Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Geographic Distribution of The cagA, vacA, iceA, oipA and dupA Genes of Helicobacter Pylori Strains Isolated in China

Zhijing Xue
Chinese Center for Disease Control and Prevention https://orcid.org/0000-0002-9459-6663
Hong Yang
Peking Union Medical College Hospital
Dongxing Su
Nanning Second Peoples Hospital
Xiangfeng Song
Rushan Peoples Hospital
Xin Deng
Yiyang central hospital
Changhong Yu
The First Affiliated Hospital of Jiamusi Medical University
Chunhua Sun
the peoples hospital of huzhu Tu Ethnic Autonomous county
Lihua He
Chinese Center for Disease Control and Prevention
Yuanhai You
Chinese Center for Disease Control and Prevention
Yanan Gong
Chinese Center for Disease Control and Prevention
Dongjie Fan
Chinese Center for Disease Control and Prevention
Lu Sun
Chinese Center for Disease Control and Prevention
Xiurui Han
Chinese Center for Disease Control and Prevention
Ruyue Fan
Chinese Center for Disease Control and Prevention
Kangle Zhai
Chinese Center for Disease Control and Prevention
Yaming Yang
Chinese Center for Disease Control and Prevention
Maojun Zhang
Chinese Center for Disease Control and Prevention
Xiaomei Yan
Chinese Center for Disease Control and Prevention
Jiaming Qian
Peking Union Medical College Hospital
Jianzhong Zhang (zhangjianzhong@icdc.cn)
Chinese Center for Disease Control and Prevention https://orcid.org/0000-0001-7056-8206

Research

Keywords: Helicobacter pylori, genotype, virulence genes, PCR, China

Posted Date: December 30th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-136079/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: There were geographical differences in the distribution of *Helicobacter pylori* (*H. pylori*) genotypes (*cagA*, *vacA*, *iceA*, *oipA* and *dupA*, et al). The population in different regions in China have grant different patterns of gastroduodenal diseases which are associated with these genotypes, but the geographical characteristics of *H. pylori* genotypes were still unknown.

Materials and Methods: Gastric biopsy specimens were obtained from 348 patients from five regions in China. The regional distribution was 89 patients from Shandong, 91 from Guangxi, 57 from Hunan, 58 from Qinghai and 53 from Heilongjiang. DNA extracted from cultured isolates were analyzed by polymerase chain reaction (PCR) to determine the presence of *cagA*, *vacA*, *iceA*, *oipA* and *dupA* genotypes.

Results: A total of 269 *H. pylori* isolates were obtained, of which 74 isolates were from Shandong, 78 from Guangxi, 46 from Hunan, 33 from Qinghai and 38 from Heilongjiang. The cagA gene was predominant in all the five regions (e.g. 100% in Hunan, Qinghai and Heilongjiang). The predominant vacA genotypes in the 269 isolates were s1a (88.1%) and m1(72.1%). vacA s1b was not detected in our study. In strains from Guangxi and Hunan, s1c was dominant; in contrast, s1a was dominant in Shandong, Qinghai and Heilongjiang. The prevalence of m1 strains in Heilongjiang (92.1%) was significantly higher (P<0.001) than in Shandong (60.8%) and Qinghai (51.5%). The dominant vacA subtype combination was s1a/m1 (62.8%) and detection of vacA s1a/m1 was significantly high 34 (89.5%) in Heilongjiang strains (P<0.001). The prevalence of iceA alleles in Hunan and Qinghai was much higher than that in the other three regions, and the difference was statistically significant. The *oipA*-positive strains were more prevalent in Guangxi (100%) and Hunan (100%) than in Qinghai (78.8%) (P<0.001). Conversely, the *dupA*-positive strains were less than half in Guangxi (15.4%) and Shandong (32.4%), whereas it was 73.9% in Hunan and 81.8% in Qinghai (P<0.001).

Conclusions: There are significant geographic differences in the distribution of *H. pylori* genotypes. These datas may be used to explain the gastroduodenal diseases patterns in different geographic regions of China.

Background

Helicobacter pylori (*H. pylori*) is a chronic infectious pathogen that can lead to gastroduodenal diseases such as chronic gastritis (CG), peptic ulcer disease (PUD), gastric cancer (GC) and mucosa associated lymphoid tissue (MALT) lymphoma [1]. Owing to the carcinogenicity of *H. pylori*, it was classified as a grade I carcinogen in humans by the World Health Organization [2]. It has been proved that more than 50% of the world's population are infected with *H. pylori* and its prevalence rates range from 20–40% in developed countries and up to 90% in China and other developing countries [3]. *H. pylori* is characterized by genetic diversity, but the clinical symptoms caused by different strains are variable and considered to be related to the genetic susceptibility and living environment of the host, mainly due to the bacterial virulence factors [4].

Several *H. pylori* virulent factors have been identified that play an important role in the pathogenicity of *H. pylori* such as *cagA*, *vacA*, *iceA*, *oipA* and *dupA* [5]. The *cagA* (cytotoxin-associated gene A) has been considered as an important carcinogen of *H. pylori* and *cagA*-positive strains can increase the risk of PUD or GC. There are EPIYA repeat sequences in its 3' variable region, which is the tyrosine phosphorylation site of CagA protein. According to the difference of the amino acid sequences flanking the EPIYA motifs, CagA 3' variable region can be divided into four different segments: EPIYA-A, EPIYA-B, EPIYA-C and EPIYA-D [6, 7].

The *vacA* (vacuolating cytotoxin gene) is associated with damaging epithelial cells by inducing the formation of vacuoles [8]. The *vacA* includes at least two different parts: the signal (s) region (s1a, s1b and s1c, s2) and the middle (m) region (m1, m2a and m2b) [9, 10]. The combination of *vacA* s and m region genotypes determines the production of cytotoxic activity and constitutes mosaic gene structure [11]. Strains with s1/m1 produce high levels of toxin in vitro, followed by s1/m2, while s2/m1 strains produce low toxicity and s2/m2 strains produce little or no toxin [12]. Some studies have shown that s1/m1 subtype is highly correlated with PUD and GC [13]. There are geographic differences in the distribution of *vacA* genotypes in different regions. Many researchers have shown that *vacA* s1a and s1c are predominant in Asia, northern Europe and USA, whereas s1b is common in South America, USA, Southern Europe and South Africa [10–16]. These differences may lead to the distinct prevalence of gastroduodenal diseases from different geographic regions. The *iceA* (induced by contact with epithelium) has two different alleles: *iceA1* and *iceA2*, which are distributed in different *H. pylori* strains [10]. Some studies have suggested that the *iceA1* may be related to PUD, whereas others have come to different conclusions [11]. This inconsistent conclusion may be due to geographical differences. The *oipA* (outer inflammatory protein) is closely related to the clinical symptoms, specifically manifested in the bacterial colonization density, severe neutrophil infiltration and high level of IL-8 [17]. Researches have shown that the prevence of *oipA* in duodenal ulcer (DU) and GC is higher, suggesting that *oipA* is not only associated with inflammation, but also the development of GC [18]. The *dupA* (duodenal ulcer promoting gene A), first recognized as a marker of *H. pylori* specific disease, can induce DU and inhibit GC [19].

Studies have shown that *H. pylori* infection is very common in China where the incidence rate of GC is higher than that in the western countries. There are many studies focusing on the detection of *H. pylori*, the prevalence and clinical outcomes at present. However, only a few studies regarded information on the relationship between *H. pylori* virulence genotypes and different geographic regions. We therefore investigated the distribution of *vacA*, *cagA*, *iceA*, *oipA* and *dupA* genotypes in different areas of China and compared the association among the genotypes.

Materials And Methods

Study subjects

A total of 348 patients were involved in this study including 89 patients from Rushan People's Hospital (Weihai, Shandong Province, China), 91 from the Second Nanning People's Hospital (Nanning, Guangxi Province, China), 57 from Yiyang Central Hospital (Yiyang, Hunan Province, China),58 from the People's Hospital of Huzhu Tu Ethnic Autonomous County (Haidong, Qinghai Province, China) and 53 from The First Affiliated Hospital of Jiamusi University (Jiamusi, Heilongjiang Province, China). Their gastric biopsy specimens were obtained during upper gastrointestinal endoscopy with informed consent. This study was approved by the Research and Ethical committees. One gastric biopsy was used for *H. pylori* culture and two for pathological examination.

H. pylori **culture and DNA extraction**

Gastric biopsy specimens were homogenized thoroughly in brain heart infusion (BHI) broth and then streaked onto the Karmali blood agar base plates under a biological safety cabinet (Thermo Scientific). The Karmali Agar base (Oxoid, CM 0935) was supplemented with 5% defibrinated sheep blood, and 1% combined antibiotics comprising of TMP (150 mg/L), vancomycin (125 mg/L), amphotericin B (100 mg/L) and polymyxin B (100 mg/L). The plates were incubated at 37 °C in a microaerobic atmosphere (5% O_2, 10% CO_2 and 85% N_2) for 3–5 days. *H. pylori* colonies were identified according to its morphological characteristics, negative Gram staining and positive for catalase, oxidase, and urease. The confirmed isolates were preserved in sterile BHI broth with 20% glycerol and frozen at -80 °C until the genomic DNA was extracted with the QIAamp DNA Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. The extracted DNA was stored at -20°C and used directly for polymerase chain reaction (PCR).

PCR amplification

The PCR reaction was carried out in a total volume of 25 μl containing 1 μl each of primer, 1 μl template DNA, 12.5 μl Go Taq® Green Master Mix (Promega, USA) and 9.5 μl nuclease-free water. Each PCR amplification was under the following conditions: initial denaturation at 94 °C for 5 min, 35 cycles of denaturation at 94 °C for 30 s, annealing at (T_m ± 5) °C for 30 s, extension at 72 °C for 40 s, and a final extension at 72 °C for 10 min. The presence of the *cagA*, *iceA*, *oipA* and *dupA* genes was determined by PCR as previously described [19, 22–24]. The genotypes of *vacA* s1a, s1b, s1c, s2, *vacA* m1, m2a, m2b, and *cagA* were also determined by PCR as previously described [4,19–21]. The amplified products were analyzed in 1% agarose gel containing 1 × TAE, stained with GelStain and visualized by electrophoresis at 110 V for 30 min using the gel documentation system (Bio-Rad, USA). Positive PCR products were sequenced using ABI 3730xl DNA sequencer and the sequences obtained were analyzed using MEGA (version 7.0.18, USA).

Statistical Analysis

Statistical datas were analyzed by SPSS software (version 20, Chicago, USA). The chi-square test and Fisher's exact test were used to assess the association among the genotypes and between specific genotypes and geographic distribution. P-value < 0.05 was considered of a statistically significant difference.

Results

A total of 269 *H. pylori* isolates out of 348 gastric biopsy specimens from five geographic regions in China were obtained, of which 74 isolates were from Shandong, 78 from Guangxi, 46 from Hunan, 33 from Qinghai and 38 from Heilongjiang. As is shown in Table 1, the highest number of *H. pylori* positives (85.7%) was obtained in Guangxi, followed by Shandong (83.1%), while Qinghai had the lowest *H. pylori* positives (56.9%). Virulence genes *cagA*, *vacA*, *iceA*, *oipA* and *dupA* were detected by PCR in all *H. pylori* obtained from culture isolates and *H. pylori* genotypes results were summarized in Table 1.

Table 1

Detection and distribution of *cagA*, *vacA*, *iceA*, *oipA* and *dupA* genes in China[a]

	No. of isolates[b]					
	SD (n = 89)	GX (n = 91)	HN (n = 57)	QH (n = 58)	HL (n = 53)	Total (n = 348)
H. pylori positive	74(83.1)	78(85.7)	46(80.7)	33(56.9)	38(71.7)	269(77.3)
cagA	66(89.2)	74(94.9)	46(100)	33(100)	38(100)	263(97.8)
vacA						
s1	69(93.2)	71(91)	44(95.7)	29(87.9)	38(100)	251(93.3)
s2	2(2.7)	0	1(2.2)	3(9.1)	2(5.3)	8(3)
m1	45(60.8)	56(71.8)	41(89.1)	21(63.6)	35(92.1)	198(73.6)
m2	66(89.2)	43(55.1)	39(84.8)	16(48.5)	30(78.9)	194(72.1)
s1m1	42(56.8)	49(62.8)	40(87)	14(42.4)	35(92.1)	180(66.9)
s1m2	62(83.8)	38(48.7)	37(80.4)	15(45.5)	30(78.9)	182(67.7)
s2m1	2(2.7)	0	1(2.2)	3(9.1)	2(5.3)	8(3)
s2m2	2(2.7)	0	1(2.2)	2(6.1)	2(5.3)	7(2.6)
iceA						
iceA1	59(79.7)	52(66.7)	44(95.7)	28(84.8)	29(76.3)	212(78.8)
iceA2	39(52.7)	66(84.6)	46(100)	31(97)	32(84.2)	214(79.6)
iceA1 + *iceA2*	27(36.5)	40(51.3)	44(95.7)	26(78.8)	23(60.5)	160(59.5)
oipA	73(98.6)	78(100)	46(100)	26(78.8)	35(92.1)	258(95.9)
dupA	24(32.4)	12(15.4)	34(73.9)	27(81.8)	25(65.8)	122(45.4)

[a]Isolates were from five regions of China: Shandong (SD), Guangxi (GX), Hunan (HN), Qinghai (QH) and Heilongjiang (HL).

[b]Values in parentheses are percentages.

Detection and distribution of virulence genes

cagA

Overall, 263 (97.8%) patients were infected with *cagA*-positive strains (Table 1). Of these *cagA*-positive strains, 66 (89.2%) were isolated from Shandong, 74 (94.9%) from Guangxi, 100% from Hunan, Qinghai and Heilongjiang. The prevalence of *cagA*-positive strains was lower in Shandong and Guangxi than in other three regions, but the difference was not statistically significant ($\chi^2 = 4.635$, $P > 0.05$). The sequencing results showed that CagA type was mainly East Asian type, accounting for 91.8% (236/257), whereas only 9 strains were Western type, and 55.6% (5/9) were isolated from Heilongjiang. Among 236 East Asian type CagA, 234 were of the ABD subtype and 2 of the AABD subtype. The distribution of CagA sequence types is shown in Table 2. The CagA type was mainly the ABD subtype, accounting for 91.1% and the ABD subtypes show a significant association with different regional areas ($\chi^2 = 19.772$, $P < 0.01$). Most of the Western strains were from Heilongjiang and there was a significant difference between the ABC subtype and Heilongjiang isolates ($\chi^2 = 15.512$, $P < 0.01$).

Table 2

Distribution of CagA types in different regions of China[a]

Region	EPIYA motif types[b]					Total
	AB	ABD[c]	AABD	ABC[c]	ABCC	
SD	9	55	1	1	0	66
GX	0	71	1	1	1	74
HN	0	46	0	0	0	46
QH	0	32	0	1	0	33
HL	3	30	0	5	0	38
Total	12(4.7)	234(91.1)	2(0.8)	8(3.1)	1(0.4)	257

[a] *cagA*-positive strains were from five regions of China: Shandong (SD), Guangxi (GX), Hunan (HN), Qinghai (QH) and Heilongjiang (HL).

[b] values in parentheses are percentages.

[c] indicating a significant difference between different regional areas($p < 0.01$).

vacA

The geographic distribution of *vacA* s- and m-alleles is shown in Table 1. In the 269 isolates, multiple *vacA* genotypes were detected in either the s-region (80.7%), the m-region (25.6%), or both (17.1%). The prevalence of *vacA* s- and m- region genotypes varied between the five regions of China. The most common *vacA* s-region genotype was s1 (93.3%): 93.2% of isolates from Shandong, 91% from Guangxi, 95.7% from Hunan, 87.9% from Qinghai and 100% from Heilongjiang. For the *vacA* m-region, 198 patients (73.6%) were infected with m1 *H. pylori* strains. The prevalence of *vacA* m1 genotype from Heilongjiang was the highest and that of Shandong was the lowest (92.1% and 60.8%). In the combination of *vacA* s and m region genotypes, s1m2 was predominant and found in 182 (67.7%) of the *H. pylori* strains and the s1/m1 was detected in 180 (66.9%).

For the *vacA* s subtype, the dominant genotype in 269 isolates was s1a (88.1%). The *vacA* s1b was not detected in our study. Among the 74 *vacA* s1 strains from Shandong, the prevalence of s1a (86.5%) and s1c (82.4%) was nearly similar. The distribution of s1a (85.9%) and s1c (89.7%) in Guangxi resembled that of Shandong. Of the 46 s1 strains from Hunan, 40 (87%) were s1a and 43 (93.5%) were s1c. In Qinghai, 87.9% of the s1 strains were s1a and 69.7% were s1c. Among strains from Heilongjiang, 37 (97.4%) of the 38 s1 strains were s1a, 33 (86.8%) was s1c. Of all the cultures, 252 (93.7%) could be genotyped as either m1, m2a or m2b. Among the 152 s1 strains from Guangxi, Hunan and Heilongjiang, m1 (76.5%) was more prevalent than m2a (48.1%). Of the 69 s1 strains from Shandong, the prevalence of m1 (55.4%) and m2a (59.6%) was about similar. Among the 29 *vacA* s1 strains from Qinghai, 14 (42.4%) contained m1 and 11 (33.3%) were m2a (Table 3 and Fig. 1).

Table 3
Geographic distribution of *vacA* subtypes of s- and m-region in China[a,b]

Region	Total	s1a			s1c			s2		Multiple genotypes		
		m1	m2a	m2b	m1	m2a	m2b	m2a	m2b	s	m	s + m
SD	74	39(52.7)	39(52.7)	31(41.9)	35(47.3)	38(51.4)	30(40.5)	2(2.7)	2(2.7)	50(67.6)	5(6.8)	4(5.4)
GX	78	46(60)	28(35.9)	14(18)	49(62.8)	32(41)	16(20.5)	0	0	65(83.3)	9(11.5)	7(9)
HN	46	36(78.3)	22(47.8)	20(43.5)	39(84.8)	23(50)	21(45.7)	1(2.2)	0	39(84.8)	9(19.6)	8(17.4)
QH	33	14(42.4)	10(30.3)	7(21.2)	11(33.3)	9(27.3)	6(18.2)	3(9.1)	2(6)	26(78.8)	8(24.2)	6(18.2)
HL	38	34(89.5)	22(57.9)	17(44.7)	30(79)	18(47.4)	15(39.5)	2(5.3)	2(5.3)	30(79)	27(71.1)	21(55.3)
Total	269	169(62.8)	121(45)	89(33.1)	164(61)	120(44.6)	88(32.7)	8(3)	6(2.2)	210(78.1)	58(25.6)	46(17.1)

[a] *H. pylori* isolates were from five regions of China: Shandong (SD), Guangxi (GX), Hunan (HN), Qinghai (QH) and Heilongjiang (HL).

[b] Values in parentheses are percentages.

The geographic distribution of *vacA* s and m subtypes was different in the five regions of China (Table 3 and Fig. 1). The prevalence of s1c in the isolates from Hunan was 93.5%, significantly higher ($\chi^2 = 10.760$, $P < 0.05$) than that in Shandong (82.4%), Guangxi (89.7%), Qinghai (69.7%) and Heilongjiang (86.8%). In contrast, s1a was more frequent in isolates from Heilongjiang (97.4%) than the other four regions, but no statistical significance was noted ($\chi^2 = 3.718$, $P > 0.05$). The prevalence of m1 strains in Heilongjiang (92.1%) was significantly higher ($\chi^2 = 15.668$, $P < 0.001$) than in Shandong (60.8%) and Qinghai (51.5%). The frequency of the m2b subtype in Hunan (50%) was significantly higher ($\chi^2 = 10.714$, $P < 0.01$) than in Guangxi (23.1%) and Qinghai (24.2%) and there was no significant difference in prevalence of *vacA* m2a ($\chi^2 = 5.450$, $P > 0.05$). We also examined s1a/m1, s1c/m1, s1/m2a, s1/m2b different combinations in patients according to analysis of the *vacA* subtypes. The dominant *vacA* subtype combination in the five regions was s1a/m1 (62.8%) and detection of *vacA* s1a/m1 was significantly high 34 (89.5%) in Heilongjiang strains ($\chi^2 = 32.218$, $P < 0.001$).

iceA

Of the 269 isolates, *iceA1* was found in 212 isolates (78.8%) and *iceA2* in 214 isolates (79.6%). In these isolates 160 (59.5%) contained both *iceA1* and *iceA2*, which indicating mixed infection. The distribution of *iceA* varied considerably for isolates from five different geographic regions of China. The *iceA1* was present in 95.7% and 84.8% of *H. pylori* strains isolated from Hunan and Qinghai, respectively, whereas only 66.7% of isolates from Guangxi were infected with *iceA1* positive strains. This difference was statistically significant ($\chi^2 = 15.510$, $P < 0.001$). The *iceA2* frequency was significantly more prevalent in Hunan (100%) and Qinghai (97%) strains than in Shandong (52.7%) strains ($\chi^2 = 42.147$, $P < 0.001$). The prevalence rates of mixed *iceA* genotypes in Shandong, Guangxi, Hunan, Qinghai and Heilongjiang were 36.5%, 51.3%, 95.7%, 78.8% and 60.5%, respectively and the mixed *iceA* was more prevalent in Hunan strains than in the other four regions strains ($\chi^2 = 48.502$, $P < 0.001$) (Table 1).

oipA

258 (95.9%) patients were infected with *oipA*-positive strains (Table 1). The *oipA*-positive isolates were present in 100% of Guangxi and Hunan isolates and the prevalence rates in Shandong, Qinghai and Heilongjiang were 98.6%, 78.8% and 92.1%, respectively. The *oipA* gene was more prevalent in Guangxi and Hunan strains than in Qinghai strains ($\chi^2 = 27.531$, $P < 0.001$).

dupA

122 (45.4%) patients were infected with *dupA*-positive strains (Table 1). The *dupA*-positive isolates were present in 73.9% of Hunan, 81.8% of Qinghai and 65.8% of Heilongjiang. In contrast, only 32.4% of Shandong and 15.4% of Guangxi isolates were infected with *dupA*-positive *H. pylori* ($\chi^2 = 72.497$, $P < 0.001$).

Association among the genotypes in H. pylori **strains**

The prevalence of *cagA* gene is relatively high in China and it was independent of *iceA* and *dupA* gene. The *cagA* was present in 204 out of 212 *iceA1* genes (96.2%), 209 out of 214 *iceA2* genes (97.7%) and 117 out of 122 *dupA* genes (95.9%). Table 4 shows the association of *vacA* with *dupA* and *iceA* genotypes in *H. pylori* strains. It was found that the infection rate of *vacA* s1c was higher than *vacA* s1a in *dupA* and *iceA* positives strains. Similarly, *vacA* m1 positive strains detected a significantly high *dupA* 91 (74.6%), *iceA1* 154(72.6%) and *iceA2* 162 (75.7%) genotypes. Furthermore, there was a high prevalence rate of *dupA* 71.3% (87/122), *iceA1* 66.5% (141/212) and *iceA2* 77.1% (165/214) in the *vacA* s1a/m1 positive strains.

Table 4
Association of *vacA* with *dupA* and *iceA* genotypes

	SD n (%)			GX n (%)			HN n (%)			QH n (%)			
vacA	*dupA*	*iceA1*	*iceA2*	*dupA*	*iceA1*	*iceA2*	*dupA*	*iceA1*	*iceA2*	*dupA*	*iceA1*	*iceA2*	*d*
s-region													
s1a	19(25.7)	52(70.3)	33(44.6)	10(12.8)	42(53.8)	58(74.4)	28(60.9)	39(84.8)	40(87)	23(70)	23(70)	25(75.8)	2
s1c	19(25.7)	48(64.9)	28(37.8)	9(11.5)	47(60.2)	60(76.9)	32(70)	41(89.1)	43(93.5)	20(60.6)	18(54.5)	21(63.6)	2
m-region													
m1	17(23)	32(43.2)	28(37.8)	7(9)	41(52.6)	47(60.3)	30(65.2)	39(84.8)	41(89.1)	14(42.4)	15(45.5)	17(51.5)	2
m2a	17(23)	31(41.9)	24(32.4)	3(3.8)	21(26.9)	28(35.9)	16(34.8)	23(50)	25(54.3)	6(18.2)	10(30.3)	12(36.4)	1
m2b	9(12.2)	29(39.2)	21(28.4)	3(3.8)	12(15.4)	17(21.8)	18(39.1)	23(50)	23(50)	7(21.2)	8(24.2)	7(21.2)	1
s/m region													
s1am1	15(20.3)	26(35.1)	26(35.1)	9(11.5)	35(44.9)	40(51.3)	24(52.2)	35(76.1)	40(87)	11(33.3)	11(33.3)	13(39.4)	2
s1am2a	13(17.6)	28(37.8)	21(28.4)	4(5.1)	17(21.8)	25(32.1)	13(28.3)	20(43.5)	22(47.8)	6(18.2)	7(21.2)	9(27.3)	1
s1am2b	5(6.8)	27(36.5)	17(23)	3(3.8)	9(11.5)	13(16.7)	14(30.4)	20(43.5)	20(43.5)	6(18.2)	7(21.2)	5(15.2)	1
s1cm1	14(18.9)	26(35.1)	23(31.1)	9(11.5)	36(46.2)	41(52.6)	28(60.9)	36(78.3)	39(84.8)	9(27.3)	8(24.2)	11(33.3)	2
s1cm2a	16(21.6)	24(32.4)	20(27)	4(5.1)	19(24.4)	28(35.9)	14(30.4)	26(56.5)	23(50)	6(18.2)	4(12.1)	8(24.2)	1
s1cm2b	7(9.5)	24(32.4)	16(21.6)	3(3.8)	10(12.8)	15(19.2)	16(34.8)	21(45.7)	21(45.7)	6(18.2)	6(18.2)	5(15.2)	1

We examined different combinations based on the analysis of the *vacA* subtypes (s1a, s1c, m1, m2a, m2b), *cagA* and *iceA* (*iceA1*, *iceA2*) in patients. The most prevalent combination s1cm1/*cagA*$^+$/*iceA2* was present in 51.7% (139/269) including 25.7% (19/74) of Shandong, 56.4% (44/78) of Guangxi, 84.8% (39/46) of Hunan, 33.3% (11/33) of Qinghai and 68.4% (26/38) of Heilongjiang. The predominant common combination genotypes in Guangxi and Hunan was s1cm1/*cagA*$^+$/*iceA2* (56.4% and 84.8%, $\chi^2 = 5.409$, $P < 0.05$) and that was s1am1/*cagA*$^+$/*iceA2* (42.4% and 73.7%, $\chi^2 = 7.143$, $P < 0.01$) in Qinghai and Heilongjiang, while the most common in Shandong was s1am2a/*cagA*$^+$/*iceA1* (36.5%) (Table 5).

Table 5
Combination genotypes of *cagA*, *vacA* and *iceA* in China

Combination	No. of isolates					
	SD n (%)	GX n (%)	HN n (%)	QH n (%)	HL n (%)	Total n (%)
s1am1/*cagA*$^+$/*iceA1*	26(35.1)	33(42.3)	35(76.1)	12(36.4)	26(68.4)	132(49.1)
s1am1/*cagA*$^+$/*iceA2*	22(29.7)	37(47.4)	36(78.3)	14(42.4)	28(73.7)	137(50.9)
s1am2a/*cagA*$^+$/*iceA1*	27(36.5)	19(24.4)	21(45.7)	8(24.2)	18(47.4)	93(34.6)
s1am2a/*cagA*$^+$/*iceA2*	19(25.7)	26(33.3)	22(47.8)	10(30.3)	17(44.7)	94(34.9)
s1am2b/*cagA*$^+$/*iceA1*	23(31.1)	8(10.3)	20(43.5)	7(21.2)	11(28.9)	69(25.7)
s1am2b/*cagA*$^+$/*iceA2*	15(20.3)	10(12.8)	20(43.5)	6(18.2)	14(36.8)	65(24.2)
s1cm1/*cagA*$^+$/*iceA1*	24(32.4)	39(50)	37(80.4)	9(27.3)	24(63.2)	133(49.4)
s1cm1/*cagA*$^+$/*iceA2*	19(25.7)	44(56.4)	39(84.8)	11(33.3)	26(68.4)	139(51.7)
s1cm2a/*cagA*$^+$/*iceA1*	26(35.1)	23(29.5)	21(45.7)	6(18.2)	16(42.1)	92(34.2)
s1cm2a/*cagA*$^+$/*iceA2*	19(25.7)	30(38.5)	23(50)	8(24.2)	15(39.5)	95(35.3)
s1cm2b/*cagA*$^+$/*iceA1*	21(28.4)	11(14.1)	21(45.7)	6(18.2)	11(28.9)	70(26)
s1cm2b/*cagA*$^+$/*iceA2*	13(17.6)	14(17.9)	21(45.7)	5(15.2)	14(36.8)	67(24.9)

Discussion

This study aims to investigate the distribution of virulence genes of *H. pylori* isolated from patients living in different geographical regions of China and the association among these genotypes. *H. pylori* was detected for the presence of the genes for *cagA*, *vacA*, *iceA*, *oipA* and *dupA*. The present study demonstrates that there was obvious geographical diversity of *H. pylori* genotypes within China, emphasizing that even within a country genetic diversity still exists. The *H. pylori* isolates were cultured from five different geographic regions of China, Shandong in the east, Guangxi in the south, Hunan in the central, Qinghai in the west and Heilongjiang in the north. As far as we know, this was the first comprehensive study on the distinct virulence genes of *H. pylori* in different geographical regions of China and the association among genotypes.

Geographic distribution versus genotypes

In East Asia, more than 90% of *H. pylori* isolates carry *cagA* gene, which may be the reason why the incidence rate of GC in East Asian countries is higher than that in western countries. In the present study, 97.8% patients were infected with *cagA*-positive strains. This result was similar to studies in other Asian countries and some regions of China where the prevalence of *cagA*-positive strains was above 90% [25, 26]. However, this was different from reports in some European and American countries where the prevalence of *cagA*-positive strains ranged from 50–70% (Fig. 2A) [14,27–30]. Furthermore, we found that the majority of CagA types were East Asian type, only 3.5% were Western type.

The presence of the *vacA* genotypes was also different in distinct geographical regions. 93.3% of *H. pylori* carried the *vacA* s1 genotype similar to previous studies in other regions of China [31, 32]. In the present study, there was a high prevalence of s1a 88.1% and s1c 85.5% in the *vacA*-positive strains. The result was slightly different from some reports in which the prevalence of s1a and s1c was a little lower [14, 33]. The *vacA* s1b was not detected in our study, whereas the prevalence of s1b subtype was almost 100% in South America, 80% in Spain and Portugal strains, very few in East Asia [14] (Fig. 2B). The *vacA* s2 was prevalent in Africa [34] and consistent with studies in some European and American countries [14], but the s2 detected in this study was very low, further revealing the geographic diversity of *vacA* gene (Fig. 2C). The presence of *vacA* m1 strains was significantly higher in Heilongjiang, which may be the reasons for the high incidence of gastric cancer. These findings were different from some countries such as Japan, Korean, Singapore and some European and American strains [11, 14, 35] (Fig. 2D), suggesting the differences between Chinese strains and foreign strains. The *vacA* s1m1 genotype was predominant in Heilongjiang and s1m2 strains in Shandong, which was consistent with other studies in China, such as Xi'an, Beijing, Taiwan and Hong Kong [33,36–38], but different from studies in America, Netherlands and Germany [14, 28, 39]. The reason for the difference may be association with geographic diversity of *vacA* genotypes.

The *iceA1* gene was common in Japan and Korea while the *iceA2* gene was predominant in the America, Colombia, Brazil, Europe and South Africa [9, 40, 41]. The prevalence of *iceA1* was 78.8% in all studied strains, consistent with studies reported from Thailand, Korea and Tunisia [42–44]. The *oipA* is closely related to severe inflammatory response and the induction of IL-8 secretion. In the present study, *oipA* was present in most strains, which was similar to previous studies [45]. The *dupA* is considered as a marker of duodenal diseases but in some studies the relationship is not linked [46]. The presence of *dupA* gene was also different in different geographical regions (e.g. 84.8% in the South Africa, 43.7% in the Belgium and 70% in the United States) [47]. Similarly, in the present study, the prevalence of the *dupA* was different in different regions in China. We detected 81.8% *dupA* positive isolates in Qinghai, while the lower prevalence

of *dupA* (15.4%) was in Guangxi and 32.4% in Shandong. The reasons for the difference in prevalence of *dupA* gene in the five geographic regions in China is unclear.

Gastroduodenal diseases versus genotypes

Studies have shown that more than 50% of the world's population are infected with *H. pylori*, but not all people develop gastrointestinal diseases, which is associated with living environment, host factors and bacterial virulence genes. *cagA*, as an important carcinogen, was reported to be closely related to PUD and GC as previously described. There are obvious regional differences in the incidence of GC, mainly in northeast China, Shandong, Henan and other regions. According to the data of the national disease surveillance system in 2018, the mortality rate of GC was the highest in eastern China, accounting for 22.71%, while in central and western China, the mortality rate of gastric cancer was 18.97% and 16.06%, respectively.

In the present study, the majority virulent genotypes of *H. pylori* isolated from Shandong were *vacA* s1m2/*cagA*/*iceA*1 positive. This genotype combination may be responsible for the high incidence of gastrointestinal diseases in Shandong Province. Some studies showed that *vacA* s1m1 and *cagA* positive strains were more frequently isolated from GC patients than from PUD and MALT patients [48]. Among all the five regions studied, detection of *vacA* s1m1/*cagA* was significantly higher in Heilongjiang than in other regions. This may be the reason for high incidence of GC in Heilongjiang. Some researchers considered that the *iceA*1 was more common in patients with PUD while the *iceA*2 was most frequently isolated from CG patients [49]. Hussein proved that *dupA* gene was a high risk factor for DU by detecting *dupA* gene from 2358 samples [24]. In this study, the presence of *iceA* and *dupA* was significantly higher in Hunan and Qinghai where the incidence rate of CG and PUD may be higher.

Conclusions

The present study demonstrated that there were significant geographic distribution differences of *H. pylori* genotypes in China, which can lead to gastroduodenal diseases differences in different regions. The next step is to clearly reveal the relationship between *H. pylori* genotypes and clinical outcomes. Researches on *H. pylori* virulence factors in China are important for a clinical and epidemiological survey to better understand the pathogenic mechanism.

Abbreviations

H. pylori: *Helicobacter pylori*

PCR: polymerase chain reaction

CG: chronic gastritis

GC: gastric cancer

PUD: peptic ulcer disease

MALT: mucosal-associated lymphoid tissue

cagA: cytotoxin-associated gene A

vacA: vacuolating cytotoxin gene

iceA: induced by contact with epithelium

oipA: outer inflammatory protein

dupA: duodenal ulcer promoting gene A

Declarations

Availability of data and materials

Not applicable.

Acknowledgements

Not applicable.

Funding

This work was supported by the Health Research and Special Projects Grant of China (No. 201002020, No.201502005) and the National Science and Technology Major Project of China (2018ZX10712-001).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Full consent is given for publication in Gut Pathogens.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ZX and HY developed the idea, designed the study, analyzed the datas and drafted the manuscript. DS, XS, XD, CY and CS collected the samples. LH, YY, YG, DF and LS performed the DNA extraction. XH, RF, KZ and YY analyzed the datas, MZ and XY designed the study and analyzed the results. JQ and JZ designed the study, reviewed and revised the manuscript. All authors read and approved the final manuscript.

References

1. Ernst PB, Gold BD. The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastroduodenal ulcer and gastric cancer. Annu Rev Microbiol. 2000;54(1):615-40.
2. Olfat FO, Zheng Q, Oleastro M, V Petra B Thomas, K Riita, et al. Correlation of the *Helicobacter pylori* adherence factor BabA with duodenal ulcer disease in four European countries. FEMS Immunol Med Microbiol. 2005;44(2):151-6.
3. Hooi JKY, Lai WY, Ng WK, Suen MMY Underwood FE, Tanyingoh D, et al. Global Prevalence of *Helicobacter pylori* Infection: Systematic Review and Meta-Analysis. Gastroenterol. 2017;23(4):420-9.
4. Doorn LJ, Figueiredo C, Sanna R, Plaisier A, Quint W. Clinical relevance of the *cagA*, *vacA*, and *iceA* status of *Helicobacter pylori*. Gastroenterol. 1998;115(1):58-62.
5. Mobley J, Harry LT. Defining *Helicobacter pylori* as a Pathogen: Strain Heterogeneity and Virulence. Am J Med. 1996;100(46):2S-11S.
6. Kanada R, Uchida T, Tsukamoto Y, Nguyen LT, Moriyama M. Genotyping of the *cagA* gene of *Helicobacter pylori* on immunohistochemistry with East Asian *cagA*-specific antibody. Pathol Int. 2008;58(2):218-25.
7. Hatakeyama M. Anthropological and clinical implications for the structural diversity of the Helicobacter pylori CagA oncoprotein. Cancer Sci. 2011;102(9):36-43.
8. Ciara U, Rainer H. *vacA*'s induction of *vacA*-containing vacuoles (VCVs) and their immunomodulatory activities on human T cells. Toxins. 2016;8(6):190-4.
9. Tanih NF, Mcmillan M, Naidoo N, Ndip LM, Weaver LT, Ndip RN. Prevalence of *Helicobacter pylori vacA*, *cagA* and *iceA* genotypes in South African patients with upper gastrointestinal diseases. Acta Tropica. 2010;116(1):68-73.
10. Tan HJ, Rizal AM, Rosmadi MY, Goh KL. Distribution of *Helicobacter pylori cagA*, *cagE* and *vacA* in different ethnic groups in Kuala Lumpur, Malaysia. J Gastroenterol Hepatol. 2010;20(4):589-94.
11. Yamaoka Y. Relationship between *Helicobacter pylori iceA*, *cagA*, and *vacA* status and clinical outcome: studies in four different countries. J Clin Microbiol. 1999;37(7):2274-9.
12. Franco AT, Johnston E, Krishna U, Yamaoka Y, Israel DA, Nagy TA, et al. Regulation of gastric carcinogenesis by *Helicobacter pylori* virulence factors. Cancer Res. 2008;68(2):379-87.
13. Atherton JC, Peek RM, Tham KT, Cover TL, Blaser MJ. Clinical and pathological importance of heterogeneity in *vacA*, the vacuolating cytotoxin gene of *Helicobacter pylori*. Gastroenterol. 1997;112(1):92-9.
14. Doorn LJV, Figueiredo C, Francis M, Pena S, Quint WGV. Geographic distribution of *vacA* allelic types of *Helicobacter pylori*. Gastroenterol. 1999;116(4):823-30.
15. Ito Y, Azuma T, Ito S, Miyaji H, Kuriyama M. Analysis and typing of the *vacA* gene from *cagA*-positive strains of *Helicobacter pylori* isolated in Japan. J Clin Microbiol. 1997;35(7):1710-4.
16. Shimoyama T, Yoshimura T, Mikami T. Evaluation of *Helicobacter pylori vacA* genotype in Japanese patients with gastric cancer. J Clin Pathol. 1998;51(4):299-301.
17. Yamaoka Y, Kikuchi S, Hala MZ, Gutierrez O, Osato MS, Graham DY. Importance of *Helicobacter pylori oipA* in clinical presentation, gastric inflammation, and mucosal interleukin 8 production. Gastroenterol. 2002; 123(2):414-24.
18. Yamaoka Y. *Helicobacter pylori* outer membrane proteins and gastroduodenal disease. Gut. 2006;55(6):775-81.
19. Chen CY, Wang FY, Wan HJ, Jin XX, Wei J, Wang ZK, et al. Amino acid polymorphisms flanking the EPIYA-A motif of *Helicobacter pylori* CagA C-terminal region is associated with gastric cancer in east China: experience from a single center. J Dig Dis. 2013;14(7):358-65.
20. Selander RK, Caugant DA, Ochman H, Musser JM, Whittam TS. Methods of Multilocus Enzyme Electrophoresis for bacterial Population genetics and Systematics. Appl Environ Microb. 1986;51(5):873-84.
21. Lu H, Hsu PI, Graham DY, Yamaoka Y. Duodenal ulcer promoting gene of *Helicobacter pylori*. Gastroenterol. 2005;128(4):833-48.
22. Jr PR, Thompson SA, Donahue JP, Tham KT, Atherton JC, Blaser MJ, et al. Adherence to gastric epithelial cells induces expression of a *Helicobacter pylori* gene, *iceA*, that is associated with clinical outcome. P Assoc Am Physician. 1998;110(6):531-44.
23. Breurec S, Michel R, Seck A, Brisse S, Dieye FB, Garin B, et al. Clinical relevance of *cagA* and *vacA* gene polymorphisms in *Helicobacter pylori* isolates from Senegalese patients. Clin Microbiol Infect. 2012;18(2):153-9.

24. Salih AM, Goreal A, Hussein NR, Bdullah SM, Assafi M. The distribution of *cagA* and *dupA* genes in *Helicobacter pylori* strains in Kurdistan region, northern Iraq. Ann Saudi Med. 2013;33(3):290-3.
25. Pan ZJ, Berg DE, Hulst RWMVD，Su WW, Raudonikiene A, Xiao SD, et al. Prevalence of vacuolating cytotoxin production and distribution of distinct *vacA* alleles in *Helicobacter pylori* from China. J Infect Dis. 1998;178(1):220-6.
26. Yang H, Wu SV, Pichuantes S, Song M, Wang J, Zhou D, et al. High prevalence of *cagA*-positive strains in *Helicobacter pylori*-infected, healthy, young Chinese adults. J Gastroenterol Hepatol. 1999;14(5):476-80.
27. van DLJ, Figueiredo C, Sanna R, Plaisier A, Quint W. Clinical relevance of the *cagA*, *vacA*, and *iceA* status of *Helicobacter pylori*. Gastroenterol. 1998;115(1):58-66.
28. Mukhopadhyay AK, Kersulyte D, Jeong JY, Datta S, Ito Y, Chowdhury A, et al. Distinctiveness of genotypes of *Helicobacter pylori* in Calcutta, India. J Bacteriol. 2000;182(11):3219-27.
29. Owen RJ, Peters TM, Varea R, Teare EL, Saverymuttu S. Molecular epidemiology of *Helicobacter pylori* in England: prevalence of cag pathogenicity island markers and IS605 presence in relation to patient age and severity of gastric disease. FEMS Immunol Med Microbiol. 2001;30(1):65-71.
30. Miehlke S, Kirsch C, Agha-Amiri K, Günther T, Bayerdrffer E. The *Helicobacter pylori vacA* s1, m1 genotype and *cagA* is associated with gastric carcinoma in Germany. Int J Cancer. 2000;87(3):322-7.
31. He Y, Hu PJ, He XX, Zeng ZR, Chen W, Peng XZ. Subtypes of *H. pylori vacA* gene in Guangdong area and the association with gastroduodenal disease. Chin J Gastroenterol. 2000;5(3):90-2(Chinese).
32. Lee XB, Liu WZ, Xiao SD, Xu WW. Genetic analysis and clinical evaluation of *vacA* and *cagA* 3′ region of *H. pylori* Shanghai Chin J Gastroenterol. 2000;5(2):86-9 (Chinese).
33. Wang J, van Doorn LJ, Robinson PA, Ji X，Wang D, Wang Y, et al. Regional variation among *vacA* alleles of *Helicobacter pylori* in China. J Clin Microbiol. 2003;41(5):1942-5.
34. Assumpcao MB, Martins LC, Barbosa M, Antonia K, Santos D, Sintia B, et al. *Helicobacter pylori* in dental plaque and stomach of patients from Northern Brazil. World J Gastroenterol. 2010;16(4):3033-9.
35. Zheng PY, Hua J, Yeoh KG, Ho B. Association of peptic ulcer with increased expression of Lewis antigens but not *cagA*, *iceA*, and *vacA* in *Helicobacter pylori* isolates in an Asian population. Gut. 2000;47(1):18-22.
36. Wong BC, Yin Y, Berg DE, Xia HH，Zhang JZ, Wang WH, et al. Distribution of distinct *vacA*, *cagA* and *iceA* alleles in *Helicobacter pylori* in Hong Kong. Helicobacter. 2001;6(2):317-24.
37. Perng CL, Lin HJ, Sun IC, Tseng GY. *Helicobacter pylori cagA*, *iceA* and *vacA* status in Taiwanese patients with peptic ulcer and gastritis. J Gastroenterol Hepatol. 2003;18(5):1244-9.
38. Qiao W, Hu JL, Xiao B, Wu KC，Peng DR，Atherton JC, et al. *cagA* and *vacA* genotype of *Helicobacter pylori* associated with gastric diseases in Xi'an area. World J Gastroenterol. 2003;9(7):1762-6.
39. Miernyk K, Morris J, Bruden D, Mcmahon B, Hurlburt D, Sacco F, et al. Characterization of *Helicobacter pylori cagA* and *vacA* genotypes among Alaskans and their correlation with clinical disease. J Clin Microbiol. 2011;49(7):3114-21.
40. Podzorski RP, Podzorski DS, Wuerth A, Tolia V. Analysis of the *vacA*, *cagA*, *cagE*, *iceA*, and *babA2* genes in *Helicobacter pylori* from sixty-one pediatric patients from the Midwestern United States. Diagn Microbiol Infect Dis. 2003;46(9):83-8.
41. Ashour AA, Collares GB, Mendes EN, Ribeiro DGV, Queiroz DMDM, Magalhaes PP, et al. *iceA* genotypes of *Helicobacter pylori* strains isolated from Brazilian children and adults. J Clin Microbiol. 2001;39(3):1746-50.
42. Kim SY, Woo CW, Lee YM, Son BR，Kim JW，Chae HB，et al. Genotyping *cagA*, *vacA* subtype, *iceA1*, and *babA* of *Helicobacter pylori* isolates from Korean patients, and their association with gastroduodenal diseases. J Korean Med Sci. 2001;16(4):579-84.
43. Chomvarin C, Namwat W, Chaicumpar K, Mairiang P，Sangchan A, Sripa B, et al. Prevalence of *Helicobacter pylori vacA*, *cagA*, *cagE*, *iceA* and *babA2* genotypes in Thai dyspeptic patients. Int J Infect Dis. 2008;12(5):30-6.
44. Ben MK, Fendri C, Zribi M, Masmoudi A, Burucoa C. Prevalence of *Helicobacter pylori vacA*, *cagA*, *iceA* and *oipA* genotypes in Tunisian patients. Ann Clin Microbiol Antimicrob. 2010;9(3):10-3.
45. Si YG, Sun B, Wang SH. Detection of *Helicobacter pylori* virulence genes from different digestive diseases and its clinical significance. Chin J Nosocomiol. 2019;29(4):489-97(Chinese).
46. Pereira WN, Ferraz MA, Zabaglia LM, De RW, Orcini WA. Association among *H. pylori* virulence markers *dupA*, *cagA* and *vacA* in Brazilian patients. J Venom Anim Toxins Incl Trop Dis. 2014;20(1):1-5.
47. Argent RH, Burette A, Miendje DVY, Atherton JC. The presence of *dupA* in *Helicobacter pylori* is not significantly associated with duodenal ulceration in Belgium, South Africa, China, or North America. Clin Infect Dis. 2007;45(9):1204-6.
48. Blaser MJ. Intrastrain differences in *Helicobacter pylori*: a key question in mucosal damage? Ann Med. 1995; 27(2):559-63.
49. Momenah AM, Tayeb MT. Relationship between *Helicobacter pylori vacA* genotypes status and risk of peptic ulcer in Saudi patients. Saudi Med J. 2006;27(6):804-7.

Figures

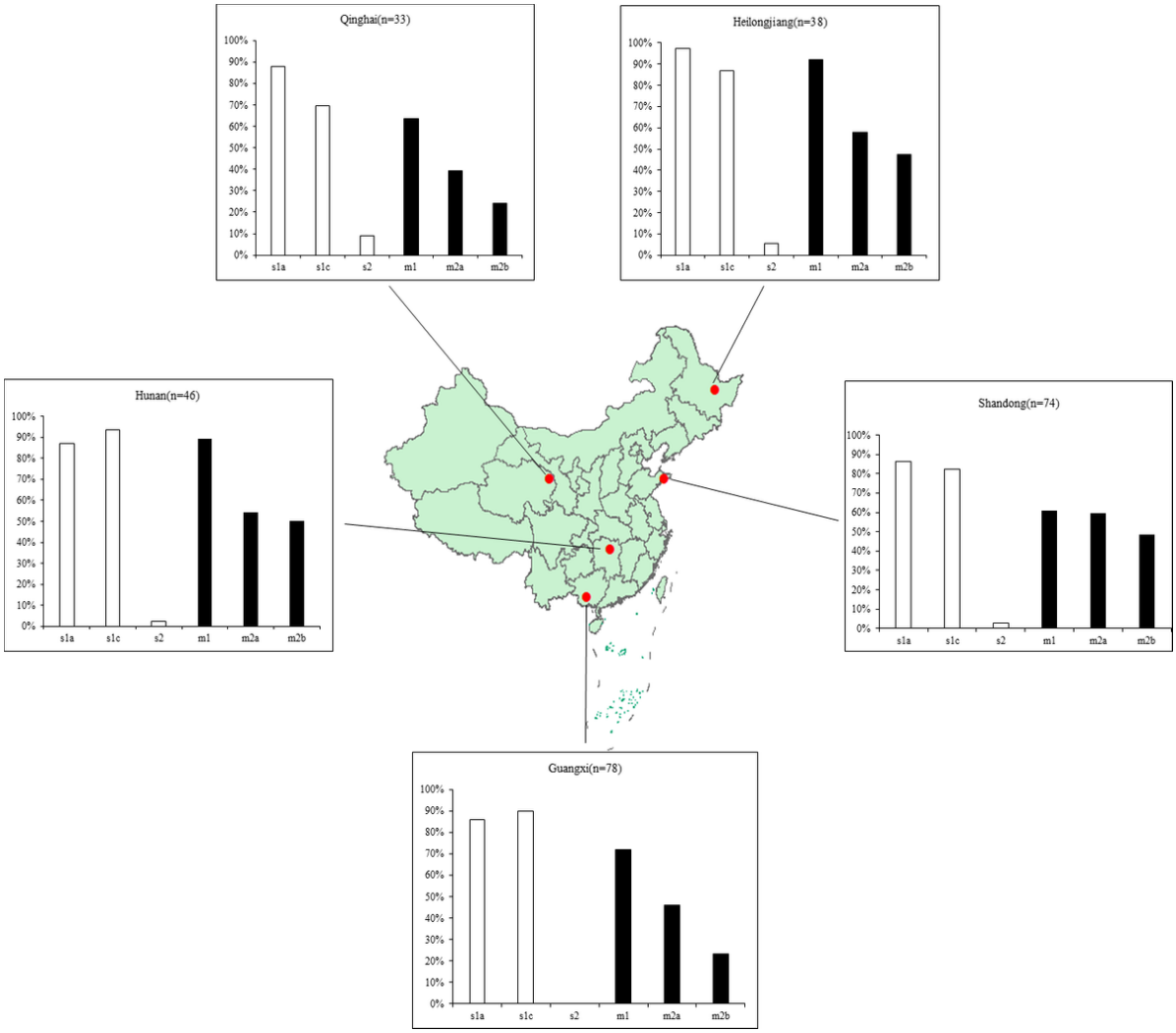

Figure 1

Geographic distribution of vacA s- and m-region genotypes of H. pylori strains from China. The prevalence of vacA subtypes (s1a, s1c, s2, m1, m2a and m2b) are given as a percentage of the total number of strains (shown in parentheses). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

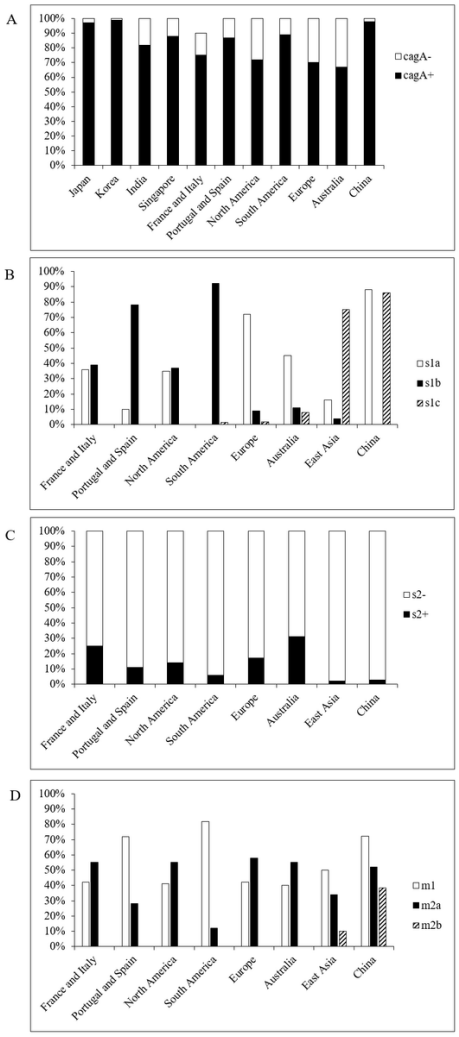

Figure 2

Distribution of vacA and cagA genotypes of H. pylori from different regions of the world. (A) cagA; (B) vacA s alleles; (C) vacA s2 allele and (D) vacA m alleles.