statistically significant difference.

Next, we also executed STRUCTURE no-admixture model on 68 Dominican Republic strains assigned to hpAfrica1 and 229 reference hpAfrica1 strains. At  $K = 2$ , hspWAfrica and hspSAfrica were identified according to the origin of 229 reference hpAfrica1 strains (Additional file 12: Figure S4). At  $K = 3$ , hspWAfrica was divided into two groups, but hspCAfrica subpopulation reported in the previous study was not formed [46], these three subpopulations showed different ratio of AE1 component (Fig. 3). The pink colored strains in Fig. 3 had a significantly higher ratio of AE1 component than

other groups (Kruskal-Wallis test followed by Steel-Dwass post-hoc test,  $P < 0.001$ , respectively). Thus, these subpopulations were considered to be hspWAfrica, hspSAfrica and hybrid between hspWAfrica and hpEurope (Additional file 13: Figure S5).

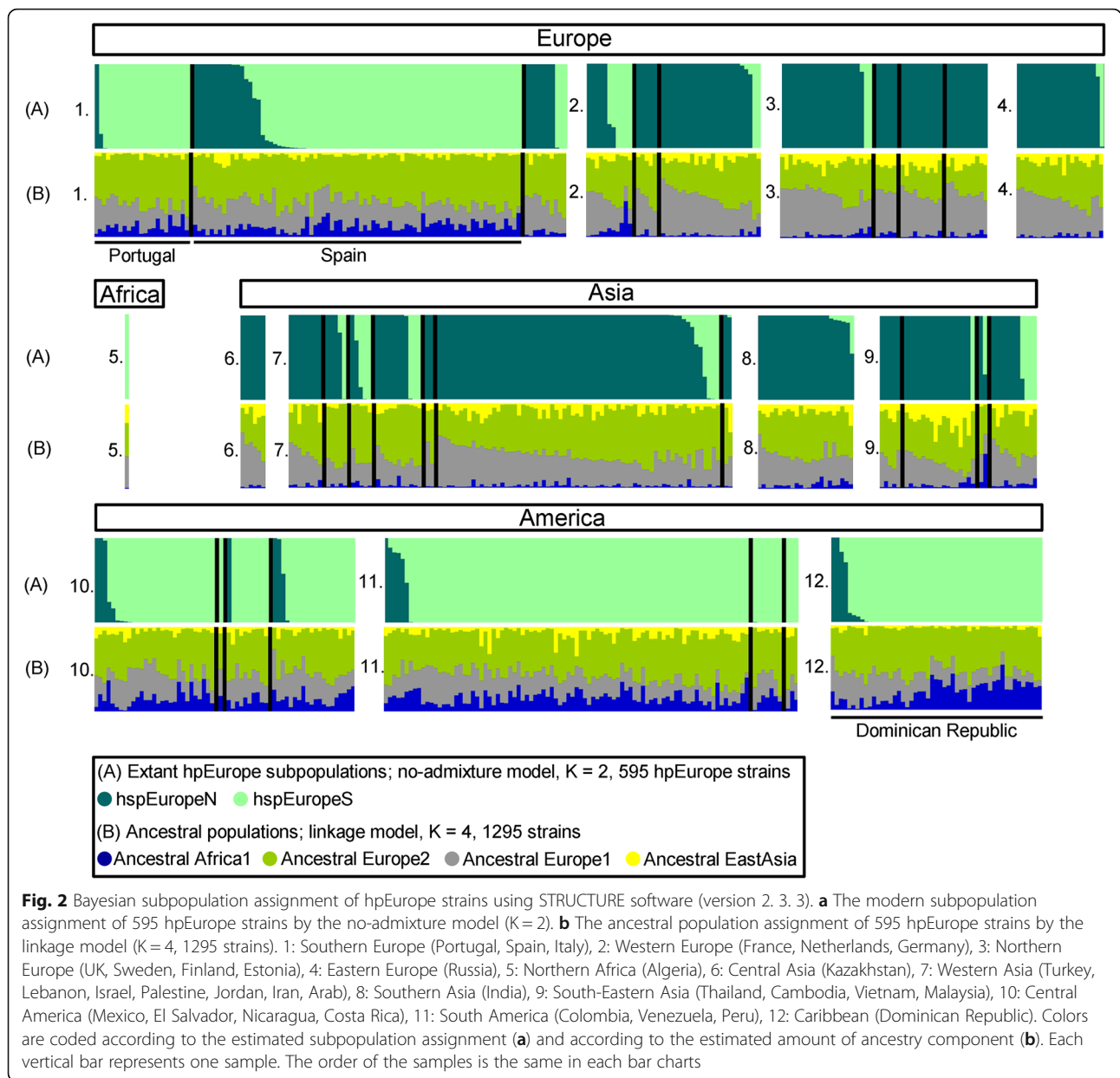
In the Dominican Republic, 68 hpAfrica1 strains were assigned to hspWAfrica (67.6%, 46/68), hspSAfrica (7.4%, 5/68) or hspWAfrica/hpEurope hybrid (25.0%, 17/68). On the contrary, 96.7% (29/30) and 90.0% (9/10) were hybrid (hspWAfrica/hpEurope) in Central America and South America, respectively. We compared the ratio of ancestral component of the hybrid (hspWAfrica/hpEurope) strains by geographic region: African continent ( $n = 24$ ), Central America ( $n = 29$ ), South America ( $n = 9$ ), Dominican Republic ( $n = 17$ ). Statistical test showed that Central America group had a significantly higher AEA component than Dominican Republic group (Kruskal-Wallis test followed by Steel-Dwass post-hoc test,  $P < 0.001$ ; Additional file 14: Figure S6). At  $K = 4$ , 10 strains of classified as hybrid (hspWAfrica/hpEurope) at  $K = 3$  formed a novel subpopulation composed of Nicaragua ( $n = 8$ ), Guatemala ( $n = 1$ ) and Costa Rica ( $n = 1$ ), shown in red color in Additional file 12: Figure S4. This novel subpopulation group had a significantly higher AEA component than the other groups (Wilcoxon rank sum test,  $P = 0.003$ ; Additional file 15: Figure S7). Next, we constructed a phylogenetic tree to investigate the origin of hpAfrica1 strains in the Dominican Republic (Fig. 4). Many of the Dominican Republic hspWAfrica strains appeared near branch of the strains from Burkina Faso, Gambia and Senegal. Four Dominican Republic hspSAfrica strains formed a cluster distant from other African strains.

#### Comparison of genetic diversity by each *H. pylori* population

The number of polymorphic site and the nucleotide diversity of the 68 hpAfrica1 strains were 575 and  $0.031 \pm 0.014$ , respectively, and the number of polymorphic site and the nucleotide diversity of the 51 hpEurope strains were 535 and  $0.036 \pm 0.018$ , respectively. The genetic diversity of hpEurope population was significantly greater than hpAfrica1 population ( $P < 0.001$ , Wilcoxon rank sum test; Additional file 16: Figure S8).

#### Relationship between phylogeographical classification of *H. pylori* and host human

Figure 5a shows the number of *H. pylori* population type in each mtDNA haplogroup. In both African and European mtDNA haplogroups, hpAfrica1 *H. pylori* was predominant. In contrast, in Amerindian mtDNA haplogroup, hpEurope was predominant. The difference of *H. pylori* populations was significant between the African and the Amerindian mtDNA haplogroups ( $P = 0.023$ , Fisher's exact test followed by Bonferroni post-



hoc test). In addition, Fig. 5b shows the box plot diagram of the ratio of European ancestral components (AE1 + AE2) of *H. pylori* in each mtDNA haplogroup. The AE1 + AE2 ratio was significantly different between Amerindian and African mtDNA haplogroups ( $P = 0.026$ , Kruskal-Wallis test followed by Steel-Dwass post-hoc test).

On the contrary, no significant difference was found between *H. pylori* population and Y-haplogroup (Additional file 17: Figure S9).

**Nucleotide sequencing**

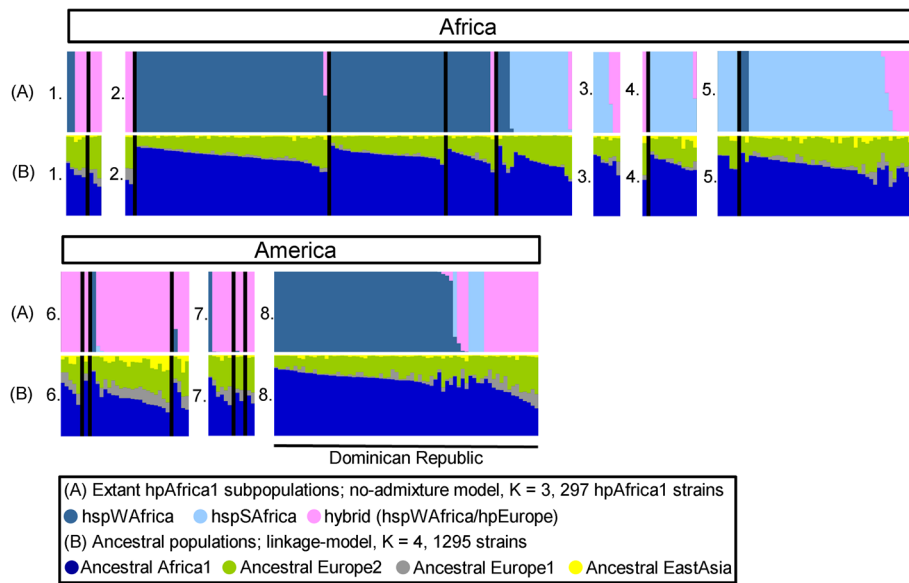
Sequencing data for seven housekeeping genes of *H. pylori* and mtDNA of human DNA are available under

DDBJ accession numbers LC321074- LC321906 and LC319790- LC320027, respectively.

**Discussion**

**Host human DNA**

In the analysis of mtDNA haplogroup, both diversity and frequency of African type were larger than that of Amerindian type and European type. Many African mtDNA haplogroups (L1, L2, L3) observed in this study were frequent in West Africa [58, 59]. Therefore, these results reflected the history that the slave trade to the Caribbean was mainly from West Africa. All the four major Amerindian mtDNA haplogroups (A, B, C, D) [60] were observed in Amerindian type mtDNA of 14



**Fig. 3** Bayesian subpopulation assignment of hpAfrica1 strains using STRUCTURE software (version 2.3.3). **a** The modern subpopulation assignment of 297 hpAfrica1 strains by the no-admixture model (K = 3). **b** The ancestral population assignment of 297 hpAfrica1 strains by the linkage model (K = 4, 1295 strains). 1: Northern Africa (Morocco, Algeria), 2: Western Africa (Cape Verde, Senegal, Gambia, Burkina Faso, Cameroon), 3: Middle Africa (Angola), 4: Eastern Africa (Mozambique, Madagascar), 5: Southern Africa (Namibia, South Africa), 6: Central America (Mexico, Guatemala, Nicaragua, Costa Rica), 7: South America (Colombia, Venezuela, Brazil), 8: Caribbean (Dominican Republic). Colors are coded according to the estimated subpopulation assignment (**a**) and according to the estimated amount of ancestry component (**b**). Each vertical bar represents one sample. The order of the samples is the same in each bar charts

individuals. This suggests that the human population in the Dominican Republic has a trace of ancient migrants from East Asia via the Bering Strait before Columbus. The four European mtDNA haplogroups (H, HV, J1, U5) observed in 9 individuals are common in the Iberian Peninsula where the colonial settlers were originated from [61]. Thus, the result is also consistent with the history. In Y-haplogroup analysis of 34 men, three European haplogroups (I, J, R1b) were observed, which were commonly recorded in the Iberian Peninsula [62]. The African Y-haplogroup E is common in West Africa. In contrast to mtDNA haplogroup, Amerindian Y-haplogroup was not observed at all.

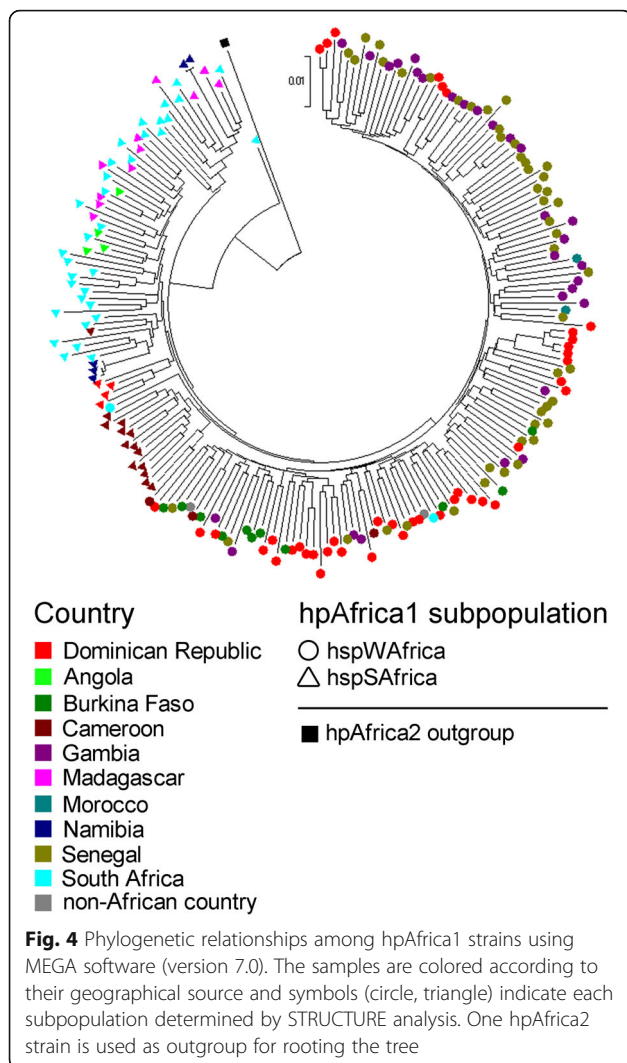
In this study, African haplogroup was dominant in mtDNA and European haplotype was dominant in Y chromosome. These observations were consistent with the previous studies in the Dominican Republic [63–65]. Probably African slave men and American aboriginal men could not leave many descendants because of battle, slavery, or infectious diseases brought by Europeans during the slavery era that lasted more than 3 centuries. European mtDNA haplogroup is rare probably because only a small number of women immigrated from Europe in the past. Possible reasons why Amerindian Y-haplogroup was not observed while Amerindian mtDNA was observed is a bias for the survival of Amerindian women, the smaller sample size of Y than mtDNA or a

result of miscegenation between Amerindian women and European men.

**H. pylori population structure**

We confirmed that 119 *H. pylori* strains isolated in the Dominican Republic were classified into hpEurope and hpAfrica1. Furthermore, hpEurope strains were divided into two subpopulations: hspEuropeN and hspEuropeS. Modern hpEurope strain is a hybrid of ancestral Europe 1 (AE1), which is a main component of hpAsia2, and ancestral Europe 2 (AE2), which is a main component of hpNEAfrica. Comparing the ratio of the ancestral component of reference Eurasian continent strains used in this analysis, hspEuropeS, mainly observed in Iberian Peninsula, had higher AE2 than AE1. In contrast, hspEuropeN, mainly observed in northern Europe and Asia, had higher AE1 than AE2. In addition, hspEuropeS had a relatively high component of ancestral Africa1 (AA1), while hspEuropeN had a relatively high component of ancestral EastAsia (AEA). Furthermore, the largest subpopulation among hpEurope strains in American continent was hspEuropeS. These results were consistent with the history that the origin of the colonial settlers to American Continent was Iberian Peninsula.

Interestingly, hspEuropeS strains in Central America and Dominican Republic contain significantly higher AA1 component than that of Iberian Peninsula



(Additional file 10: Figure S2). In addition, a high proportion of AA1 component was observed in part of the hspEuropeS strains, similarly to the report about the Portuguese speaking countries by Oleastro et al. [18]. This result suggests that there was gene flow from hpAfrica1 strains brought by African slaves to hpEurope strains brought by the colonial settlers in Central America and Dominican Republic. On the other hand, hspEuropeS strains in Central and South America had significantly higher AEA components than that of the Iberian Peninsula and the Dominican Republic. This suggests that hspEuropeS strains in Central and South America underwent genetic exchange with hspAmerind strains hosted by American aborigines. The ratio of AEA component in the Dominican Republic strains is as low as the strains in the Iberian Peninsula.

These results show that hpEurope strains in the Dominican Republic is different from other countries in that they were highly affected by hpAfrica1 but not by

hspAmerind. This country was dominated by colonial settlers first in the American continent, and at the same time, many slaves were forcibly brought from Africa. Therefore, cohabitation history of Europeans and Africans is the longest, and American aborigines might have disappeared in the earliest in the Dominican Republic. Furthermore, genetic exchange with neighboring countries was restricted because of the island environment.

hpAfrica1 strains were divided into three subpopulations: hspWAfrica, hspSAfrica, and hybrid (hspWAfrica/hpEurope). The hybrid (hspWAfrica/hpEurope) strains mainly observed in the country of the Mediterranean coast of the African continent and in the American continent. In the Dominican Republic, pure hspWAfrica and hspSAfrica strains were observed more frequently than other American continent countries. In addition, hybrid (hspWAfrica/hpEurope) strains in the Dominican Republic had significantly less AEA component than that of other American continent countries (Additional file 14: Figure S6). The reason may be the same to the reason of less AEA component in hpEurope strains in the Dominican Republic.

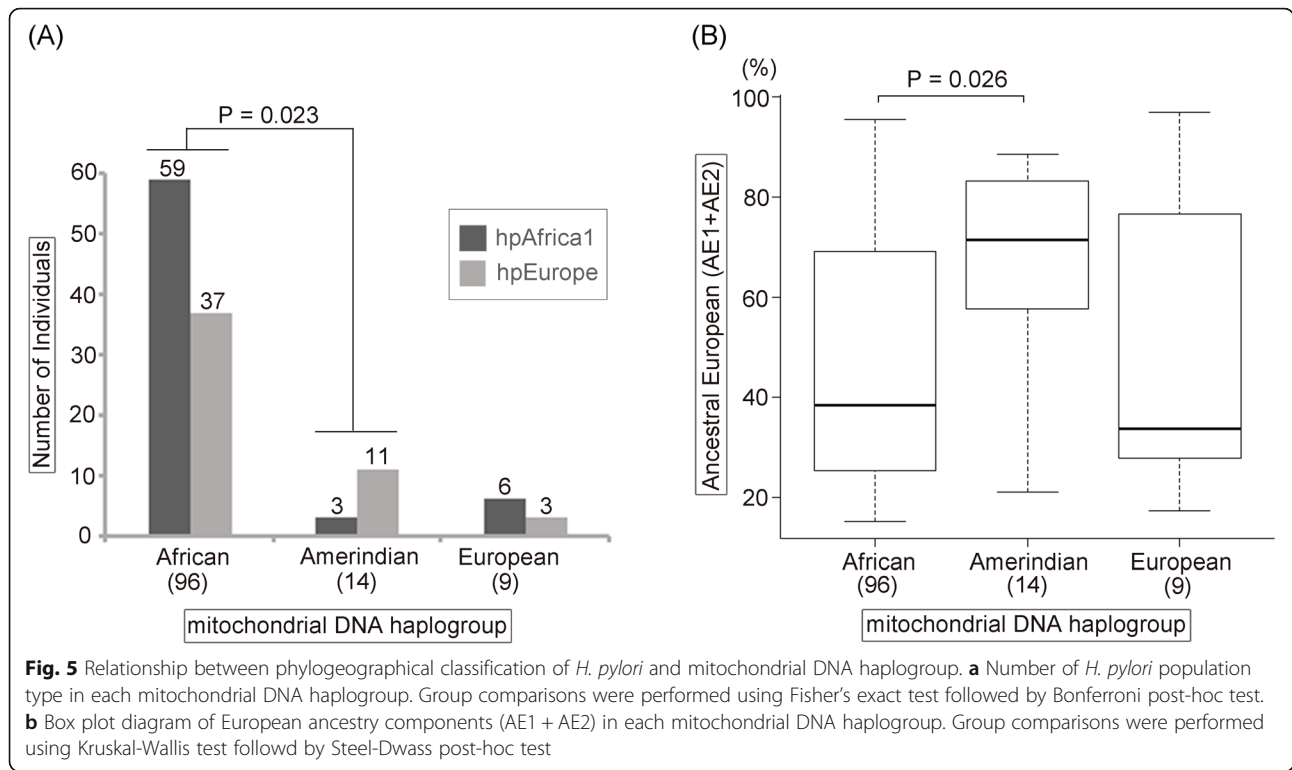
A distinct bacterial population appeared at  $K = 4$  predominant with Nicaragua strains (Additional file 12: Figure S4), which corresponds to hspAfrica1Nicaragua in a previous study [21]. hspAfrica1Nicaragua was reported to be a subpopulation that appeared due to the rapid evolution of the hpAfrica1 strains in American continent. Although a previous study [21] on whole genome data of *H. pylori* reported that there is no evidence that hspAmerind strains contributed DNA to other extraneous New World strains, this study demonstrated that hspAfrica1Nicaragua strains contained high AEA components. This observation supports a hypothesis that the hspAmerind strain disappeared through strain subversion by transformation [22].

The phylogenetic tree showed that hspWAfrica strains in the Dominican Republic were close to the strains in Burkina Faso, Gambia and Senegal (Fig. 4). These countries are located in West Africa where slave trade was thriving. The four hspSAfrica strains in the Dominican Republic that formed a cluster distant from other African strains may derive from a region of African continent with yet no survey of *H. pylori*.

#### Relationship between phylogeographical classification of *H. pylori* and host human

The mtDNA is inherited maternally because only mitochondria in the ova are passed on to the child but there is no contribution from the father. Contrary, Y chromosome is inherited paternally to the son. Intra family infection of *H. pylori* occurs often mainly by vertical transmission from a mother who contact to her child intimately. Thus, we initially anticipated that the infection





pattern of *H. pylori* might resemble to the inheritance of mtDNA haplogroup rather than Y-haplogroup.

The results that the ratio of hpAfrica1 strains was the highest (57.1%) in *H. pylori* and the ratio of African haplogroup was the highest (80.7%) in mtDNA supported the dominance of mother-to-child infection. However, we observed discrepancy between Amerindian mtDNA and *H. pylori* genotype. Amerindian mtDNA was observed in 11.8% of the individuals but none of them had *H. pylori* that belong to hspAmerind. The reason may be explained by a hypothesis that the adaptability of the hspAmerind strains is lower than that of the hpEurope strains [22, 23]. In addition, a previous study reported that hpAfrica1 strains are as adaptive as hpEurope strains [47]. Therefore, hspAmerind may be the weakest.

The proportion of hpEurope in *H. pylori* (42.0%) was much higher than that of European mtDNA haplogroup (7.6%) because 38.5% of African mtDNA haplogroup and 78.6% of Amerindian mtDNA haplogroup had hpEurope *H. pylori*. These results suggest that infection occurs not only from mother to child but also from father to child, or horizontal transmission occurs from environment such as water supply. However, Amerindian mtDNA haplogroup had significantly higher ratio of hpEurope *H. pylori* and higher amount of European ancestral component (AE1 + AE2) than African mtDNA haplogroup. Thus, the original hspAmerind strains infected to Amerindian aborigines might be replaced by hpEurope

strain during the past 500 years. This hypothesis also reflects the history that many aboriginal American women had intermarriage with settler men [65].

In European mtDNA haplogroup, hpAfrica1 *H. pylori* was predominant (6/9 = 66.7%). Although the infection of hpAfrica1 to the patient of European mtDNA could be caused by horizontal transmission, it is also probable that hpAfrica1 might have higher adaptability than hpEurope strains. A previous study predicted that higher genetic diversity is advantageous to adapt the host range [22]. In the Dominican Republic, the genetic diversity of hpEurope was higher than that of hpAfrica1. Thus, genetic diversity does not explain the excess of hpAfrica1 *H. pylori* among individuals of European mtDNA. Further study is needed to clarify the reason.

To our knowledge, this report is the first comparative study between *H. pylori* and mtDNA and Y haplogroups of admixture population in South America. However, there are several limitations in this study. Firstly, the number of male patients was small. For a better understanding of the relationship between Y-haplogroup and *H. pylori*, larger number of male samples will be necessary. Secondly, our samples were taken at a hospital in Santo Domingo, the capital city of the Dominican Republic. The physical and cultural landscape varies by region in the Dominican Republic. Therefore, our results cannot be generalized across the entire region of the Dominican Republic.

## Conclusions

We found that *H. pylori* in the Dominican Republic consists of two populations: hpAfrica1 and hpEurope. Although the Amerindian type of mtDNA haplogroup was observed in 11.8% of the patients, Amerindian type (hspAmerind) of *H. pylori* was not observed. This result supports the hypothesis that hspAmerind strains have lower adaptability than other groups because of low genetic diversity. *H. pylori* strains in the Dominican Republic have different characteristics from South and Central American countries in that they have high component of African ancestry but poor Amerindian component, which reflects the history and the geographic condition of this country.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12862-019-1526-9>.

**Additional file 1: Table S1.** Primers used for DNA sequencing of human mitochondrial DNA.

**Additional file 2: Table S2.** Primers used for multiplex amplification of human Y-SNPs.

**Additional file 3: Table S3.** Primers used for minisequencing of Human Y-SNPs.

**Additional file 4: Table S4.** Primers used for DNA sequencing of *H. pylori*.

**Additional file 5: Table S5.** Sources and population assignment of *H. pylori* global reference strains.

**Additional file 6: Table S6.** Results of mitochondrial DNA haplotypes and haplogroups.

**Additional file 7: Table S7.** List of strain information, mitochondrial DNA haplogroup, Y chromosomal haplogroup and population assignment of *H. pylori* in the Dominican Republic.

**Additional file 8: Table S8.** Results of Y chromosomal DNA haplotypes and haplogroups.

**Additional file 9: Figure S1.** Phylogenetic tree of 119 Dominican Republic *H. pylori* strains with 1293 global reference strains.

**Additional file 10: Figure S2.** Box plot diagram of ancestral Africa1 components (AA1) in hspEuropeS subpopulation divided by regions: Iberian Peninsula ( $n = 85$ ), Central America ( $n = 53$ ), South America ( $n = 93$ ), Dominican Republic ( $n = 47$ ). The difference of AA1 ratio between regions was investigated by Kruskal-Wallis test followed by Steel-Dwass post-hoc test.

**Additional file 11: Figure S3.** Box plot diagram of ancestral EastAsia components (AEA) in hspEuropeS subpopulation divided by regions: Iberian Peninsula ( $n = 85$ ), Central America ( $n = 53$ ), South America ( $n = 93$ ), Dominican Republic ( $n = 47$ ). The difference of AEA ratio between regions was investigated by Kruskal-Wallis test followed by Steel-Dwass post-hoc test.

**Additional file 12: Figure S4.** Bayesian subpopulation assignment of 297 hpAfrica1 strains using the no-admixture model ( $K = 2$ ,  $K = 3$ ,  $K = 4$ ) of STRUCTURE software (version 2.3.3). 1: Northern Africa (Morocco, Algeria), 2: Western Africa (Cape Verde, Senegal, Gambia, Burkina Faso, Cameroon), 3: Middle Africa (Angola), 4: Eastern Africa (Mozambique, Madagascar), 5: Southern Africa (Namibia, South Africa), 6: Central America (Mexico, Guatemala, Nicaragua, Costa Rica), 7: South America (Colombia, Venezuela, Brazil), 8: Caribbean (Dominican Republic). Colors are coded according to the estimated subpopulation assignment. Each vertical bar represents one sample. The order of the samples is the same in each bar charts.

**Additional file 13: Figure S5.** Box plot diagram of ancestral Europe 1 components (AE1) in the three subpopulations classified by STRUCTURE analysis (no-admixture model,  $K = 3$ ) of 297 hpAfrica1 strains. The difference of AE1 ratio between subpopulations was investigated by Kruskal-Wallis test followed by Steel-Dwass post-hoc test.

**Additional file 14: Figure S6.** Box plot diagram of ancestral EastAsia components (AEA) in hybrid (hspWAfrica/hpEurope) subpopulation divided by regions: African continent ( $n = 24$ ), Central America ( $n = 29$ ), South America ( $n = 9$ ), Dominican Republic ( $n = 17$ ). The difference of AEA ratio between regions was investigated by Kruskal-Wallis test followed by Steel-Dwass post-hoc test.

**Additional file 15: Figure S7.** Box plot diagram of ancestral EastAsia components (AEA) in the two groups within hybrid (hspWAfrica/hpEurope) subpopulation that indicated by STRUCTURE analysis (no-admixture model,  $K = 4$ ) of 297 hpAfrica1 strains. The difference of AEA ratio between groups was investigated by Wilcoxon rank sum test.

**Additional file 16: Figure S8.** Box plot diagram of pairwise genetic distances between *H. pylori* strains grouped by bacterial population. The difference of pairwise genetic distance between the bacterial populations was investigated by Wilcoxon rank sum test.

**Additional file 17: Figure S9.** Relationship between phylogeographical classification of *H. pylori* and Y chromosomal haplogroup. (A) Number of *H. pylori* population type in each Y chromosomal haplogroup. Group comparisons were performed using Fisher's exact test. (B) Box plot diagram of European ancestry components (AE1 + AE2) in each Y chromosomal haplogroup. Group comparisons were performed using Wilcoxon rank sum test.

## Abbreviations

*atpA*: ATP synthase subunit alpha; *efp*: Elongation factor P; *H. pylori*: *Helicobacter pylori*; HV1: Hypervariable Region 1; HV2: Hypervariable Region 2; MLST: Multi Locus Sequence Typing; mtDNA: mitochondrial DNA; *mutY*: A/G-specific adenine glycosylase; *ppa*: Inorganic pyrophosphatase; SNP: Single nucleotide polymorphism; STR: Short Tandem Repeats; *trpC*: Bifunctional indole-3-glycerol phosphate synthase; *ureI*: Urease accessory protein; *yphC*: GTP-binding protein

## Acknowledgements

We thank members of our laboratories for discussions and comments.

## Authors' contributions

Conceived and designed the experiments: SS, RS, YY. Performed the experiments: TO, HN, PS. Analyzed the data: TO, RS, YY. Contributed reagents/materials/analysis tools: MC, JJ, CH. Wrote the manuscript: TO, RS, YY. All authors read and approved the final manuscript.

## Funding

This work was funded by grants-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan (221S0002, 16H06279, 15H02657 and 16H05191), by the Japan Society for the Promotion of Science (Core-to-Core Program), and by National Institutes of Health (DK62813). It was also supported in parts by a grant of The National Fund for Innovation and Development of Science and Technology (FONDO-CYT) from the Ministry of Higher Education Science and Technology (MES-CyT) of the Dominican Republic (2012–2013-2A1–65 and 2015-3A1–182) (MC).

## Availability of data and materials

Sequencing data for seven housekeeping genes of *H. pylori* and mtDNA of human DNA are available under DDBJ accession numbers LC321074-LC321906 and LC319790-LC320027, respectively.

## Ethics approval and consent to participate

Written informed consent was obtained from the all participants, and the protocol was approved by the ethics committees of Dr. Luis E. Aybar Health and Hygiene City, the Institute of Microbiology and Parasitology, IMPA, Autonomous University of Santo Domingo, UASD, in the Dominican Republic and Oita University Faculty of Medicine, Japan.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Department of Environmental and Preventive Medicine, Faculty of Medicine, Oita University, Yufu, Oita, Japan. <sup>2</sup>Criminal Investigation Laboratory, Oita Prefectural Police H.Q, Oita, Japan. <sup>3</sup>Institute of Microbiology and Parasitology, Faculty of Science, Autonomous University of Santo Domingo, Santo Domingo, Dominican Republic. <sup>4</sup>Department of Biomedical Research, National Institute of Medicine and Diagnostic Imaging, Santo Domingo, Dominican Republic. <sup>5</sup>Dominican-Japanese Digestive Disease Center, Dr Luis E. Aybar Health and Hygiene City, Santo Domingo, Dominican Republic. <sup>6</sup>Department of Gastroenterology, Hokkaido Cancer Center, Sapporo, Hokkaido, Japan. <sup>7</sup>Department of Medicine, Gastroenterology and Hepatology Section, Baylor College of Medicine, Houston, TX, USA.

Received: 27 September 2018 Accepted: 15 October 2019

Published online: 01 November 2019

**References**

- 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;445:915–8.
- Moodley Y, Linz B, Bond RP, Nieuwoudt M, Soodyall H, Schlebusch CM, et al. Age of the association between *Helicobacter pylori* and man. *PLoS Pathog*. 2012;8:e1002693.
- Suerbaum S, Michetti P. *Helicobacter pylori* infection. *N Engl J Med*. 2002;347:1175–86.
- Wirth T, Meyer A, Achtman M. Deciphering host migrations and origins by means of their microbes. *Mol Ecol*. 2005;14:3289–306.
- 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. *Mol Microbiol*. 1999;32:459–70.
- Li WH, Sadler LA. Low nucleotide diversity in man. *Genetics*. 1991;129(2):513–23.
- Yamaoka Y. *Helicobacter pylori* typing as a tool for tracking human migration. *Clin Microbiol Infect*. 2009;15:829–34.
- Suzuki R, Shiota S, Yamaoka Y. Molecular epidemiology, population genetics, and pathogenic role of *Helicobacter pylori*. *Infect Genet Evol*. 2012;12:203–13.
- Nagasawa S, Motani-Saitoh H, Inoue H, Iwase H. Geographic diversity of *Helicobacter pylori* in cadavers: forensic estimation of geographical origin. *Forensic Sci Int*. 2013;229:7–12.
- Ikegaya H. Geographical identification of cadavers by human parasites. *Forensic Sci Int Genet*. 2008;2:83–90.
- 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*. 2004;101:4746–51.
- Goh KL, Chan WK, Shiota S, Yamaoka Y. Epidemiology of *Helicobacter pylori* infection and public health implications. *Helicobacter*. 2011;16(Suppl 1):1–9.
- Eusebi LH, Zagari RM, Bazzoli F. Epidemiology of *Helicobacter pylori* infection. *Helicobacter*. 2014;19(Suppl 1):1–5.
- Rothenbacher D, Inceoglu J, Bode G, Brenner H. Acquisition of *Helicobacter pylori* infection in a high-risk population occurs within the first 2 years of life. *J Pediatr*. 2000;136(6):744–8.
- Weyermann M, Rothenbacher D, Brenner H. Acquisition of *Helicobacter pylori* infection in early childhood: independent contributions of infected mothers, fathers, and siblings. *Am J Gastroenterol*. 2009;104:182–9.
- Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, et al. Traces of human migrations in *Helicobacter pylori* populations. *Science*. 2003;299:1582–5.
- World Health Organization and Unicef. Progress on Sanitation and Drinking Water. Update and MDG assessment. Geneva: World Health Organization; 2015. p. 2015.
- Oleastro M, Rocha R, Vale FF. Population genetic structure of *Helicobacter pylori* strains from Portuguese-speaking countries. *Helicobacter*. 2017;22:1–10.
- Breurec S, Raymond J, Thiberge JM, Hem S, Monchy D, Seck A, et al. Impact of human migrations on diversity of *Helicobacter pylori* in Cambodia and New Caledonia. *Helicobacter*. 2013;18:249–61.
- Israel DA, Salama N, Krishna U, Rieger UM, Atherton JC, Falkow S, et al. *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. *Proc Natl Acad Sci U S A*. 2001;98:14625–30.
- Thorell K, Yahara K, Berthenet E, Lawson DJ, Mikhail J, Kato I, et al. Rapid evolution of distinct *Helicobacter pylori* subpopulations in the Americas. *PLoS Genet*. 2017;13:e1006546.
- Domínguez-Bello MG, Pérez ME, Bortolini MC, Salzano FM, Pericchi LR, Zambrano-Guzmán O, et al. Amerindian *Helicobacter pylori* strains go extinct, as European strains expand their host range. *PLoS One*. 2008;3:e3307.
- Yamaoka Y, Orito E, Mizokami M, Gutierrez O, Saitou N, Kodama T, et al. *Helicobacter pylori* in north and South America before Columbus. *FEBS Lett*. 2002;517:180–4.
- Ghose C, Perez-Perez GI, Dominguez-Bello M-G, Pride DT, Bravi CM, Blaser MJ. East Asian genotypes of *Helicobacter pylori* strains in Amerindians provide evidence for its ancient human carriage. *Proc Natl Acad Sci U S A*. 2002;99:15107–11.
- Shiota S, Suzuki R, Matsuo Y, Miftahussurur M, Tran TTH, Binh TT, et al. *Helicobacter pylori* from gastric cancer and duodenal ulcer show same phylogeographic origin in the Andean region in Colombia. *PLoS One*. 2014;9:e105392.
- Moodley Y, Linz B, Yamaoka Y, Windsor HM, Breurec S, Wu J-Y, et al. The peopling of the Pacific from a bacterial perspective. *Science*. 2009;323:527–30.
- Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. *Nature*. 1987;325:31–6.
- Stringer CB, Andrews P. Genetic and fossil evidence for the origin of modern humans. *Science*. 1988;239:1263–8.
- Cann RL. Genetic clues to dispersal in human populations: retracing the past from the present. *Science*. 2001;291:1742–8.
- Cavalli-Sforza LL, Feldman MW. The application of molecular genetic approaches to the study of human evolution. *Nat Genet*. 2003;33(suppl):266–75.
- Breurec S, Guillard B, Hem S, Brisse S, Dieye FB, Huerre M, et al. Evolutionary history of *Helicobacter pylori* sequences reflect past human migrations in Southeast Asia. *PLoS One*. 2011;6:e22058.
- Miftahussurur M, Sharma RP, Shrestha PK, Suzuki R, Uchida T, Yamaoka Y. Molecular epidemiology of *Helicobacter pylori* infection in Nepal: specific ancestor root. *PLoS One*. 2015;10:e0134216.
- Kodaman N, Pazos A, Schneider BG, Piazzuelo MB, Mera R, Sobota RS, et al. Human and *Helicobacter pylori* coevolution shapes the risk of gastric disease. *Proc Natl Acad Sci U S A*. 2014;111:1455–60.
- Shiota S, Cruz M, Abreu JAJ, Mitsui T, Terao H, Disla M, et al. Virulence genes of *Helicobacter pylori* in the Dominican Republic. *J Med Microbiol*. 2014;63(PART 9):1189–96.
- Central Intelligence Agency. The World Factbook. <https://www.cia.gov/library/publications/the-world-factbook/geos/dr.html>. Accessed 13 Sep 2017.
- Lalueza-Fox C, Calderón FL, Calafell F, Morera B, Bertranpetit J. MtDNA from extinct Tainos and the peopling of the Caribbean. *Ann Hum Genet*. 2001;65(Pt 2):137–51.
- Nagashima H, Iwatani S, Cruz M, Jiménez Abreu JA, Tronilo L, Rodríguez E, et al. Differences in interleukin 8 expression in *Helicobacter pylori*-infected gastric mucosa tissues from patients in Bhutan and the Dominican Republic. *Hum Pathol*. 2015;46:129–36.
- Levin BC, Cheng H, Reeder DJ. A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics*. 1999;55:135–46.
- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;290:457–65.
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 1999;23:147.
- Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:1870–4.
- Weissensteiner H, Pacher D, Kloss-Brandstätter A, Forer L, Specht G, Bandelt HJ, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res*. 2016;44:W58–63.
- van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 2009;30:386–94.
- Chaitanya L, Van Oven M, Weiler N, Hartevelde J, Wirken L, Sijen T, et al. Developmental validation of mitochondrial DNA genotyping assays for

- adept matrilineal inference of biogeographic ancestry at a continental level. *Forensic Sci Int Genet.* 2014;11:39–51.
45. Geppert M, Roewer L. SNaPshot® minisequencing analysis of multiple ancestry-informative Y-SNPs using capillary electrophoresis. *Methods Mol Biol.* 2012;830:127–40.
  46. Nell S, Eibach D, Montano V, Maady A, Nkwescheu A, Siri J, et al. Recent acquisition of *Helicobacter pylori* by Baka pygmies. *PLoS Genet.* 2013;9:e1003775.
  47. Linz B, Vololonantenainab CRR, Seck A, Carod JF, Dia D, Garin B, et al. Population genetic structure and isolation by distance of *Helicobacter pylori* in Senegal and Madagascar. *PLoS One.* 2014;9:1–7.
  48. Secka O, Moodley Y, Antonio M, Berg DE, Tapgun M, Walton R, et al. Population genetic analyses of *Helicobacter pylori* isolates from Gambian adults and children. *PLoS One.* 2014;9:e109466.
  49. Vale FF, Vadelu J, Oleastro M, Breurec S, Engstrand L, Perets TT, et al. Dormant phages of *Helicobacter pylori* reveal distinct populations in Europe. *Sci Rep.* 2015;5:14333.
  50. 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:e9645.
  51. Devi SM, Ahmed I, Francalacci P, Hussain MA, Akhter Y, Alvi A, et al. Ancestral European roots of *Helicobacter pylori* in India. *BMC Genomics.* 2007;8:184.
  52. Tay CY, Mitchell H, Dong Q, Goh K-L, Dawes IW, Lan R. Population structure of *Helicobacter pylori* among ethnic groups in Malaysia: recent acquisition of the bacterium by the Malay population. *BMC Microbiol.* 2009;9:126.
  53. Molina-Castro SE, Herrera D, Malespín-Bendaña W, Ramírez V, Une C. The geographic origin of *Helicobacter pylori* isolated from Costa Rican patients. *Gut Microbes.* 2014;5:517–21.
  54. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987;4(4):406–25.
  55. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol.* 1980;16:111–20.
  56. 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.
  57. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol Ecol Resour.* 2010;10:564–7.
  58. Salas A, Richards M, Lareu M-V, Scozzari R, Coppa A, Torroni A, et al. The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet.* 2004;74:454–65.
  59. Tishkoff SA, Gonder MK, Henn BM, Mortensen H, Knight A, Gignoux C, et al. History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. *Mol Biol Evol.* 2007;24:2180–95.
  60. Fagundes NJR, Kanitz R, Eckert R, Valls ACS, Bogo MR, Salzano FM, et al. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am J Hum Genet.* 2008;82:583–92.
  61. Barral-Arca R, Pischedda S, Gómez-Carballeda A, Pastoriza A, Mosquera-Miguel A, López-Soto M, et al. Meta-analysis of mitochondrial DNA variation in the Iberian Peninsula. *PLoS One.* 2016;11:e0159735.
  62. Adams SM, Bosch E, Balaresque PL, Ballereau SJ, Lee AC, Arroyo E, et al. The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet.* 2008;83:725–36.
  63. Tajima A, Hamaguchi K, Terao H, Oribe A, Perrotta VM, Baez CA, et al. Genetic background of people in the Dominican Republic with or without obese type 2 diabetes revealed by mitochondrial DNA polymorphism. *J Hum Genet.* 2004;49:495–9.
  64. Torroni A, Rengo C, Guida V, Cruciani F, Sellitto D, Coppa A, et al. Do the four clades of the mtDNA Haplogroup L2 evolve at different rates? *Am J Hum Genet.* 2001;69:1348–56.
  65. Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, et al. Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc Natl Acad Sci.* 2010;107(Suppl 2):8954–61.

## Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

