



Prevalence of *Helicobacter pylori* *babA*, *oipA*, *sabA*, and *homb* genes in isolates from Chinese patients with different gastroduodenal diseases

Qiaoyun Zhao^{1,2,3} · Conghua Song^{1,2,3,5} · Ke Wang⁴ · Donghong Li⁴ · Yang Yang^{1,2,3} · Dongsheng Liu^{2,3} · Le Wang^{2,3} · Nanjin Zhou⁴ · Yong Xie^{1,2,3}

Received: 30 January 2020 / Accepted: 16 March 2020 / Published online: 26 March 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Disease outcome is associated with virulence factors of *Helicobacter pylori* (*H. pylori*), which are partially attributed to the outer membrane protein (OMP). This study aimed to investigate the correlation between the four OMP genes (*babA*, *oipA*, *sabA*, and *homb*) and gastroduodenal diseases. One hundred and seventy-seven *H. pylori* strains were isolated from Chinese patients with different gastroduodenal diseases (49 chronic gastritis, 19 gastric ulcer, 33 gastric cancer, and 76 duodenal ulcer), 94 of which contained pathological information (41 superficial gastritis, 24 intestinal hyperplasia, and 29 gastric adenocarcinoma). The full-length amplification of *babA*, *oipA*, *sabA*, and *homb* genes was acquired and sequenced. Then, the genetic polymorphism was analyzed to compare with the reference strains from the GenBank database. Functional status and cluster analysis were also performed to evaluate the impact of genetic polymorphism on disease outcome. The prevalence of *babA*, *oipA*, *sabA*, and *homb* genes were 91.5%, 100%, 94.0%, and 95.5%, respectively. The four OMP genes were characterized by genetic polymorphism and in the status of positive selection ($Ka/Ks > 1$). The proportion of strains with functional status on for *oipA* and *sabA* gene was 100% and 76.2%, respectively. The sequences of four OMP genes were mainly clustered together with the East Asian references. The four OMP genes were not different in patients with gastroduodenal diseases and pathologic changes ($P > 0.05$). *H. pylori* *babA*, *oipA*, *sabA*, and *homb* genes were common in the Chinese populations, but did not seem to be involved in the development of gastroduodenal diseases.

Keywords *Helicobacter pylori* · Outer membrane protein · Gastroduodenal diseases · Gene sequence · Genetic polymorphism · Phylogeny

Introduction

Helicobacter pylori (*H. pylori*) is a slow-growing, micro-aerophilic, Gram-negative, spiral-shaped and flagellated bacterium, which colonizes the mucus layer of the gastric

epithelium [1]. Approximately one-half of the global human population was colonized by this organism, especially in developing countries [2]. Although *Helicobacter pylori* (*H. pylori*) infection can induce inflammatory responses of host cells, the majority of infected individuals remain asymptomatic throughout life. It is reported that *H. pylori*-infected individuals are having the risk of developing peptic ulcers in 15–20%, dyspepsia in 5–10%, gastric cancer in 1%, and primary gastric mucosa-associated lymphoid tissue lymphoma in 0.1% [3]. However, the mechanism of different disease outcomes caused by *H. pylori* infection remains unclear.

Currently, multiple different factors (i.e., pathogen, host, and environment) are all thought to account for the distinct clinical phenotypes. The differences in bacterial genotype and virulence intensity constitute the main points of pathogenic factors. *H. pylori* has been proved to be one of the most genetically diverse bacterial species known [4].

Edited by: Volkhard A. J. Kempf.

Qiaoyun Zhao and Conghua Song are joint first authors of this paper.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00430-020-00666-2>) contains supplementary material, which is available to authorized users.

✉ Yong Xie
xieyong_tfahoncu@163.com

Extended author information available on the last page of the article

Geographic differences in gastric cancer incidence could be partly attributed to the differences between *H. pylori* strains [5]. Biological diversity of isolates was found in patients with different diseases [6]. Polymorphisms of *H. pylori* reflect the geographic origin and correlate with the status of virulence factors [7]. Furthermore, the genetic heterogeneity was also identified in genes encoding outer membrane proteins (OMPs) [8]. Phylogenetic analysis revealed that OMP genes (i.e., *homB* and *homA*) were heterogeneously distributed worldwide, with a marked difference between East Asian and Western strains [9]. Therefore, the differential prevalence of *H. pylori* genes and the subsequent alterations in OMPs may play a vital role in the different clinical outcome of chronic infection [10].

The *H. pylori* genome is predicted to encode > 60 different OMPs, but most have not been studied in detail [8, 11]. In specific, the blood group antigen-binding adhesin A (*babA*) gene is generally associated with the presence of *H. pylori* adhesion [12]. Attachment to the gastric mucosa is the first step in establishing bacterial colonization of *H. pylori*, which is crucial for colonization and persistent infection. Subsequently, the continuous colonization of *H. pylori* mediates the chronic inflammatory response of gastric epithelial cells. The protein encoded by outer inflammatory protein A (*oipA*) was identified as a factor that induces a pro-inflammatory response [13]. In addition, the sialic acid-binding adhesin A (*sabA*) gene is important for the induction of inflammatory reaction in the gastric mucosa [14]. The helicobacter outer membrane B (*homB*) codes for a putative OMP and was shown to induce activation of interleukin-8 secretion in vitro, which was found to be associated with peptic ulcers [15].

It is known that *babA*, *oipA*, *sabA*, and *homB* genes widely exist in different strains of *H. pylori* [16]. The genetic polymorphism and the functional status of these four genes affect the expression of corresponding OMPs, which may be associated with the regional prevalence of gastroduodenal diseases. However, the available results about the relationship between the four OMP genes (*babA*, *oipA*, *sabA*, and *homB*) and gastroduodenal diseases are controversial. Based on similar studies [17–19], the inconsistent association between these OMP genes and disease outcomes may be related to geographic diversity, sample sizes and/or detection methods (e.g., PCR, western blot, incomplete sequencing). Herein, we examined more than one hundred *H. pylori* strains isolated from Chinese patients with different gastroduodenal diseases. The full-length amplification of *babA*, *oipA*, *sabA*, and *homB* genes was acquired and sequenced. The findings of *babA*, *oipA*, *sabA*, and *homB* gene in clinical outcomes would greatly benefit the study of pathogenesis, biomarkers, therapeutic targets, or vaccine candidates.

Materials and methods

Strains and genomic DNA extraction

One hundred and seventy-seven preserved *H. pylori* strains (Table S1) in this study were previously isolated from Chinese patients with different gastroduodenal diseases (49 chronic gastritis, 19 gastric ulcer, 33 gastric cancer, 76 duodenal ulcer), 94 of which contained pathological information (41 chronic superficial gastritis, 24 intestinal hyperplasia, and 29 gastric cancer) [20]. The strains were successfully cultured on Campylobacter agar base (OXOID Co., UK) plates containing 5% defibering sheep blood and selective antibiotic medium. Then, the plates were placed in a microaerobic incubator with a condition of 5% O₂ and 10% CO₂ at 37 °C. The identification of *H. pylori* was confirmed by colony morphology and Gram-Staining. Genomic DNA of isolates was extracted by the QIAamp DNA Mini Kit (QIAGEN Co., France) according to the manual instruction.

Amplification and detection of *babA*, *oipA*, *sabA*, and *homB* genes

Amplification of *babA*, *oipA*, *sabA*, and *homB* genes was performed with the multi-primer PCR assay. All amplifications were carried in a reaction volume of 25 µl mix reagents that contained 2 µl (20 ng/µl) of *H. pylori* genome DNA, 5 µl Prime STAR buffer, 2 µl (2.5 mM each) dNTP Mixture, 1 µl (10 µM) forward primer and 1 µl (10 µM) reverse primer, 0.5 µl Prime STAR polymerase, and 13.5 µl sterile water. The conditions of thermal cycler program used for *babA*, *oipA*, *sabA*, and *homB* consisted of the following five steps: initial denaturation at 95 °C for 5 min followed by 35 cycles of 45 s at 95 °C (denaturation), and 45 s at 64 °C for *babA*, 30 s at 62 °C for *oipA*, 30 s at 58 °C for *sabA*, 40 s at 60 °C for *homB* (all annealing steps), followed by 30 s at 72 °C (extension step) and a final extension step was 5 min at 72 °C. The information of PCR primer for each amplicon is summarized in Table 1. All amplification products were examined by 1.5% agarose gel electrophoresis (TransGen Biotech Co., China) and detected by ChemiDoc MP System (Bio-Rad, USA).

The full-length sequence of *babA*, *oipA*, *sabA*, and *homB* genes

Samples without target genes of *babA*, *oipA*, *sabA*, and *homB* were discarded at the identification stage of the PCR product. The positive PCR products of isolates were

Table 1 Primers used for the detection of *H.pylori oipA*, *babA*, *sabA*, and *homB* genes

Genes	Primer	Sequence (5'–3')	Products size (bp)
<i>oipA</i>	Forward	CGCACTTCTTCTATAGGCACATTCGC	1400
	Reverse	CATCAAAGAAACGCACACCTGCTATCA	
<i>babA</i>	Forward	ATGAAAAAACACATCCTTTCATTAGCT	2202
	Reverse	TTAGTAAGCGAACACATAGTTCAAATACAC	
<i>sabA</i>	Forward	TTTGCGAGCGTTTCGGCGTTGT	2539
	Reverse	CTTTGGCTCTAGCAATGTGTGGC	
<i>homB</i>	Forward	TTTCGCCCAAATACCCTACTGA	2834
	Reverse	GAAGTGCCGCCATAACATG	

purified and then the DNA samples were submitted and performed as a service by the Beijing Genomics Institute (BGI Tech., China) for the full-length sequence of *babA*, *oipA*, *sabA*, and *homB* genes.

Analysis of genetic polymorphism and the functional status

Sequences of *babA*, *oipA*, *sabA*, and *homB* genes were automatically and manually aligned by ClustalW in MEGA v7.0 software (MEGA Inc., USA). DnaSP v6.0 software (Julio Rozas & Universitat de Barcelona, Spain) was used to conduct analyses of genetic polymorphism. The nonsynonymous (Ka) to synonymous (Ks) substitution rate ratio ($\omega = Ka/Ks$) can be used as an estimator for selective pressure on DNA sequence evolution (negative/purifying selection $Ka/Ks < 1$, neutral selection $Ka/Ks = 1$ and positive/adaptive selection $Ka/Ks > 1$) [21]. The ratios of Ka/Ks and Pi(a)/Pi(s) were evaluated by the modified Nei Gojori method and the Jukes–Cantor correction in MEGA to evaluate the evolutionary selection. In addition, parameters involving haplotypes (H), haplotype diversity (Hd), nucleotide diversity (Pi), nucleotide differences (K), and Tajima'D test were obtained to estimate the functional status of the four OMP genes. The presence of an early stop codon (TAA/TGA/TAG) in the 5' region of sequences (signal peptide function fragments) means that the functional process of the gene to protein translation is turned off. Also, expression of *oipA* and *sabA* genes is regulated by the slipped-strand repair mechanism based on the number of CT dinucleotide repeats in the 5' region of the gene (switch on = functional, switch off = non-functional) [22].

Phylogenetic analysis of *babA*, *oipA*, *sabA*, and *homB* genes

To compare the differences between the strains, a total of 27 reference strains with complete genetic sequence (known sequences listed in the GenBank database) and geographic information were obtained by reviewing the literature [23,

24]. The detailed information of these reference strains were as follows: hspEAsia: 51, 52, 83, 35A, F16, F32, F57; hspAmerind: cuz20, Pecan4, Puno135, sat464, shi169, shi417, shi470, v225d; hpEurope: 26695, B8, HPAG1, G27, Lithuania75, SJM180; hspWAfrica: J99, Pecan18; hspSouthIndia: India7, SNT49; hpAfrical: SouthAfrica7, SouthAfrica20. We performed the phylogenetic analysis of *babA*, *oipA*, *sabA*, and *homB* genes in isolates from two aspects: different gastroduodenal diseases (177/177 isolates) and different pathological parameters (94/177 isolates). The neighbor-joining method was used and bootstrap analysis was performed with 1000 replications to reconstruct the phylogenetic tree in MEGA. To enhance the image contrast, the phylogenetic tree was ulteriorly edited by online tools (Evolview, <http://www.evolgenius.info/evolview>).

Statistical analysis

The one-way analysis of variance, Chi-square test or Fisher's exact test was employed for the comparison of the corresponding variables across groups in IBM SPSS 21.0 statistics software (IBM Corp., USA). A test result was two-sided where a *P* value of < 0.05 was considered statistically significant.

Results

Prevalence of *babA*, *oipA*, *sabA*, and *homB* genes

The prevalence of *babA*, *oipA*, *sabA*, and *homB* genes in 177 isolates from Chinese patients with different gastroduodenal diseases (49 chronic gastritis, 19 gastric ulcer, 33 gastric cancer, and 76 duodenal ulcer) was 91.5%, 100%, 94.0%, and 95.5%, respectively. Besides, the prevalence of *babA*, *oipA*, *sabA*, and *homB* genes in 94 isolates with different pathological information (41 superficial gastritis, 24 intestinal hyperplasia, and 29 gastric adenocarcinoma) was 92.5%, 100%, 89.4%, and 97.9%, respectively. However, neither gastroduodenal disease nor pathological parameters were

related to the prevalence of these four OMP genes (*babA*, *oipA*, *sabA*, and *homB*) (all $P > 0.05$).

Genetic polymorphism and evolutionary selection

There were 250 (27.4%) variable sites found in the 177 *oipA* sequence with 912 bases, which contained 141 (15.5%) parsimony informative sites and 109 (11.95%) singleton variable sites. Also, the variable sites were found in *babA*, *sabA*, and *homB* was 55.3%, 43.94%, and 43.16%, respectively. The Pi of the *oipA* gene among Chinese

isolates was 0.023, which was far lower than that in the reference strains (0.271). The average K of the *oipA* gene was also lower in the Chinese isolates than that in the reference strains (20.82 vs. 246.13). The details of these parameters in the remaining three genes (*babA*, *sabA*, and *homB*) are shown in the Table 2. In addition, the Ka/Ks ratios were calculated to identify the status of *babA*, *oipA*, *sabA*, and *homB* genes. In general, all four OMP genes were in the status of positive selection ($Ka/Ks > 1$). As shown in Table 3, neither genetic polymorphism nor

Table 2 Polymorphic analysis of *oipA*, *babA*, *sabA*, and *homB* genes

Gene	Stains	N	Pi	K	Mean Ka	Mean Ks	Ka/Ks
<i>oipA</i>	China	177	0.023	20.820	0.063	0.061	1.033
	Ref.	28	0.271	246.130	0.698	0.554	1.260
	Total	205	0.061	55.205	0.139	0.136	1.022
<i>babA</i>	China	146	0.061	118.219	0.068	0.051	1.333
	Ref.	27	0.103	219.963	0.188	0.118	1.593
	Total	173	0.070	133.886	0.078	0.06	1.300
<i>sabA</i>	China	157	0.105	189.791	0.133	0.131	1.015
	Ref.	21	0.070	126.981	0.108	0.095	1.137
	Total	178	0.107	194.258	0.136	0.135	1.007
<i>homB</i>	China	169	0.073	133.801	0.116	0.092	1.261
	Ref.	15	0.046	86.314	0.074	0.063	1.175
	Total	184	0.071	131.354	0.095	0.09	1.056

The polymorphic analysis was performed by DnaSP software; Pi: nucleotide diversity; K: average number of nucleotide differences; Ka: the number of nonsynonymous substitutions per nonsynonymous site; Ks: the number of synonymous substitutions per synonymous site

Table 3 Polymorphism of *oipA*, *babA*, *sabA*, and *homB* genes and its correlation with pathological parameters

Gene	Disease	n (%)	Pi	P_1^{Δ}	Ka/Ks	P_2^{Δ}
<i>oipA</i>	CSG	41 (100)	0.023	0.486	1.033	0.486
	IM	24 (100)	0.024	0.474	1.123	0.474
	GC	29 (100)	0.021	0.479	1.127	0.479
	Total	94 (100)	0.023	Ref.	1.197	Ref.
<i>babA</i>	CSG	36 (87.8)	0.068	0.485	1.194	0.485
	IM	23 (95.8)	0.068	0.475	1.106	0.475
	GC	25 (86.2)	0.071	0.477	1.167	0.477
	Total	84 (100.0)	0.068	Ref.	1.156	Ref.
<i>sabA</i>	CSG	39 (95.1)	0.110	0.486	1.062	0.486
	IM	23 (95.8)	0.111	0.474	1.011	0.474
	GC	24 (82.8)	0.109	0.476	1.005	0.476
	Total	86 (100.0)	0.108	Ref.	1.000	Ref.
<i>homB</i>	CSG	41 (97.9)	0.080	0.486	1.012	0.486
	IM	24 (100)	0.075	0.475	1.019	0.475
	GC	27 (100)	0.068	0.478	1.107	0.478
	Total	92 (100)	0.075	Ref.	1.052	Ref.

The polymorphic analysis was performed by DnaSP software; CSG: chronic superficial gastritis; IM: intestinal metaplasia; GC: gastric cancer; Pi: nucleotide diversity; Ka: the number of nonsynonymous substitutions per nonsynonymous site; Ks: the number of synonymous substitutions per synonymous site; Δ : single sample *t* test

evolutionary selection of the four OMP genes was related to the pathological parameters (all $P > 0.05$).

The functional status of *babA*, *oipA*, *sabA*, and *hombB* genes

In all sequences, *oipA* showed an on functional status and 12 repeated patterns of CT dinucleotide (Table 4). In addition, the repeated pattern of CT dinucleotide in the coding region was found and classified. In 166 *sabA* sequences, 128 (77.1%) of which had a switch on status while 38 (22.9%) had an off status. Furthermore, the repeated pattern of CT dinucleotide varied in the sequences which were different from *oipA* (Table 5). In 147 *babA* sequences, 132 (89.8%) of which had an on functional status while 15 (10.2%) had an off status. The function state on and off were 161 (95.3%) and 8 (4.7%) respectively in 169 *hombB* sequences. No repeated pattern of CT dinucleotide or other regular pattern was found in both *babA* and *hombB* sequences with functional status off in which the sequences mainly showed irregular single nucleotide polymorphism.

Phylogenetic analysis of four OMP genes

Phylogenetic analysis based on the full-length sequence of *oipA*, *babA*, *sabA*, and *hombB* genes was performed respectively to distinguish the correlation between isolates and gastroduodenal diseases. Among which, phylogenetic tree showed that strains with the *oipA* gene was clustered together containing six clades (Fig. 1): three main clades including most Chinese strains and all HpAsia clades (as Asian clade); one strain from patient with duodenal ulcer solely clustered in an independent clade; two strains were overlapped with western strains (as Western clade); hspAmerind strains grouped into the last clade (as hspAmerind clade). The phylogenetic tree of the *babA* gene (Fig. 2) showed a similar feature as *oipA*. However, both the phylogenetic tree of the *sabA* and the *hombB* displayed a different pattern compared with the *oipA*. The phylogenetic tree of the *sabA* gene was mainly grouped into three clades (Fig. 3): the first clade was a sole strain isolated from a patient with gastric cancer; the second clade included half of the Chinese strains; the third clade included a few Chinese strains and

Table 4 The switch status of the signal sequence-coding region for the *oipA* gene

The sequence of the signal peptide coding region	CT pattern	Switch	n (%)
CTAA present^a			
ATGAAAAAAGCTCTCTTACTAATTCTCTTTTTCTCGTTTTGGCTCCACGCTGAAAGGAATGGAT TTTATTTA	2+1	on	1 (0.56)
ATGAAAAAAGCTCTCTTACTAACTCTCTTTTTCTCGTTTTGGCTCCACGCTGAAAGGAATGGAT TTTATTTA	3+1	on	109 (61.58)
ATGAAAAAAGCTCTCTTACTAACTCTCTTTCTCTCGTTTTGGCTCCACGCTGAAAGGAATGGAT TTTATTTA	3+2	on	5 (2.82)
ATGAAAAAAGCTCTCTTACTAACCTCTCTCTTTCTCGTTTTGGCTCCACGCTGAAAGGAATGGAT TTTATTTG	4+1	on	3 (1.69)
ATGAAAAAAGCTCTCTTACTAACTCTCTCTCTTTTTTTTTGGCTCCACGCTGAAAGGAATGG ATTTTATTTA	5	on	5 (2.82)
ATGAAAAAAGCTCTCTTACTAACTCTCTCTCTCTTTTTTTTTGGCTTCATGCTGAAAGGAATGGAT TTTATTTG	6	on	2 (1.13)
CTAA absent^b			
ATGAAAAAAACCCTTTTACTTTTTCTGTCTCTCTCGTTTTGGCTTCATGCTGAAAGGAATGGAT TTTATTTA	1+1+3	on	1 (0.56)
ATGAAAAAAACCCTTTTACTCTTTTTGTCTTTCTCGTTTTGGCTCCACGCTGAAAGGAATGGAT TTTA	2+1+1	on	1 (0.56)
ATGAAAAAAACCCTTTTACTCTTTCTGTCTTTCTCGTTTTGGCTCCACGCTGAAAGGAATGGAT TTTA	2+1+1+1	on	36 (20.34)
ATGAAAAAAACCCTTTTACTCTTTCTGTCTCTCTCGTTTTGGCTCCACGCTGAAAGGAATGGA TTTTA	2+1+3	on	8 (4.52)
ATGAAAAAAACCCTTTTACTCTTTCTCTTTTTCTCGTTTTGGCTCCACGCTGAAAGGAATGGAT TTTA	2+2+1	on	4 (2.26)
ATGAAAAAAACCCTTTTACTCTTTCTCTCTTTCTTGTTTTTGGCTCCACGCTGAAAGGAATGGAT TTTA	2+3+1	on	2 (1.13)

Switch: switch on = functional; switch off = non-functional (encoding protein)

^a4-base sequence (CTAA) was found in the upstream of CT repeats which began at position + 23

^b4-base sequence (CTAA) was not found in the upstream of CT repeats which began at position + 19

Table 5 The switch status of the signal sequence-coding region for the *sabA* gene

The sequence of the signal peptide coding region	CT pattern	Switch	n (%)
ATGAAAAATATATTGATATGGTTTATTAATAATTTGTTTATCTTGGTTCTAGCTATTCCTATTTTGCTC TTTGGTGTTTTGGCATTATTGG	None	On	1 (0.63)
ATGAAAAAACGATTTTACTTTGTCTATCCC-----TTGCGGCTTCATTGCTTTGTGCTG AAGACAACGGCTTTTTTGTGA	1+1	On	2 (1.25)
ATGAAAAAACGATTTTACTTTCTCTATCCC-----TTATGGCGTCATTGCTTTATGCTG AAGACAATGGCTTTTTTATGA	1+2	On	89 (55.6)
ATGAAAAAACGATTTTACTTTCTCTATCCC-----TTATGGCGTCATTGCTTTATGCTG AAGACAATGGCTTTTTTATGA	1+2	Off	3 (1.88)
ATGAAAAACCGATTTTACTCTCGCTTGGCG-----TACTATCGCTCTACGCTGAAG ACAACGGCTTTTTTGTGA	2+1	On	1 (0.63)
ATGAAAAAACGATTTTACTCTCTATATCCC-----TTGCGGCATCATTGCTTTGCGCTG AAGACAACGGCTTTTTTGTGA	3	On	1 (0.63)
ATGAAAAAACGATTTTACTTTCTCTCTCGC-----TTGCGGTATCATCGCTCCACGCTG AAGACAACGGCTTTTTTGTGA	1+3	On	2 (1.25)
ATGAAAAAGACAATTTACTCTCTCTCGC-----TTCATCGCTCTGACGCTGA AGACAATGGCTTTTTTGTGA	4	On	10 (6.25)
ATGAAAAAGACAATTTACTCTCTCTCGC-----TTGCGGTATCATCGCTCCACGCTG AAGACAATGGCTTTTTTGTGG	4	Off	1 (0.63)
ATGAAAAAACGATTTTACTCTCTCTCTCGC-----TTGCGGTATCATCGCTCCACACTG AAGACAACGGCTTTTTTGTGA	5	On	3 (1.88)
ATGAAAAAGACAATTTACTCTCTCTCTCGC-----TTCATCGCTCTTGCATGCTGA AGACAACGGCTTTTTTGTGA	5	Off	6 (3.75)
ATGAAAAACAATTTTACGCTCTCTCTCTCGC-----TTCATCGCTCTTGCACGCTGA AGACAACGGCTTTTTTATGA	6	on	1 (0.63)
ATGAAAAAACCTTTTACTCTCTCTCTCTCGC-----TTCATCGCTTTTAAACGCTGA AGACAACGGCTTTTTTGTGG	6	Off	7 (4.4)
ATGAAAAAGACAATTTACTCTCTCTCTCTCGC-----TTCATCGCTCTTGCACGCTGA AGACAACGGCTTTTTTGTGA	7	On	4 (2.5)
ATGAAAAAACGATTTTACTCTCTCTCTCTCTCGC-----TTGCGGTATCATCGCTCCACGCTG AAGACAACGGCTTTTTTGTGA	7	Off	1 (0.63)
ATGAAAAAACGATTTTACTTTCTCTCTCTCTCTCGC-----TTGCGGTATCATCGCTCCATGCTG AAGACAATGGCTTTTTTGTGA	1+6	On	1 (0.63)
ATGAAAAAACGATTTTACTTTCTCTCTCTCTCTCGC-----TTGCGGTATCATCGCTCCACGCTG AAGACAACGGCTTTTTTGTGA	1+6	Off	2 (1.25)
ATGAAAAAACGATTTTACTCTCTCTCTCTCTCTCGC-----TTGCGGTATCATCGCTCCACGCTG AAGACAACGGCTTTTTTGTGA	8	On	2 (1.25)
TTGAAAAAAGAGTTTATTCTCTCTCTCTCTCTC-----GCTTGGCGGTATCATCGCTCCACGCTGA AGACAATGGCTTTTTTGTGA	8	Off	6 (3.75)
ATGAAAAAGACAATTTACTCTCTCTCTCTCTCTC-----GCTTCATCGCTTTTGGCGGCT GAAGACAATGGCTTTTTTGTGA	1+8	Off	2 (1.25)
ATGAAAAACAATTTTACTCTCTCTCTCTCTCTCTCGC-----GGTATCATCGCTCCACGCTGA AGACAACGGCTTTTTTGTGG	9	On	2 (1.25)
ATGAAAAAACGATTTTACTCTCTCTCTCTCTCTCTCGC-----TTCATCGCTCTTGCACGCTGA AGACAATGGCTTTTTTATGA	9	Off	8 (5.0)
ATGAAAAAACGATTTTACTTTCTCTCTCTCTCTCTCTCTCGCGTGCGGTATCATCGCTCCACGCTG AAGACAGTGGCTTTTTTGTGA	1+9	On	1 (0.63)
ATGAAAAACAATTTTACTCTCTCTCTCTCTCTCTCTCGC-----TTCATCACTCTTGCACGCTGA AGACAACGGCTTTTTTGTGA	10	On	3 (3.75)
ATGAAAAAACGATTTTACTCTCTCTCTCTCTCTCTCTCGC---GTGCGGTATCATCGCTCCGTGCTG AAGACAACGGCTTTTTTGTGA	10	Off	1 (0.63)

-----: The termination codon was not displayed for a long sequence

Switch: switch on = functional; switch off = non-functional (encoding protein)

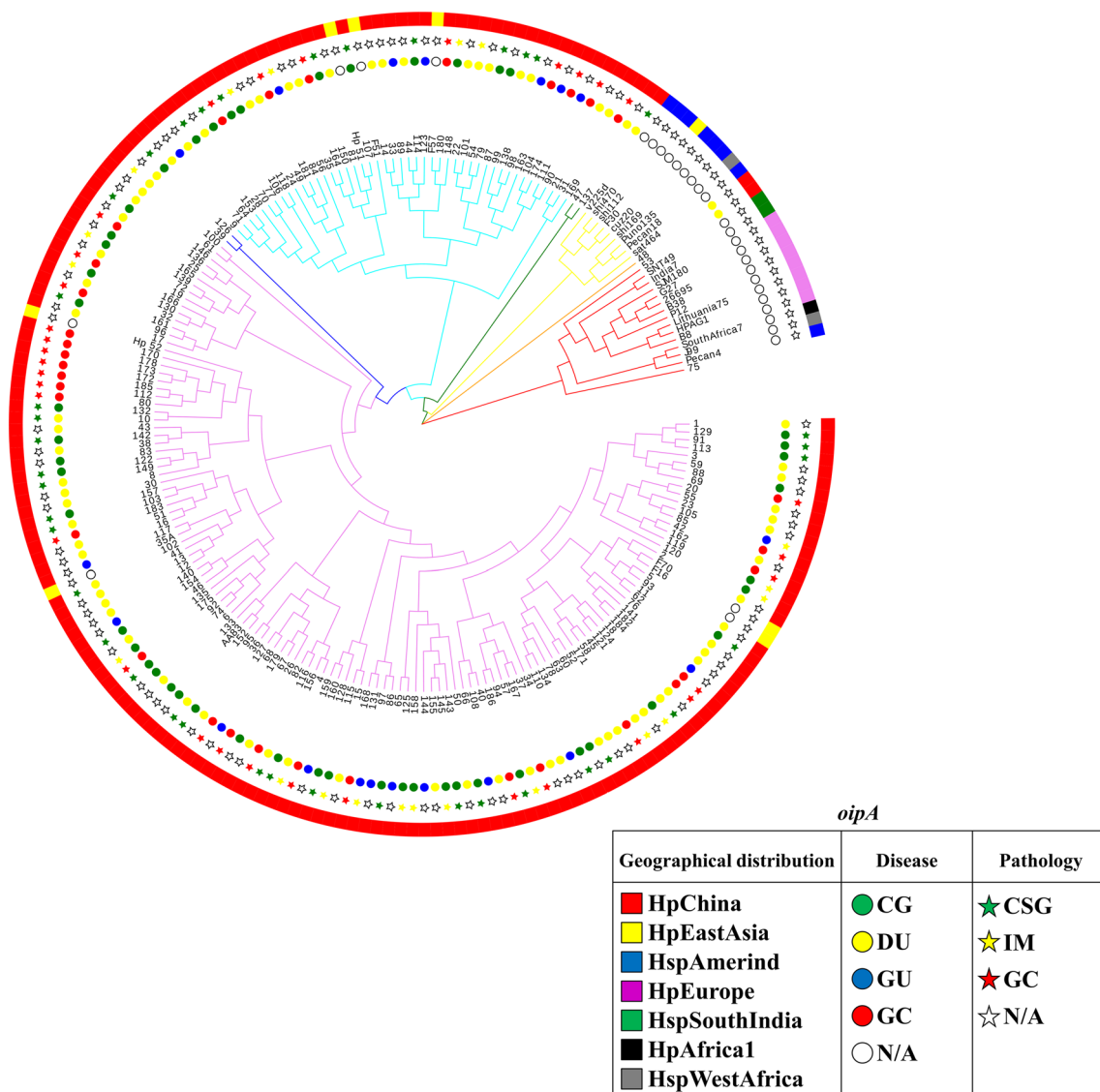


Fig. 1 Phylogenetic analysis of *oipA* gene sequence. HpChina: isolates from Chinese patients with different clinical information and pathological parameters; the others: reference strains from the Gen-

Bank database; CG: chronic gastritis; GU: gastric ulcer; DU: duodenal ulcer; CSG: chronic superficial gastritis; IM: intestinal metaplasia; GC: gastric cancer

all reference strains. Similarly, the phylogenetic tree of the *homB* gene (Fig. 4) appeared in the same characters as the *sabA* gene. However, the sequences with different disease information were not clustered together as expected in the phylogenetic analysis.

Discussions

It was reported that some OMPs play a vital role in the pathogenesis of *H. pylori* [8, 25]. Moreover, OMPs could cooperate with other virulent factors to lead diverse disease outcomes including gastric cancer. However, the strains in these studies were isolated from different geographical

regions and the genes of the four outer membrane proteins were investigated in different ways. The reported results from the PCR analysis were the most common, but the correlation between OMP genes (i.e., *babA*, *oipA*, *sabA*, and *homB*) and different gastroduodenal diseases was controversial [26, 27]. The possible reason was that these genes had not been sequenced at full length and the relationship between single nucleotide polymorphisms and different diseases had not been well analyzed.

The full-length sequence of *babA*, *oipA*, *sabA*, and *homB* genes was detected respectively in our study with the multiprimer PCR assay, and the result showed that the prevalence of *H. pylori* *babA*, *oipA*, *sabA*, and *homB* genes in isolates from Chinese patients was 91.5%, 100%, 94.0%, and 95.5%,

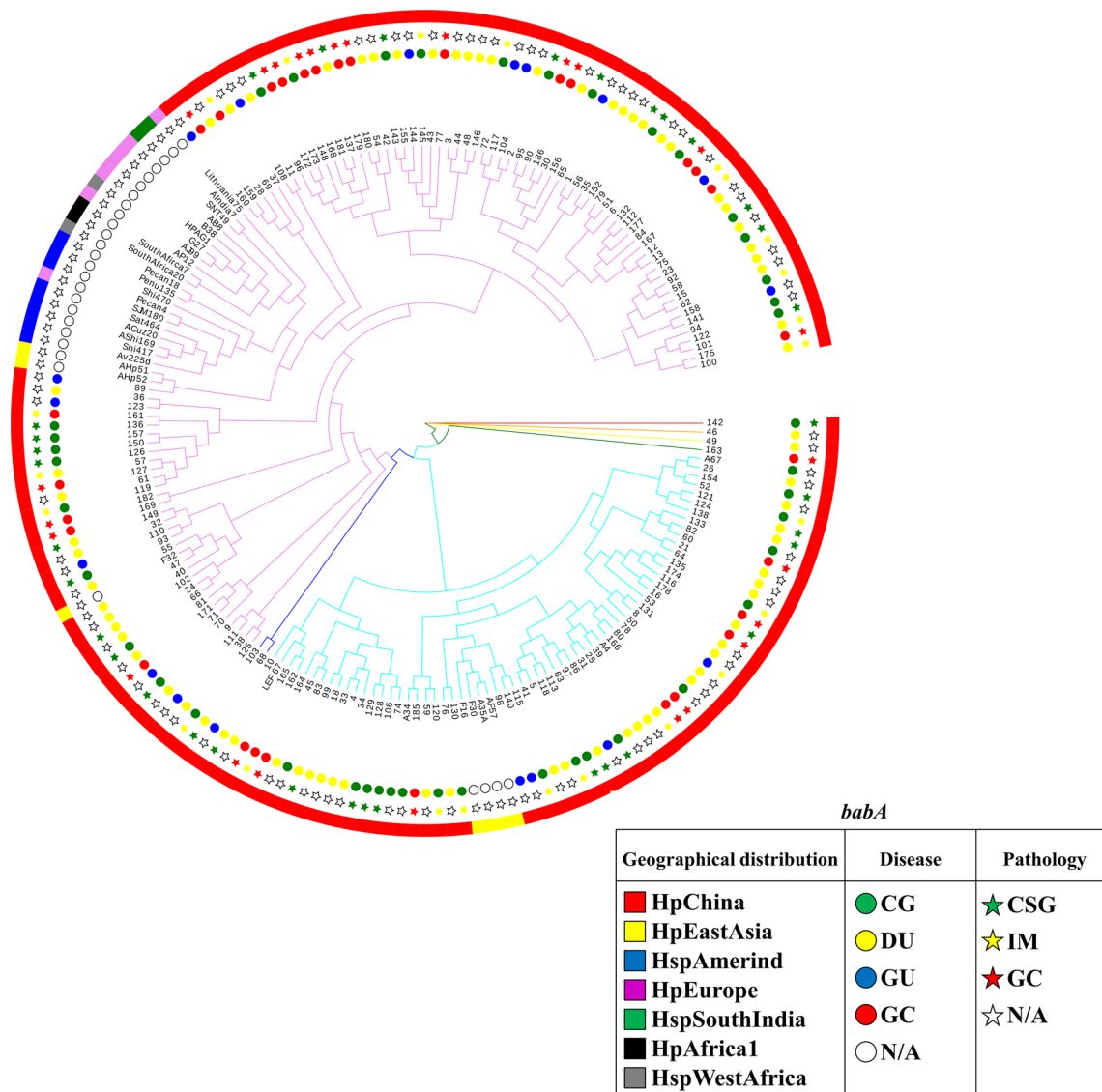


Fig. 2 Phylogenetic analysis of *babA* gene sequence. HpChina: isolates from Chinese patients with different clinical information and pathological parameters; the others: reference strains from the Gen-

Bank database; CG: chronic gastritis; GU: gastric ulcer; DU: duodenal ulcer; CSG: chronic superficial gastritis; IM: intestinal metaplasia; GC: gastric cancer

respectively. The results were largely consistent with literature reported from some Asian countries (> 90%, South Korea, Japan) [28, 29], but higher than those from Western countries (< 80%, America, Germany) [17, 19, 30]. The prevalence of gastric cancer varies with geographic area and a high incidence of gastric cancer in East Asians countries may be explained, at least in part, by the differences in genotypes of *H. pylori* strains [5]. Therefore, the differences in strains could be further explained by the different prevalence of *H. pylori* OMP genes.

The high prevalence of *babA*, *oipA*, and *sabA* genes had an active influence over the correlation between *H. pylori* and gastroduodenal diseases [18, 31]. Yamaoka et al. [17] found that *oipA* was associated with gastritis, peptic ulcer,

and gastric cancer, while *sabA* was associated with gastric cancer and intestinal metaplasia in American. Gerhard et al. [32] found that *babA* was associated with peptic ulcer and gastric cancer in German. The *homB* gene might be susceptible to gastric cancer development in American and Colombian and contribute to the determination of clinical outcomes in patients from East Asia and Western countries [33]. However, not all relationships between the four OMP genes and gastroduodenal diseases were consistent with the above results [19, 34–36]. Similar to the latter, no correlation between the prevalence of *babA*, *oipA*, *sabA*, and *homB* genes and gastroduodenal diseases with pathological parameters was observed in the present study ($P > 0.05$).

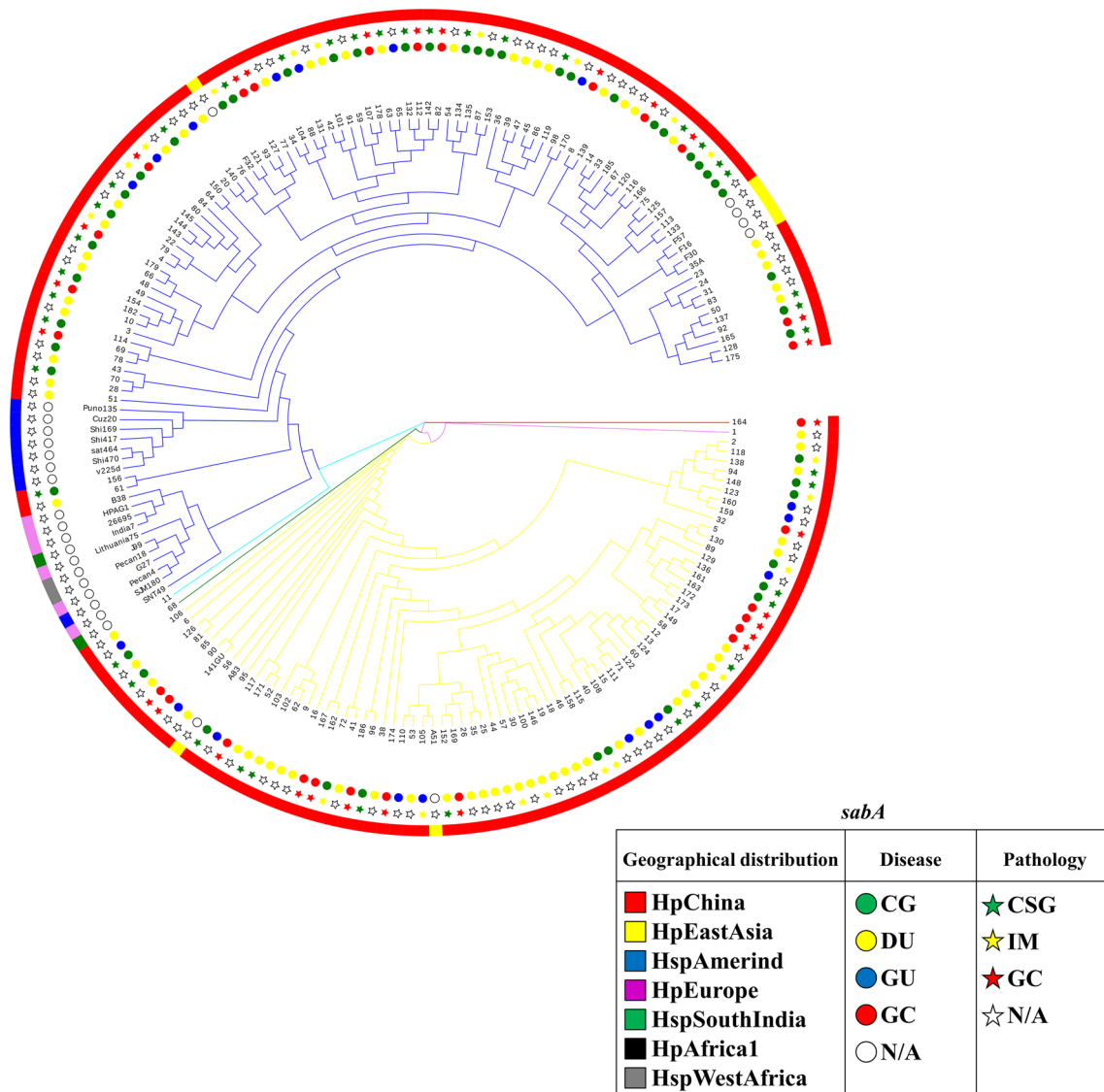


Fig. 3 Phylogenetic analysis of *sabA* gene sequence. HpChina: isolates from Chinese patients with different clinical information and pathological parameters; the others: reference strains from the Gen-

Bank database; CG: chronic gastritis; GU: gastric ulcer; DU: duodenal ulcer; CSG: chronic superficial gastritis; IM: intestinal metaplasia; GC: gastric cancer

It was necessary to analyze genetic diversity and evolutionary selection to further identify the relationship between OMP genes and gastroduodenal diseases. Compared with the *oipA* gene ($P_i < 0.1$), the other three OMP genes (*babA*, *sabA*, and *homB*) had higher genetic diversity ($P_i > 0.1$). The genetic diversity of *oipA* and *babA* genes was lower when compared with the reference sequence, while that of *sabA* and *homB* genes was higher. *H. pylori* strains with genetic polymorphism have an extraordinary capability to survive under intense selective pressure to colonize in hostile gastric conditions [37, 38]. *H. pylori* genes were highly divergent between East Asian and non-Asian strains, and 86% among which exhibit a positive selection [39]. Our results showed that all four OMP genes were in the status of

positive selection ($K_a/K_s > 1$). A study found that the Asian strain was more evolutionarily capable than the European strain, and suggested that the number of genes in a positive selection state was proportional to the stress of adaptation [40]. These results indicated that the pathogenic factors of *H. pylori* were in a positive state of evolutionary selection, thus gaining the ability to adapt and survive in gastric mucosa.

We also considered all mutation events and predicted functional states at the genetic level, including one or more non-synonymous mutations. Most switch status of the four OMP genes in this study was on > 90 . The switch status of the gene may affect bacterial characteristics such as virulence [22]. Protein OipA was identified as a factor that induces a pro-inflammatory response. The functional

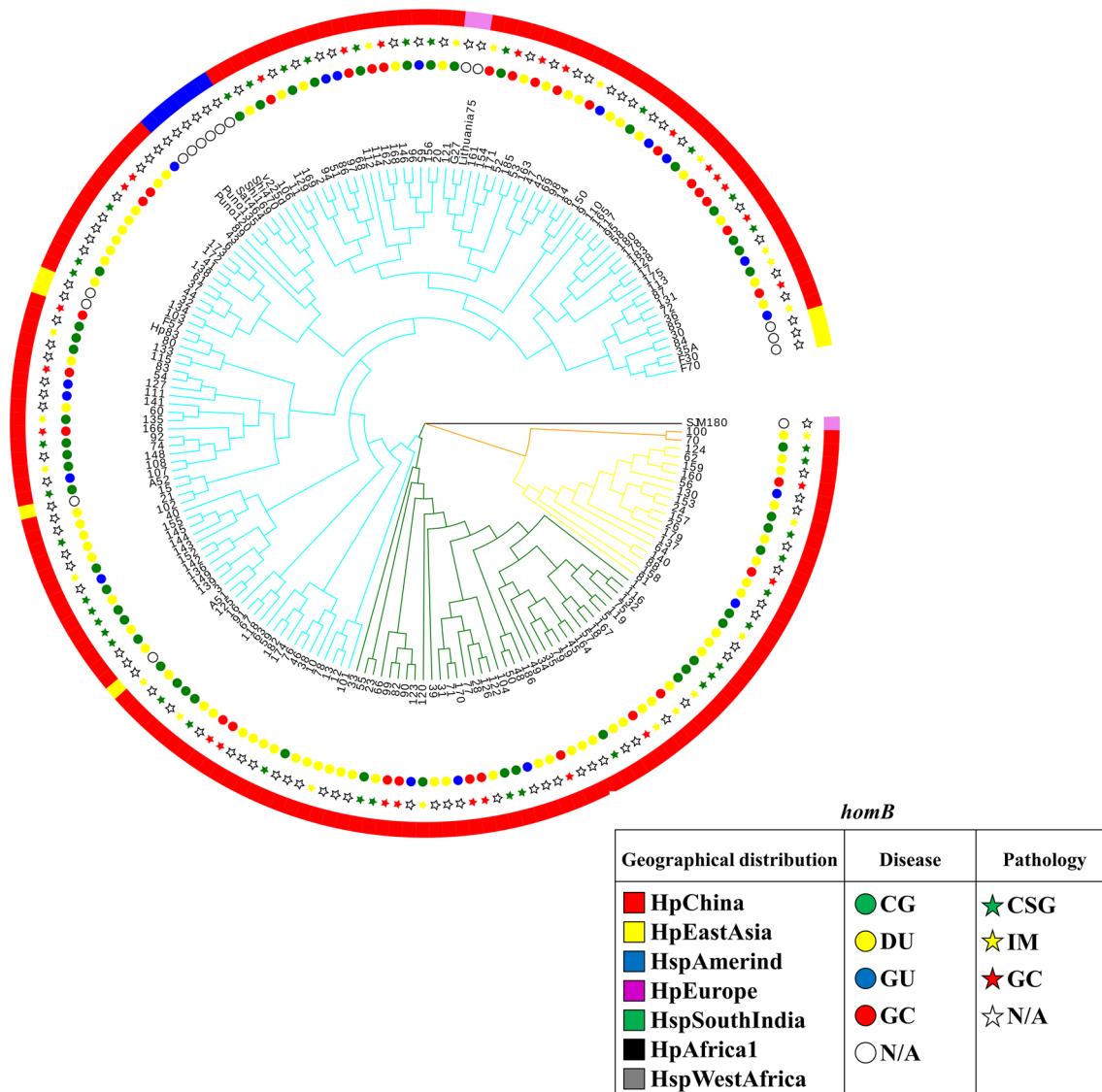


Fig. 4 Phylogenetic analysis of *homB* gene sequence. HpChina: isolates from Chinese patients with different clinical information and pathological parameters; the others: reference strains from the Gen-

Bank database; CG: chronic gastritis; GU: gastric ulcer; DU: duodenal ulcer; CSG: chronic superficial gastritis; IM: intestinal metaplasia; GC: gastric cancer

status of the *oipA* gene was relevant to gastritis and gastric cancer. The sequence of *oipA* and *sabA* was regulated by the number of CT dinucleotide repeats in the N-terminal signal peptide coding region [41, 42]. All strains in this study had the functional status for the *oipA* gene, which was consistent with the results from Japanese strains. A total of 13 different CT patterns were found in 177 sequences for the *oipA* gene, of which no more than 6 CT repeats were found and the most prevalent CT pattern was 3 + 1. Ando et al. [7] studied 109 strains from nine countries, and found that 67% (31/46) strains from East Asia have no more than 6 CT repeats, while 70% (23/33) strains from Western countries showed a 6 to 12 CT repeats. The results reveal the difference and uniqueness in CT numbers

between East and Western, implying a geographical distribution characteristic of *oipA* gene.

Unlike *oipA* gene, various CT patterns of *sabA* gene were detected in 166 sequenced genes. The results suggested that the *sabA* gene was of great high flexibility to switch on/off status. It is reported that the *sabA* gene is of great flexibility and diversity. A study from the Taiwanese showed that 80% of strains were *sabA* gene-positive but only 31% expressed SabA protein [43]. SabA was first found to mediate binding to the sialyl-Lewis antigen whereas this interaction was weaker than BabA-Leb adherence. Both the changeable on/off switching and weakness binding to sLex may benefit *H. pylori* to escape from the host's tough immune response, leading to chronic inflammatory response and subsequent

gastric mucosal lesions by successfully colonizing in the stomach for a long time.

The phylogenetic tree of the four OMP gene sequences containing different disease information was constructed respectively in this study, but no clustering of similar disease was observed. The results indirectly showed that none of the four OMP genes could effectively distinguish the differences between gastroduodenal diseases. *H. pylori* has gradually evolved into different geographical clusters with the development of human evolution and society [39]. Together with reference strains from other regions of the world, both the phylogenetic tree of *oipA* and *babA* genes showed a characteristic of the regional cluster. However, the clustering results of *sabA* and *homB* genes were relatively dispersed. In addition, almost all Chinese strains and East Asian strains were clustered together and could be classified as the main branch. The four OMP genes showed multiple subbranches in the clustering of Chinese and East Asian strains, which seemed to have more evolutionary diversity than the reference strains from the West. The evolutionary diversity of OMP genes may be related to the regional environment so that *H. pylori* can better adapt to the living conditions.

Our study also had some limitations that could not be ignored. First, we were unable to determine the sequential order of isolated strains and gastroduodenal diseases. Second, there was no homologous pathological information about duodenal ulcer in our sample. Finally, we could not get the strains without diseases due to pathogenic characteristics of *H. pylori*, but with chronic superficial gastritis instead in direct comparison.

Conclusion

BabA, *oipA*, *sabA*, and *homB* of *H. pylori* strain isolated from Chinese patients were highly polymorphic and in the status of positive selection ($K_a/K_s > 1$), but not a significant association with gastroduodenal diseases and pathologic changes in the present study. Nevertheless, more potential OMPs that could modulate interactions between *H. pylori* and gastroduodenal diseases were deserved further investigation.

Author contributions YX conceived and designed research. QZ and CS conducted the experiment and wrote the manuscript. KW, DL, YY, and NZ contributed new reagents and analytical tools. DL and LW contributed to the source of strains. All authors read and approved the manuscript.

Funding Funding was provided by the National Key Research and Development Program of China (No. 2016YFC1302201), National Natural Science Foundation of China (Nos. 81970502, 81860107, and 81260076), Leading Talent Training Plan of the Gan-Po Outstanding Talents 555 Project of Jiangxi Province (2010-3-61), Natural Science

Foundation of Fujian Province of China (No. 2017J01347), and Special Funds for Graduate Student Innovation of Jiangxi Province (No. YC2017-B024).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Fagoonee S, Pellicano R (2019) *Helicobacter pylori*: molecular basis for colonization and survival in gastric environment and resistance to antibiotics. A short review. *Infect Dis* 51(6):399–408. <https://doi.org/10.1080/23744235.2019.1588472>
2. Hooi JKY, Lai WY, Ng WK, Suen MMY, Underwood FE, Tanyingoh D, Malfertheiner P, Graham DY, Wong VWS, Wu JCY, Chan FKL, Sung JJY, Kaplan GG, Ng SC (2017) Global prevalence of *Helicobacter pylori* infection: systematic review and meta-analysis. *Gastroenterology* 153(2):420–429. <https://doi.org/10.1053/j.gastro.2017.04.022>
3. Liu WZ, Xie Y, Lu H, Cheng H, Zeng ZR, Zhou LY, Chen Y, Wang JB, Du YQ, Lu NH, Chinese Society of Gastroenterology CSGoHp, Peptic U (2018) Fifth Chinese National Consensus Report on the management of *Helicobacter pylori* infection. *Helicobacter* 23(2):e12475. <https://doi.org/10.1111/hel.12475>
4. Munoz-Ramirez ZY, Mendez-Tenorio A, Kato I, Bravo MM, Rizzato C, Thorell K, Torres R, Aviles-Jimenez F, Camorlinga M, Canzian F, Torres J (2017) Whole genome sequence and phylogenetic analysis show *Helicobacter pylori* strains from Latin America have followed a unique evolution pathway. *Front Cell Infect Microbiol* 7:50. <https://doi.org/10.3389/fcimb.2017.00050>
5. Yamaoka Y, Kato M, Asaka M (2008) Geographic differences in gastric cancer incidence can be explained by differences between *Helicobacter pylori* strains. *Intern Med* 47(12):1077–1083. <https://doi.org/10.2169/internalmedicine.47.0975>
6. Cover TL (2016) *Helicobacter pylori* diversity and gastric cancer risk. *MBio* 7(1):e01869-01815. <https://doi.org/10.1128/mBio.01869-15>
7. Ando T, Peek RM, Pride D, Levine SM, Takata T, Lee YC, Kusugami K, van der Ende A, Kuipers EJ, Kusters JG, Blaser MJ (2002) Polymorphisms of *Helicobacter pylori* HP0638 reflect geographic origin and correlate with cagA status. *J Clin Microbiol* 40(1):239–246. <https://doi.org/10.1128/jcm.40.1.239-246.2002>
8. Matsuo Y, Kido Y, Yamaoka Y (2017) *Helicobacter pylori* outer membrane protein-related pathogenesis. *Toxins*. <https://doi.org/10.3390/toxins9030101>
9. Oleastro M, Menard A (2013) The role of *Helicobacter pylori* outer membrane proteins in adherence and pathogenesis. *Biology* 2(3):1110–1134. <https://doi.org/10.3390/biology2031110>
10. Whitmire JM, Merrell DS (2019) *Helicobacter pylori* genetic polymorphisms in gastric disease development. *Adv Exp Med Biol* 1149:173–194. https://doi.org/10.1007/5584_2019_365
11. Cover TL (2006) Role of *Helicobacter pylori* outer membrane proteins in gastroduodenal disease. *J Infect Dis* 194(10):1343–1345. <https://doi.org/10.1086/508432>
12. Hage N, Howard T, Phillips C, Brassington C, Overman R, Debreczeni J, Gellert P, Stolnik S, Winkler GS, Falcone FH (2015) Structural basis of Lewis(b) antigen binding by the

- Helicobacter pylori* adhesin BabA. *Sci Adv* 1(7):e1500315. <https://doi.org/10.1126/sciadv.1500315>
13. Al-Maleki AR, Loke MF, Lui SY, Ramli NSK, Khosravi Y, Ng CG, Venkatraman G, Goh KL, Ho B, Vadivelu J (2017) *Helicobacter pylori* outer inflammatory protein A (OipA) suppresses apoptosis of AGS gastric cells in vitro. *Cell Microbiol*. <https://doi.org/10.1111/cmi.12771>
 14. Benktander J, Barone A, Johansson MM, Teneberg S (2018) *Helicobacter pylori* SabA binding gangliosides of human stomach. *Virulence* 9(1):738–751. <https://doi.org/10.1080/21505594.2018.1440171>
 15. Oleastro M, Cordeiro R, Ferrand J, Nunes B, Lehours P, Carvalho-Oliveira I, Mendes AI, Penque D, Monteiro L, Megraud F, Menard A (2008) Evaluation of the clinical significance of homB, a novel candidate marker of *Helicobacter pylori* strains associated with peptic ulcer disease. *J Infect Dis* 198(9):1379–1387. <https://doi.org/10.1086/592166>
 16. Odenbreit S, Swoboda K, Barwig I, Ruhl S, Boren T, Koletzko S, Haas R (2009) Outer membrane protein expression profile in *Helicobacter pylori* clinical isolates. *Infect Immun* 77(9):3782–3790. <https://doi.org/10.1128/IAI.00364-09>
 17. Yamaoka Y, Ojo O, Fujimoto S, Odenbreit S, Haas R, Gutierrez O, El-Zimaity HM, Reddy R, Arnqvist A, Graham DY (2006) *Helicobacter pylori* outer membrane proteins and gastroduodenal disease. *Gut* 55(6):775–781. <https://doi.org/10.1136/gut.2005.083014>
 18. Su YL, Huang HL, Huang BS, Chen PC, Chen CS, Wang HL, Lin PH, Chieh MS, Wu JJ, Yang JC, Chow LP (2016) Combination of OipA, BabA, and SabA as candidate biomarkers for predicting *Helicobacter pylori*-related gastric cancer. *Sci Rep* 6:36442. <https://doi.org/10.1038/srep36442>
 19. Yilmaz N, Koruk Ozer M (2019) The prevalence of *Helicobacter pylori* babA, homB, aspA, and sabA genes and its relationship with clinical outcomes in Turkey. *Can J Gastroenterol Hepatol* 2019:1271872. <https://doi.org/10.1155/2019/1271872>
 20. Liu DS, Wang YH, Zeng ZR, Zhang ZY, Lu H, Xu JM, Du YQ, Li Y, Wang JB, Xu SP, Chen Y, Lan CH, Cheng H, Jiang MD, Zhang LX, Huo LJ, Chen SY, Zhang GX, Wu KC, Zhu X, Chen YX, Zhu Y, Shu X, Xie Y, Lu NH (2018) Primary antibiotic resistance of *Helicobacter pylori* in Chinese patients: a multiregion prospective 7-year study. *Clin Microbiol Infect* 24(7):780.e785–780, e788. <https://doi.org/10.1016/j.cmi.2017.11.010>
 21. Zhang Z, Li J, Yu J (2006) Computing Ka and Ks with a consideration of unequal transitional substitutions. *BMC Evol Biol* 6:44. <https://doi.org/10.1186/1471-2148-6-44>
 22. Kudo T, Nurgalieva ZZ, Conner ME, Crawford S, Odenbreit S, Haas R, Graham DY, Yamaoka Y (2004) Correlation between *Helicobacter pylori* OipA protein expression and oipA gene switch status. *J Clin Microbiol* 42(5):2279–2281. <https://doi.org/10.1128/jcm.42.5.2279-2281.2004>
 23. Kawai M, Furuta Y, Yahara K, Tsuru T, Oshima K, Handa N, Takahashi N, Yoshida M, Azuma T, Hattori M, Uchiyama I, Kobayashi I (2011) Evolution in an oncogenic bacterial species with extreme genome plasticity: *Helicobacter pylori* East Asian genomes. *BMC Microbiol* 11:104. <https://doi.org/10.1186/1471-2180-11-104>
 24. Gutierrez-Escobar AJ, Trujillo E, Acevedo O, Bravo MM (2017) Phylogenomics of Colombian *Helicobacter pylori* isolates. *Gut Pathog* 9:52. <https://doi.org/10.1186/s13099-017-0201-1>
 25. Adrian J, Bonsignore P, Hammer S, Frickey T, Hauck CR (2019) Adaptation to host-specific bacterial pathogens drives rapid evolution of a human innate immune receptor. *Curr Biol* 29(4):616–630.e615. <https://doi.org/10.1016/j.cub.2019.01.058>
 26. Vaziri F, Najar Peerayeh S, Alebouyeh M, Mirzaei T, Yamaoka Y, Molaei M, Maghsoudi N, Zali MR (2013) Diversity of *Helicobacter pylori* genotypes in Iranian patients with different gastroduodenal disorders. *World J Gastroenterol* 19(34):5685–5692. <https://doi.org/10.3748/wjg.v19.i34.5685>
 27. Torres K, Valderrama E, Sayegh M, Ramirez JL, Chiurillo MA (2014) Study of the oipA genetic diversity and EPIYA motif patterns in cagA-positive *Helicobacter pylori* strains from Venezuelan patients with chronic gastritis. *Microb Pathog* 76:26–32. <https://doi.org/10.1016/j.micpath.2014.09.006>
 28. Kim SY, Woo CW, Lee YM, Son BR, Kim JW, Chae HB, Youn SJ, Park SM (2001) Genotyping CagA, VacA subtype, IceA1, and BabA of *Helicobacter pylori* isolates from Korean patients, and their association with gastroduodenal diseases. *J Korean Med Sci* 16(5):579–584. <https://doi.org/10.3346/jkms.2001.16.5.579>
 29. Shao L, Takeda H, Fukui T, Mabe K, Han J, Kawata S, Ootani K, Fukao A (2010) Genetic diversity of the *Helicobacter pylori* sialic acid-binding adhesin (sabA) gene. *Biosci Trends* 4(5):249–253
 30. Dossumbekova A, Prinz C, Mages J, Lang R, Kusters JG, Van Vliet AH, Reindl W, Backert S, Saur D, Schmid RM, Rad R (2006) *Helicobacter pylori* HopH (OipA) and bacterial pathogenicity: genetic and functional genomic analysis of hopH gene polymorphisms. *J Infect Dis* 194(10):1346–1355. <https://doi.org/10.1086/508426>
 31. Kim JY, Kim N, Nam RH, Suh JH, Chang H, Lee JW, Kim YS, Kim JM, Choi JW, Park JG, Lee YS, Lee DH, Jung HC (2014) Association of polymorphisms in virulence factor of *Helicobacter pylori* and gastroduodenal diseases in South Korea. *J Gastroenterol Hepatol* 29(5):984–991. <https://doi.org/10.1111/jgh.12509>
 32. Gerhard M, Lehn N, Neumayer N, Boren T, Rad R, Schepp W, Miehle S, Classen M, Prinz C (1999) Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. *Proc Natl Acad Sci USA* 96(22):12778–12783. <https://doi.org/10.1073/pnas.96.22.12778>
 33. Oleastro M, Cordeiro R, Yamaoka Y, Queiroz D, Megraud F, Monteiro L, Menard A (2009) Disease association with two *Helicobacter pylori* duplicate outer membrane protein genes, homB and homA. *Gut Pathog* 1(1):12. <https://doi.org/10.1186/1757-4749-1-12>
 34. Souod N, Sarshar M, Dabiri H, Momtaz H, Kargar M, Mohammadzadeh A, Abdi S (2015) The study of the oipA and dupA genes in *Helicobacter pylori* strains and their relationship with different gastroduodenal diseases. *Gastroenterol Hepatol Bed Bench* 8(Suppl 1):S47–53
 35. Chen MY, He CY, Meng X, Yuan Y (2013) Association of *Helicobacter pylori* baba2 with peptic ulcer disease and gastric cancer. *World J Gastroenterol* 19(26):4242–4251. <https://doi.org/10.3748/wjg.v19.i26.4242>
 36. Dabiri H, Jafari F, Baghaei K, Shokrzadeh L, Abdi S, Pourhoseingholi MA, Mohammadzadeh A (2017) Prevalence of *Helicobacter pylori* vacA, cagA, cagE, oipA, iceA, baba2 and babB genotypes in Iranian dyspeptic patients. *Microb Pathog* 105:226–230. <https://doi.org/10.1016/j.micpath.2017.02.018>
 37. Abadi ATB (2017) Strategies used by *Helicobacter pylori* to establish persistent infection. *World J Gastroenterol* 23(16):2870–2882. <https://doi.org/10.3748/wjg.v23.i16.2870>
 38. Gangwer KA, Shaffer CL, Suerbaum S, Lacy DB, Cover TL, Bordenstein SR (2010) Molecular evolution of the *Helicobacter pylori* vacuolating toxin gene vacA. *J Bacteriol* 192(23):6126–6135. <https://doi.org/10.1128/JB.01081-10>
 39. Duncan SS, Valk PL, McClain MS, Shaffer CL, Metcalf JA, Bordenstein SR, Cover TL (2013) Comparative genomic analysis of East Asian and non-Asian *Helicobacter pylori* strains identifies rapidly evolving genes. *PLoS ONE* 8(1):e55120. <https://doi.org/10.1371/journal.pone.0055120>
 40. Soto-Giron MJ, Ospina OE, Massey SE (2015) Elevated levels of adaptation in *Helicobacter pylori* genomes from Japan; a link to higher incidences of gastric cancer? *Evol Med Public Health* 1:88–105. <https://doi.org/10.1093/emph/eov005>

41. Yamaoka Y, Kwon DH, Graham DY (2000) A M(r) 34,000 pro-inflammatory outer membrane protein (oipA) of *Helicobacter pylori*. Proc Natl Acad Sci USA 97(13):7533–7538. <https://doi.org/10.1073/pnas.130079797>
42. Oh JD, Kling-Backhed H, Giannakis M, Xu J, Fulton RS, Fulton LA, Cordum HS, Wang C, Elliott G, Edwards J, Mardis ER, Engstrand LG, Gordon JI (2006) The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. Proc Natl Acad Sci USA 103(26):9999–10004. <https://doi.org/10.1073/pnas.0603784103>
43. Sheu BS, Odenbreit S, Hung KH, Liu CP, Sheu SM, Yang HB, Wu JJ (2006) Interaction between host gastric Sialyl-Lewis X and *H. pylori* SabA enhances *H. pylori* density in patients lacking gastric Lewis B antigen. Am J Gastroenterol 101(1):36–44. <https://doi.org/10.1111/j.1572-0241.2006.00358.x>

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

Affiliations

Qiaoyun Zhao^{1,2,3}  · Conghua Song^{1,2,3,5} · Ke Wang⁴ · Donghong Li⁴ · Yang Yang^{1,2,3} · Dongsheng Liu^{2,3} · Le Wang^{2,3} · Nanjin Zhou⁴ · Yong Xie^{1,2,3} 

¹ Department of Gastroenterology, First Affiliated Hospital of Nanchang University, No.17 Yongwaizheng Street, Donghu District, Nanchang 330006, Jiangxi, China

² Microbiology Laboratory, Gastroenterology Institute of Jiangxi Province, Nanchang 330006, Jiangxi, China

³ Gastroenterology Research Center, Key Laboratory of Digestive Diseases of Jiangxi Province, Nanchang 330006, Jiangxi, China

⁴ Laboratory of Biochemistry and Molecular Biology, Jiangxi Institute of Medical Sciences, Nanchang 330006, Jiangxi, China

⁵ Department of Gastroenterology, Affiliated Hospital of Putian University, Putian 351100, Fujian, China