

RESEARCH ARTICLE

Distribution characteristics of the *sabA*, *hofC*, *homA*, *homB* and *frpB-4* genes of *Helicobacter pylori* in different regions of ChinaMengyang Fang^{1,2}, Zhijing Xue², Lihua He², Yuanhai You², Yanan Gong², Dongjie Fan^{1,2}, Lu Sun², Kangle Zhai^{1,2}, Yaming Yang^{1,2}, Jianzhong Zhang^{1,2}*

1 Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, China, 2 State Key Laboratory of Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Chinese Center for Disease Control and Prevention, National Institute for Communicable Disease Control and Prevention, Beijing, China

 These authors contributed equally to this work.

* zhangjianzhong@icdc.cn


 OPEN ACCESS

Citation: Fang M, Xue Z, He L, You Y, Gong Y, Fan D, et al. (2022) Distribution characteristics of the *sabA*, *hofC*, *homA*, *homB* and *frpB-4* genes of *Helicobacter pylori* in different regions of China. PLoS ONE 17(5): e0268373. <https://doi.org/10.1371/journal.pone.0268373>

Editor: Egon Anderson Ozer, Northwestern University Feinberg School of Medicine, UNITED STATES

Received: January 3, 2022

Accepted: April 28, 2022

Published: May 19, 2022

Copyright: © 2022 Fang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting Information](#) files.

Funding: National Science and Technology Major Project of China (2018ZX10712-001) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Background

Helicobacter pylori (*H. pylori*) encodes numerous outer membrane proteins (OMPs), with considerable geographic heterogeneity and related to different clinical outcomes. This study aimed to investigate the distribution characteristics of five important OMP genes (*sabA*, *hofC*, *homA*, *homB* and *frpB-4*) in different regions of China.

Materials and method

A total of 266 strains were isolated from 348 stomach biopsy specimens in Shandong, Guangxi, Heilongjiang, Hunan, and Qinghai provinces. The presence of *sabA*, *hofC*, *homA*, *homB* and *frpB-4* gene was detected by polymerase chain reaction (PCR) from *H. pylori* genomic DNA.

Results

Among the strains in five regions, the prevalence of *frpB-4* was 100% and that of *hofC* was 97.7%. The prevalence of *homB* in the isolates from Qinghai (45.5%) was significantly lower than that in Shandong (75.3%), Guangxi (76.9%) and Hunan (69.6%) ($P < 0.05$). The frequency of *homA* in Shandong (30.1%) was significantly lower than in Guangxi (57.7%) and Qinghai (63.6%) ($P < 0.05$). The prevalence of the *sabA* gene in Shandong, Guangxi, Heilongjiang, Hunan and Qinghai provinces was 21.9%, 59.7%, 45.9%, 52.2%, and 18.2%, respectively ($P < 0.05$). The *sabA* “on” status was significantly more frequent in isolates from Guangxi (46.8%), Heilongjiang (37.8%), and Hunan (47.8%) than Qinghai (3.0%) ($P < 0.05$). The presence of *homA* and *sabA* genes may be negatively correlated with the development of gastritis. There was no significant association between the *frpB-4*, *hofC*, *homB* gene and clinical outcomes.

Conclusion

The prevalence of *homA*, *homB*, and *sabA* genes and the *sabA* “on” or “off” status have significant geographical differences among five provinces in China. The presence of *homA* and *sabA* genes may be protective factors of gastritis.

Introduction

Helicobacter pylori (*H. pylori*) is a Gram-negative, helix-shaped, microaerophilic, flagellated bacterium, which colonizes the mucus layer of the gastric epithelium [1]. *H. pylori* infection leads to peptic ulcer disease (PUD) in 15%-20% of the infected population, dyspepsia in 5%-10% of the infected population, and gastric cancer (GC) or mucosa-associated lymphoid tissue (MALT) lymphoma in 1% of the infected population [2]. While the exact molecular mechanisms by which *H. pylori* infection induces a severe clinical outcome have not yet been clearly elucidated, they are thought to involve various elements, including host genetic and environmental factors, as well as certain bacterial virulence genes [3].

The outer membrane is the outer barrier of Gram-negative bacteria, which consists of two highly asymmetric layers: the inner monolayer contains only phospholipids and the outer monolayer consists mainly of outer membrane proteins (OMPs) that are resistant to the external environment [4, 5]. About 4% of *H. pylori* genome encodes OMP [6]. According to their functions, they can be divided into five homologous gene families [6]: Hop (outer membrane porins) and Hor (Hop-related proteins) proteins, Hof (*H. pylori* OMP) proteins, Hom (*H. pylori* outer membrane) proteins, iron-regulated OMPs, and efflux pump OMPs families.

HopP belongs to the Hop family, also known as sialic acid-binding adhesion (SabA). SabA specifically binds to sLex antigen on gastric mucosa epithelial cell and plays an important role in *H. pylori* colonization and inflammation mediation [7, 8]. SabA expression level can quickly adapt to changes in the human gastric environment by “on” or “off”. The “on” status of *sabA* was negatively correlated with the degree of gastric acid secretion, suggesting that the pH or antigen expression of atrophic mucosa may affect the expression of SabA.

HofC is a member of a paralogous Hof family which consists of eight members. HofC encodes a non-thermal denatured protein with 528 amino acid residues, which has a hydrophobic C-terminal sequence motif of many outer membrane proteins [6]. HofC proteins are involved in passive diffusion and adhesion of cations such as antibiotics [9]. The *hofC* gene is highly variable in global strains and shows many America-differentiated SNPs and region-differentiated SNPs within the Americas [10].

Hom family is a small OMP family composed of HomA, HomB, HomC, and HomD. The *homA* gene, which presents more than 90% identity to *homB* [6]. Interestingly, *homA* was more frequently found in strains isolated from non-ulcer dyspepsia (NUD). In vitro, HomB promotes the secretion of the interleukin-8 (IL-8) and increases *H. pylori* adhesion. The *homA* and *homB* sequences have considerable geographic heterogeneity [11]. The prevalence of the *homA* and *homB* genes is different in strains all over the world, and there are significant differences between East Asian and Western strains [12].

H. pylori contains six iron-regulated OMPs. These OMPs can be divided into two groups based on homology. One of them is homologous of the *Neisseria gonorrhoeae* ferric enterobactin receptor FrpB, encoded by *frpB-1*, *frpB-2/3*, and *frpB-4* [6, 13]. A gene containing multiple highly differentiated SNPs is *frpB-4*, encoding for the first TonB-dependent nickel transport system across a bacterial outer membrane [14]. Studies have shown that the protein

encoded by *frpB-4* will show specific amino acid changes in different regions, and these changes may affect the transport of nickel, thereby affecting the activity of urease [10].

China is a country with a vast territory and a large population. The five provinces selected in this study cover the eastern (Shandong), northern (Heilongjiang), western (Qinghai), southern (Guangxi), and central (Hunan) regions of China. The current infection rate of *H. pylori* in China is about 40–50%. *H. pylori* is one of the causes of gastric cancer (GC), which is the second most common cancer in China. Shandong (30.50/100000) and Qinghai (48.76/100000) provinces are considered to have a high incidence of GC, while Heilongjiang (15.53/100000), Guangxi (19.68/100000) and Hunan (10–20/100000) provinces have a relatively low incidence of GC [15, 16]. Meanwhile, the economic development level, living environment and dietary habits of these selected regions are quite different, which is more helpful for us to analyze the diversity of *H. pylori* distribution in China. At present, studies on OMP pay more attention to its correlation with diseases and other mechanisms. Only a few studies have analyzed the differential characteristics of *H. pylori* OMP coding genes in different regions. Therefore, we studied the distribution characteristics of *sabA*, *hofC*, *homA*, *homB*, and *frpB-4* in different regions of China and their association with clinical outcomes.

Materials and methods

Study subjects

A total of 266 patients were included in this study, including 73 cases in Weihai, Shandong province, 77 cases in Nanning, Guangxi province, 46 cases in Yiyang, Hunan province, 37 cases in Jiamusi, Heilongjiang province, and 33 cases in Haidong, Qinghai province, as previously reported [17]. Endoscopic examination showed PUD in 15 patients, chronic superficial gastritis (CSG) in 81 patients, chronic erosive gastritis (CEG) in 87 patients, chronic atrophic gastritis (CAG) in 42 patients. Their gastric biopsy specimens were obtained during upper gastrointestinal endoscopy with informed consent. This study was approved by Ethical Committee of National Institute for Communicable Disease Control and Prevention Chinese Center for Disease Control and Prevention (approval No. ICDC-2013001).

H. pylori culture and identification

The strains were obtained from patients with varying gastric diseases as diagnosed by routine endoscopy and biopsy sampling. *H. pylori* strains were cultured on Karmali agar base plates supplemented with 5% defibring sheep blood and 1% combined antibiotics. The plates were grown at 37°C under microaerobic conditions (5% O₂, 10% CO₂ and 85% N₂) approximately for 3–5 days. The organisms were identified as *H. pylori* based on colony morphology and Gram staining, as well as positive reactions in oxidase, catalase and urease biochemical tests, and also subsequently the results of *H. pylori*-specific polymerase chain reaction (PCR). All isolates were maintained at -80°C in sterile brain heart infusion (BHI) broth with 20% glycerol.

DNA isolation and PCR amplification

Genomic DNA was extracted by QIAamp DNA Mini Kit according to the manual instruction. The DNA was stored at -20°C for molecular studies. All PCR amplification reactions were performed by standard methods in a final volume of 25 µL containing forward and reverse primers (0.2 µM each), 2 ng/µL DNA template, 12.5 µL 2×EasyTaq PCR SuperMix (Transgen, China) and 9.5 µL nuclease-free water. The PCR protocol (35 cycles) included a denaturing step at 94°C for 30 sec, annealing at 55, 56, 55, 55 and 60°C for *hofC*, *frpB-4*, *homB*, *homA* and *sabA*, respectively, for 30 sec, and extension at 72°C for 1 minutes followed by a 10-minute

Table 1. Primers used for PCR amplification of *hofC*, *frpB-4*, *homB*, *homA*, and *sabA* genes.

Gene	Primer	primer sequence (5'→3')	Product size (bp)
<i>hofC</i>	<i>hofC</i> -F	GCTTGCCACTRTTGTTCACT	907
	<i>hofC</i> -R	CGACCGTATTCAGCGTTATT	
<i>frpB-4</i> [18]	<i>frpB-4</i> -F	AGCCGTCTCTTAAGGGTAAC	407
	<i>frpB-4</i> -R	TCGCTATTGCTTGGATCTTG	
<i>homB</i> [19]	<i>homB</i> -F	AGAGGGTGTTTGAAACGCTCAATA	161
	<i>homB</i> -R	GGTGAATTCTTCTGCGGTTTG	
<i>homA</i> [19]	<i>homA</i> -F	AGAGGGTGTTTGAAACGCTCAATA	128
	<i>homA</i> -R	GGTGAATTCTTCTGCGGTTTG	
<i>sabA</i> [20]	<i>sabA</i> -F	GTGGATCCCTTTAAGGAACATTTTATGAAAA	600
	<i>sabA</i> -R	GGAATTCGGTTAGGATAAAAAGCGCAAAGATTGT	

<https://doi.org/10.1371/journal.pone.0268373.t001>

extension step at 72°C. The *sabA* gene was detected by PCR, which was additionally sequenced to define its functional status as either “on” or “off”. The nucleotide sequences of the *sabA* gene were submitted to China National Microbiological Data Center. The accession numbers are NMDCN0000R8M to NMDCN0000RBU. The functional status of *sabA* is regulated by the number of CT-dinucleotide repeats. The information of the PCR primer for each amplicon is summarized in Table 1. PCR products were then electrophoresed for 35 min at 130 Volt on 1.5% agarose gel in the presence of Gelstain and illuminated by the gel documentation system (Bio-Rad, USA).

DNA sequencing

Samples without target genes of *sabA* were discarded at the identification stage of the product. The positive PCR products of isolates were purified and then the DNA samples were submitted and performed as a service by the Beijing Genomics Institute for the sequence of *sabA* genes. The *sabA* gene resulting PCR products were used for DNA sequencing to determine the number of CT-dinucleotide repeats.

Statistical analysis

Statistical analysis was performed by chi-squared test for independence. All data analysis were performed using the SPSS software version 25. All tests of significance were two-tailed with a P -value < 0.05 taken as significant.

Result

A total of 266 *H. pylori* isolates from five geographic regions of China were obtained, of which 73 isolates were from Shandong, 77 from Guangxi, 46 from Hunan, 33 from Qinghai and 37 from Heilongjiang. The *hofC*, *frpB-4*, *homB*, *homA* and *sabA* genes were detected by PCR in all isolates and the results were summarized in Table 2.

hofC status and *frpB-4* status

The 907-bp PCR product indicating the presence of *hofC* gene of *H. pylori* was determined in 260 patients (97.7%), whereas 6 (2.2%) patients were classified as *hofC*-negative. Of these *hofC*-positive strains, 70 (95.9%) strains were isolated from Shandong, 45 (97.8%) from Hunan, 31 (93.9%) from Qinghai, 100.0% from Guangxi and Heilongjiang. There were no significant differences between the *hofC* gene and clinical outcomes ($\chi^2 = 5.937$, $P > 0.05$). The *frpB-4* gene was found in all 266 (100.0%) isolates.

Table 2. Distribution of OMP genes among 267 *H. pylori*-positive patients with different geographic regions.

Genotypes	Shandong	Guangxi	Heilongjiang	Hunan	Qinghai	Total	χ^2	P-value
	(n = 73)	(n = 77)	(n = 37)	(n = 46)	(n = 33)	(n = 266)		
<i>hofC</i> ⁺	70 (95.9)*	77 (100)	37 (100)	45 (97.8)	31 (93.9)	260 (97.7)	5.937	0.204
<i>hofC</i> ⁻	3 (4.1)	0	0	1 (2.2)	2 (6.1)	6 (2.2)		
<i>frpB-A</i> ⁺	73 (100)	77 (100)	37 (100)	46 (100)	33 (100)	266 (100)		
<i>homB</i> ⁺	55 (75.3)	59 (76.6)	23 (62.2)	32 (69.6)	15 (45.5)	184 (69.2)	12.870	0.012
<i>homB</i> ⁻	18 (24.7)	18 (23.4)	14 (37.8)	14 (30.4)	18 (54.5)	82 (30.8)		
<i>homA</i> ⁺	22 (30.1)	44 (57.1)	17 (45.9)	19 (41.3)	21 (63.6)	123 (46.2)	15.766	0.003
<i>homA</i> ⁻	51 (69.9)	33 (42.9)	20 (54.1)	27 (58.7)	12 (36.4)	143 (53.8)		
<i>sabA</i> ⁺	16 (21.9)	46 (59.7)	17 (45.9)	24 (52.2)	6 (18.2)	109 (41.0)	32.024	<0.001
<i>sabA</i> ⁻	57 (78.1)	32 (40.3)	20 (54.1)	22 (47.8)	27 (81.8)	158 (59.0)		
<i>sabA</i> ^{“on”}	11 (15.1)	36 (46.8)	14 (37.8)	22 (47.8)	1 (3.0)	84 (31.6)	16.210	0.003
<i>sabA</i> ^{“off”}	5 (6.8)	10 (13.0)	3 (8.1)	2 (4.3)	5 (15.2)	25 (9.4)		

*Values in parentheses are percentages

<https://doi.org/10.1371/journal.pone.0268373.t002>

homB and *homA* status

The *homB* gene was determined in 184 isolates (69.2%), whereas 82 isolates (30.8%) were classified as *homB*-negative. The prevalence of *homB* in the isolates from Qinghai (45.5%) ($\chi^2 = 12.870, P < 0.05$) was significantly lower than that in Shandong (75.3%), Guangxi (76.6%), and Hunan (69.6%). For another hom family gene, 123 (46.2%) patients were infected with *homA* and 143 (53.8%) patients were *homA*-negative. The frequency of *homA* in Shandong (30.1%) ($\chi^2 = 15.766, P < 0.05$) was significantly lower than in Guangxi (57.1%) and Qinghai (63.6%). There was no association between the *homB* genes and clinical outcomes in all five regions. However, in the strain from all provinces, *homA* showed significant differences between CSG and CAG, suggesting that *homA* may be associated with the severity of gastritis, as shown in [S1 Table](#) ($\chi^2 = 6.707, P < 0.05$). The presence of *homA* may be a protective factor for gastritis (OR = 0.364).

sabA status

Sixteen *sabA* genes from Shandong province, 46 from Guangxi province, 17 from Heilongjiang province, 24 from Hunan province, 6 from Qinghai province contained a run of CT repeats. The *sabA* frequency was significantly more prevalent in Guangxi (59.7%), Hunan (52.2%), and Heilongjiang (45.9%) than in Shandong (21.9%) and Qinghai (18.2%) ($\chi^2 = 32.024, P < 0.001$). Moreover, the presence of the *sabA* gene in *H. pylori* isolates was correlated with the severity of gastritis. The severity of CSG, CEG, and CAG increased successively. The *sabA*-positive gene was significantly related to the differences of CSG and CAG ($\chi^2 = 10.539, P < 0.05$). The *sabA* gene was negatively correlated with the development of gastritis (OR = 0.253).

All *sabA* genes tested displayed different CT-repeat lengths ranging from 2 to 10 repeats between different *H. pylori* strains. In all cases, the CT repeats were associated with the 5' region of *sabA* genes. In addition, the repeated pattern of CT dinucleotide in the coding region was found and classified. In 109 *sabA*-positive gene sequencing results, 84 (77.1%) of which had a switch on status while 25 (22.9%) had an off status. The *sabA* functional status was significantly more frequent in isolates from Guangxi (46.8%), Heilongjiang (37.8%), and Hunan (47.8%) than Qinghai (3.0%) ($\chi^2 = 16.210, P < 0.05$). Furthermore, the repeated pattern of CT dinucleotide varied in the sequences which were different from *sabA* ([Table 3](#)). The pattern containing 2 CT repeats was the most frequently associated with the “on” status (58/109,

Table 3. Frequency of the *sabA* CT repeats patterns.

Strain	CT repeat	Sequence of the signal peptide coding region	Amino acid sequence	<i>sabA</i> status	Number(%)
HP-1	2	ATG-CTCTAT-CTGAAGATA-TTTTTG TGA	MKK-LSLSL-FL*	Off	1 (0.9)
HP-2	2	ATG-CTCTAT-GCTGAAGAC-TTTTTGTGAGC	MKK-LSL-AED-FFVS	On	58 (53.2)
HP-3	4	ATG-CTCTCTCTCG-GCTGAAGAT-TTTTTGTGAGC	MKK-LSL-AED-FFVS	On	9 (8.3)
HP-4	5	ATG-CTCTCTCTCTCG-TTT TAA	MKK-LSLSL-F*	Off	3 (2.8)
HP-5	5	ATG-CTCTCTCTCTCG-GCTGAAGAC-TTTTTGTGAGC	MKK-LSLSL-AED-FFVS	On	4 (3.7)
HP-6	6	ATG-CTCTCTCTCTCTCGC-GCACG CTGA	MKK-LSLS-AR*	Off	9 (8.3)
HP-7	7	ATG-CTCTCTCTCTCTCTCG-CG CTGA	MKK-LSLSL-R*	Off	3 (2.8)
HP-8	7	ATG-CTCTCTCTCTCTCTCG-GCTGAAGAC-TTTTTGTGAGC	MKK-LSLSL-AED-FFVS	On	6 (5.5)
HP-9	8	ATG-CTCTCTCTCTCTCTCTCG-ACGCTGAAG-TTTTT TGA	MKK-LSLSLSL-FL*	Off	3 (2.8)
HP-10	8	ATG-CTCTCTCTCTCTCTCTCG-ACGCTGAAG-TTTTTGTGA	MKK-LSLSLSL-FFVS	On	6 (5.5)
HP-11	9	ATG-CTCTCTCTCTCTCTCTCTCG-TTTTT TATGA	MKK-LSLSLS-FL*	Off	5 (4.6)
HP-12	10	ATG-CTCTCTCTCTCTCTCTCTCTCG-GCTGAAGAC-TTTTTTATAAGC	MKK-LSLSLSL-AED-FFIS	On	1 (0.9)
HP-13	10	ATG-CTCTCTCTCTCTCTCTCTCTCGCT-CG CTGA	MKK-LSLSLSL-R*	Off	1 (0.9)

*Indicates stop codon

<https://doi.org/10.1371/journal.pone.0268373.t003>

53.2%), and the pattern with 6 CT repeats was the most prevalent for a nonfunctional (“off”) *sabA* gene (9/25, 36%). However, the differences between the *sabA* status and clinical outcome were not statistically significant.

Combined presence of OMP encoding genes *frpB-4*, *hofC*, *homB*, *homA*, and *sabA* in different regions

The presence of *homA* was associated with the presence of *homB* ($\chi^2 = 124.332$, $P < 0.001$) and *sabA* ($\chi^2 = 5.482$, $P < 0.05$). A statistically significant correlation between *homA* and *homB* and *sabA* gene was detected. The relationship between *homA* and *homB* is even closer ($r = 0.564$) (Table 4).

As *frpB-4* and *hofC* genes were almost all positive in all 266 strains, and the correlation between the two genes and geographical distribution was not statistically significant, we selected strains that were both positive for *frpB-4* and *hofC* when analyzing the association between gene combination and geographical origin, and also excluded confounding factors.

Table 4. Comparison and statistical analysis of *H. pylori* OMP genes.

Comparison	χ^2	P-value	r
<i>hofC</i> vs <i>frpB-4</i>	/	/	
<i>hofC</i> vs <i>homB</i>	0.338	0.561	0.063
<i>hofC</i> vs <i>homA</i>	0.361	0.548	0.062
<i>hofC</i> vs <i>sabA</i>	0.0E0	1.000	0.024
<i>frpB-4</i> vs <i>homB</i>	/	/	/
<i>frpB-4</i> vs <i>homA</i>	/	/	/
<i>frpB-4</i> vs <i>sabA</i>	/	/	/
<i>homB</i> vs <i>homA</i>	125.597*	3.7671E⁻²⁹	0.566
<i>homB</i> vs <i>sabA</i>	2.984	0.084	0.105
<i>homA</i> vs <i>sabA</i>	5.760	0.016	0.146

*Numbers in boldface type indicate a significant correlation between the two genes

<https://doi.org/10.1371/journal.pone.0268373.t004>

Table 5. Prevalence of combined *H. pylori* *hofC*, *frpB-4*, *homA*, *homB*, *sabA* in different regions.

Genotype	Shandong	Guangxi	Heilongjiang	Hunan	Qinghai	Total	χ^2	P-value
	(n = 73)	(n = 77)	(n = 37)	(n = 46)	(n = 33)	(n = 266)		
<i>hofC</i> ⁺ & <i>frpB-4</i> ⁺ & <i>homB</i> ⁺ & <i>homA</i> ⁺ *	4 (5.5)	27 (33.8)	3 (6.8)	5 (10.9)	4 (12.1)	42 (15.8)	27.364	<0.001
<i>hofC</i> <i>frpB-4</i> <i>homB</i> <i>homA</i> ⁺ *	69 (94.5)	51 (66.2)	34 (91.9)	41 (89.1)	29 (87.9)	224 (84.2)		
<i>hofC</i> ⁺ & <i>frpB-4</i> ⁺ & <i>homB</i> ⁺ & <i>sabA</i> ⁺	9 (12.3)	34 (44.2)	7 (18.9)	17 (37.0)	1 (3.0)	68 (25.6)	33.509	<0.001
<i>hofC</i> <i>frpB-4</i> <i>homB</i> <i>sabA</i> ⁺	64 (87.7)	43 (55.8)	30 (81.1)	29 (63.0)	32 (97.0)	198 (74.4)		
<i>hofC</i> ⁺ & <i>frpB-4</i> ⁺ & <i>homA</i> ⁺ & <i>sabA</i> ⁺	6 (8.2)	30 (39.0)	10 (27.0)	8 (17.4)	4 (12.1)	58 (21.8)	24.127	<0.001
<i>hofC</i> <i>frpB-4</i> <i>homA</i> ⁺ <i>sabA</i> ⁺	67 (91.8)	47 (61.0)	27 (73.0)	38 (82.6)	29 (87.9)	208 (78.2)		
<i>hofC</i> ⁺ & <i>frpB-4</i> ⁺ & <i>homB</i> ⁺ & <i>homA</i> ⁺ & <i>sabA</i> ⁺	1 (1.4)	18 (23.4)	0	1 (2.2)	0	20 (7.5)	39.397	<0.001
<i>hofC</i> <i>frpB-4</i> <i>homB</i> <i>homA</i> ⁺ <i>sabA</i> ⁺	72 (98.6)	59 (76.6)	37 (100)	45 (97.8)	33 (100)	246 (92.5)		

*'&' indicate 'and', '|' indicate 'or', values in parentheses are percentages

<https://doi.org/10.1371/journal.pone.0268373.t005>

When considered in five genes associated combinations (Table 5), the frequency of *hofC*/*frpB-4*/*homB*/*homA* genotypes and *hofC*/*frpB-4*/*homB*/*homA*/*sabA* genotypes was significantly in Guangxi (33.8%) ($\chi^2 = 27.364$, $P < 0.001$ and $\chi^2 = 39.397$, $P < 0.001$) higher compared to the other four provinces. The frequency of *hofC*/*frpB-4*/*homB*/*sabA* genotypes in Shandong, Guangxi, Heilongjiang, Hunan and Qinghai provinces were 12.3%, 44.2%, 18.9%, 37.0%, 3.0%, respectively ($\chi^2 = 33.509$, $P < 0.001$). Furthermore, the *hofC*/*frpB-4*/*homA*/*sabA* genotypes frequency was significantly more prevalent in Guangxi (39.0%) ($\chi^2 = 24.127$, $P < 0.001$) than in Shandong (8.2%), Hunan (17.4%) and Qinghai (12.1%). The *hofC*/*frpB-4*/*homA*/*sabA* genotypes were significantly more frequent in Heilongjiang (27.0%) ($\chi^2 = 24.127$, $P < 0.001$) than in Shandong (8.2%).

The above four gene combinations were negatively correlated with gastritis, which may be protective factors of gastritis. The specific OR values are shown in S1 Table.

Discussion

H. pylori is one of the most successful pathogens that colonizes the stomach of half the world's population for a long time [21]. *H. pylori* infection can cause serious clinical consequences such as chronic gastritis, PUD, gastric atrophy, and GC [22, 23]. The prevalence of *H. pylori* depends on geographic regions, age, occupation, social and economic status, and living environment [24]. Outer membrane proteins are very important during infection and can influence the levels of bacterial colonization [25].

In this study, the distribution of *frpB-4*, *hofC*, *homB*, *homA*, and *sabA* genes in *H. pylori* isolated from patients suffering from gastric diseases in China was determined by PCR, and the relationship between these OMP encoding genes was assessed.

The results showed that there were significant statistical differences in the distribution of *homA*, *homB*, and *sabA* genes in different provinces of China. The distribution of different gene combinations in the five provinces was also showed significant statistical differences. This may be related to the selective adaptation of genes in different regions. China, as a country with vast territory, its geographical distribution of population has been in a relatively stable state for a long time, which may lead to *H. pylori* infection in different regions of China with more significant OMP distribution characteristics.

H. pylori carries two paralogous OMP, HomA and HomB, which have recently been suggested to be important determinants of disease severity [12]. In our study, the *homB* gene was significantly more prevalent than the *homA* gene ($\chi^2 = 125.597$, $P < 0.001$). The positive rate of *homB* was the highest in Shandong province and the lowest in Qinghai Province, while *homA*

was the opposite. The presence of the *homA* gene was significantly associated with the absence of *homB* gene since only 43 (16.2%) simultaneously harbored both genes. In western countries, the prevalence of *homA* was 61.9% and that of *homB* was 61.2%, which were approximately the same. There was no significant difference in *homB* prevalence between Colombian and American strains [12]. Compared with our result, the prevalence of *homA* (46.2%) was significantly different. Another study showed that *homA* and *homB* genes were heterogeneously distributed worldwide, with significant differences between East Asian and Western strains. Moreover, *homB* was found more frequently in East Asian strains than *homA*, but were not associated with clinical outcome. The presence of both genes in the same genome was detected in 10.4% of strains [26]. This is consistent with the results of our study. In addition, it was also pointed out that the copy number polymorphism of *homB* and *homA* had obvious geographical specificity [26]. Further sequence analysis using *H. pylori* strains from different geographic backgrounds in this study can assess whether different alleles could be associated with the severity of clinical outcomes or different geographic origin [26, 27].

In this study, the prevalence of *sabA*-positive gene was 41.0% (109/266) and that of *sabA* “on” was 31.6% (84/266). The prevalence of Portugal, the Netherlands, and Italy were 63.2%, 49.0%, and 35.5%, respectively [28–30]. The *sabA*-positive strains accounted for 85.3% of the Iranian studies, but the prevalence of *sabA* “on” status was not reported [31]. In contrast, the prevalence of *sabA* functional status was higher in Japan (81.5%) [20]. Analysis from Taiwan showed that 80.0% of strains had the *sabA* gene, while only 31.0% (45/145) strains carried *sabA* [8]. These large differences may be due to different disease sources of the strains or genetic diversity in different parts of the world. In addition, the association between *sabA* functional status and disease has not been fully established, and there is significant geographical diversity.

In this study, 100% of *H. pylori* strains carried the *frpB-4* gene and 97.7% carried the *hofC* gene, which has nothing to do with geographic origin and clinical outcome. Specific positive rates of *frpB-4* and *hofC* in different regions have not been reported. However, it was found that *frpB-4* was a gene containing multiple highly differentiated SNPs in strains from different regions of the world, especially in northeast China. The *hofC* gene is highly variable in global strains and shows many America-differentiated SNPs and region-differentiated SNPs within the Americas [9]. In Fujian, China, an area with a high incidence of cancer, *HofC* also contains the most differentiated SNPs [10, 32]. Therefore, we should analyze the relationship between the two and the disease and geographic distribution from a more microscopic perspective.

There are still some limitations to our research. At present, we only studied the distribution characteristics of *H. pylori* outer membrane protein encoding genes in different regions of China, mainly comparing the positive rate of genes in different regions, gene interaction, genotype and gene combination with regional distribution and clinical outcomes. We have not thoroughly studied the function and mechanism of one or several genes in *H. pylori* colonization and pathogenicity. We will supplement it in the following study.

Conclusion

The *homA*, *homB* and *sabA* genes of *H. pylori* have significant geographical differences among five provinces in China. The *homA* and *sabA* genes were negatively correlated with the severity of gastritis.

Supporting information

S1 Table. Frequency of 267 *H. pylori* OMP genes in patient with CSG, CEG and CAG. (DOCX)

Author Contributions

Conceptualization: Mengyang Fang, Zhijing Xue, Jianzhong Zhang.

Formal analysis: Mengyang Fang, Kangle Zhai, Yaming Yang.

Resources: Zhijing Xue, Lihua He, Yuanhai You, Yanan Gong, Dongjie Fan, Lu Sun.

Writing – original draft: Mengyang Fang.

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 (Lond)* 51: 399–408. <https://doi.org/10.1080/23744235.2019.1588472> PMID: 30907202
2. Liu WZ, Xie Y, Lu H, Cheng H, Zeng ZR, Zhou LY, et al. (2018) Fifth Chinese National Consensus Report on the management of *Helicobacter pylori* infection. *HELICOBACTER* 23: e12475. <https://doi.org/10.1111/hel.12475> PMID: 29512258
3. Šterbenc A, Jarc E, Poljak M, Homan M. (2019) *Helicobacter pylori* virulence genes. *World J Gastroenterol* 25: 4870–84. <https://doi.org/10.3748/wjg.v25.i33.4870> PMID: 31543679
4. Egan AJF. (2018) Bacterial outer membrane constriction. *Mol Microbiol* 107: 676–87. <https://doi.org/10.1111/mmi.13908> PMID: 29315884
5. Qiao S, Luo Q, Zhao Y, Zhang XC, Huang Y. (2014) Structural basis for lipopolysaccharide insertion in the bacterial outer membrane. *NATURE* 511: 108–11. <https://doi.org/10.1038/nature13484> PMID: 24990751
6. Alm RA, Bina J, Andrews BM, Doig P, Hancock RE, Trust TJ. (2000) Comparative genomics of *Helicobacter pylori*: analysis of the outer membrane protein families. *Infect Immun* 68: 4155–68. <https://doi.org/10.1128/IAI.68.7.4155-4168.2000> PMID: 10858232
7. Mahdavi J, Sondén B, Hurtig M, Olfat FO, Forsberg L, Roche N, et al. (2002) *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. *SCIENCE* 297: 573–8. <https://doi.org/10.1126/science.1069076> PMID: 12142529
8. Sheu BS, Odenbreit S, Hung KH, Liu CP, Sheu SM, Yang HB, et al. (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: 36–44. <https://doi.org/10.1111/j.1572-0241.2006.00358.x> PMID: 16405531
9. Bauwens E, Joosten M, Taganna J, Rossi M, Debraekeleer A, Tay A, et al. (2018) In silico proteomic and phylogenetic analysis of the outer membrane protein repertoire of gastric *Helicobacter* species. *Sci Rep* 8: 15453. <https://doi.org/10.1038/s41598-018-32476-1> PMID: 30337679
10. Thorell K, Yahara K, Berthenet E, Lawson DJ, Mikhail J, Kato I, et al. (2017) Rapid evolution of distinct *Helicobacter pylori* subpopulations in the Americas. *PLoS Genet* 13: e1006546. <https://doi.org/10.1371/journal.pgen.1006546> PMID: 28231283
11. Servetas SL, Kim A, Su H, Cha JH, Merrell DS. (2018) Comparative analysis of the Hom family of outer membrane proteins in isolates from two geographically distinct regions: The United States and South Korea. *HELICOBACTER* 23: e12461. <https://doi.org/10.1111/hel.12461> PMID: 29315985
12. Jung SW, Sugimoto M, Graham DY, Yamaoka Y. (2009) homB status of *Helicobacter pylori* as a novel marker to distinguish gastric cancer from duodenal ulcer. *J Clin Microbiol* 47: 3241–5. <https://doi.org/10.1128/JCM.00293-09> PMID: 19710266
13. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, et al. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *NATURE* 388: 539–47. <https://doi.org/10.1038/41483> PMID: 9252185
14. Fischer F, Robbe-Saule M, Turlin E, Mancuso F, Michel V, Richaud P, et al. (2016) Characterization in *Helicobacter pylori* of a Nickel Transporter Essential for Colonization That Was Acquired during Evolution by Gastric *Helicobacter* Species. *PLoS Pathog* 12: e1006018. <https://doi.org/10.1371/journal.ppat.1006018> PMID: 27923069
15. Ren S, Cai P, Liu Y, Wang T, Zhang Y, Li Q, et al. (2022) Prevalence of *Helicobacter pylori* infection in China: A systematic review and meta-analysis. *J Gastroenterol Hepatol* 37: 464–70. <https://doi.org/10.1111/jgh.15751> PMID: 34862656

16. He Y, Wang Y, Luan F, Yu Z, Feng H, Chen B, et al. (2021) Chinese and global burdens of gastric cancer from 1990 to 2019. *Cancer Med* 10: 3461–73. <https://doi.org/10.1002/cam4.3892> PMID: 33931958
17. Xue Z, Yang H, Su D, Song X, Deng X, Yu C, et al. (2021) Geographic distribution of the *cagA*, *vacA*, *iceA*, *oipA* and *dupA* genes of *Helicobacter pylori* strains isolated in China. *Gut Pathog* 13: 39. <https://doi.org/10.1186/s13099-021-00434-4> PMID: 34130751
18. Ernst FD, Stoof J, Horrevoets WM, Kuipers EJ, Kusters JG, van Vliet AH. (2006) NikR mediates nickel-responsive transcriptional repression of the *Helicobacter pylori* outer membrane proteins FecA3 (HP1400) and FrpB4 (HP1512). *Infect Immun* 74: 6821–8. <https://doi.org/10.1128/IAI.01196-06> PMID: 17015456
19. Hussein NR. (2011) A study of *Helicobacter pylori* outer-membrane proteins (hom) A and B in Iraq and Turkey. *J Infect Public Health* 4: 135–9. <https://doi.org/10.1016/j.jiph.2011.03.004> PMID: 21843859
20. Yanai A, Maeda S, Hikiba Y, Shibata W, Ohmae T, Hirata Y, et al. (2007) Clinical relevance of *Helicobacter pylori* *sabA* genotype in Japanese clinical isolates. *J Gastroenterol Hepatol* 22: 2228–32. <https://doi.org/10.1111/j.1440-1746.2007.04831.x> PMID: 18031386
21. Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, et al. (2003) Traces of human migrations in *Helicobacter pylori* populations. *SCIENCE* 299: 1582–5. <https://doi.org/10.1126/science.1080857> PMID: 12624269
22. Franceschi F, Zuccalà G, Roccarina D, Gasbarrini A. (2014) Clinical effects of *Helicobacter pylori* outside the stomach. *Nat Rev Gastroenterol Hepatol* 11: 234–42. <https://doi.org/10.1038/nrgastro.2013.243> PMID: 24345888
23. Youssefi M, Tafaghodi M, Farsiani H, Ghazvini K, Keikha M. (2021) *Helicobacter pylori* infection and autoimmune diseases; Is there an association with systemic lupus erythematosus, rheumatoid arthritis, autoimmune atrophy gastritis and autoimmune pancreatitis? A systematic review and meta-analysis study. *J Microbiol Immunol Infect* 54: 359–69. <https://doi.org/10.1016/j.jmii.2020.08.011> PMID: 32891538
24. Wang F, Meng W, Wang B, Qiao L. (2014) *Helicobacter pylori*-induced gastric inflammation and gastric cancer. *Cancer Lett* 345: 196–202. <https://doi.org/10.1016/j.canlet.2013.08.016> PMID: 23981572
25. Talebi Bezmin Abadi A, Rafiei A, Ajami A, Hosseini V, Taghvaei T, Jones KR, et al. (2011) *Helicobacter pylori* *homB*, but not *cagA*, is associated with gastric cancer in Iran. *J Clin Microbiol* 49: 3191–7. <https://doi.org/10.1128/JCM.00947-11> PMID: 21734027
26. Oleastro M, Cordeiro R, Yamaoka Y, Queiroz D, Mégraud F, Monteiro L, et al. (2009) Disease association with two *Helicobacter pylori* duplicate outer membrane protein genes, *homB* and *homA*. *Gut Pathog* 1: 12. <https://doi.org/10.1186/1757-4749-1-12> PMID: 19545429
27. Keikha M, Karbalaie M. (2021) Correlation between the geographical origin of *Helicobacter pylori* *homB*-positive strains and their clinical outcomes: a systematic review and meta-analysis. *BMC Gastroenterol* 21: 181. <https://doi.org/10.1186/s12876-021-01764-y> PMID: 33879080
28. Oleastro M, Cordeiro R, Ferrand J, Nunes B, Lehours P, Carvalho-Oliveira I, et al. (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: 1379–87. <https://doi.org/10.1086/592166> PMID: 18811585
29. de Jonge R, Pot RG, Loffeld RJ, van Vliet AH, Kuipers EJ, Kusters JG. (2004) The functional status of the *Helicobacter pylori* *sabB* adhesin gene as a putative marker for disease outcome. *HELICOBACTER* 9: 158–64. <https://doi.org/10.1111/j.1083-4389.2004.00213.x> PMID: 15068418
30. Chiarini A, Calà C, Bonura C, Gullo A, Giuliana G, Peralta S, et al. (2009) Prevalence of virulence-associated genotypes of *Helicobacter pylori* and correlation with severity of gastric pathology in patients from western Sicily, Italy. *Eur J Clin Microbiol Infect Dis* 28: 437–46. <https://doi.org/10.1007/s10096-008-0644-x> PMID: 18958508
31. Farzi N, Yadegar A, Aghdaei HA, Yamaoka Y, Zali MR. (2018) Genetic diversity and functional analysis of *oipA* gene in association with other virulence factors among *Helicobacter pylori* isolates from Iranian patients with different gastric diseases. *Infect Genet Evol* 60: 26–34. <https://doi.org/10.1016/j.meegid.2018.02.017> PMID: 29452293
32. You Y, Thorell K, He L, Yahara K, Yamaoka Y, Cha JH, et al. (2022) Genomic differentiation within East Asian *Helicobacter pylori*. *Microb Genom* 8. <https://doi.org/10.1099/mgen.0.000676> PMID: 35188454