Original article

Helicobacter pylori serological biomarkers of gastric cancer risk in the MCC-Spain case-control Study

Authors:

Nerea Fernández de Larrea-Baz[1,2], Beatriz Pérez-Gómez[1-3], Angelika Michel[4], Beatriz Romero[5,6], Virginia Lope[1-3], Michael Pawlita[4], Tania Fernández-Villa[7,8], Victor Moreno[2,9-11], Vicente Martín[2,7,8], Martina Willhauck-Fleckenstein[4], Gonzalo López-Abente[1-3], Jesús Castilla[2,12], Guillermo Fernandez-Tardón[2,13], Trinidad Dierssen-Sotos[2,14], Miguel Santibáñez[14,15], Rosana Peiró[2,16], José Juan Jiménez-Moleón[2,17,18], Carmen Navarro[2,19,20], Gemma Castaño-Vinyals[2,21-23], Manolis Kogevinas[2,21-23], Marina Pollán[1-3], Silvia de Sanjosé[2,24], Rosa del Campo[5,6], Tim Waterboer[4] and Nuria Aragonés[1-3]

Affiliations:

[1] Environmental and Cancer Epidemiology Area, National Center of Epidemiology, Carlos III Health Institute, Avda. Monforte de Lemos 5, 28029, Madrid, Spain

[2] CIBER Epidemiología y Salud Pública (CIBERESP)-Public Health and Epidemiology Biomedical Research Networking Centre (CIBERESP), Avda. Monforte de Lemos, 3-5, Pabellón 11, Planta 0, 28029, Madrid, Spain

[3] Cancer Epidemiology Research Group. Oncology and Hematology Area. Instituto de Investigación Sanitaria Puerta de Hierro (IIS Puerta de Hierro)- Puerta de Hierro Health Research Institute, Madrid, Spain

[4] Divison of Molecular Diagnostics of Oncogenic Infections, Infection, Inflammation and Cancer Program, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120, Heidelberg, Germany

[5] Department of Microbiology, Hospital Universitario Ramón y Cajal and Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS)-Ramón y Cajal Health Research Institute (IRYCIS), Ctra. de Colmenar Viejo km. 9,100, 28034, Madrid, Spain

[6] Red Española de Investigación en Patología Infecciosa (REIPI)-Spanish Network for Research in Infectious Diseases (REIPI)

[7] The Research Group in Gene - Environment and Health Interactions, University of León, Spain

[8] Department of Biomedical Sciences, Area of Preventive Medicine and Public Health, University of León, Campus de Vegazana s/n, 24071, León, Spain

[9] Cancer Prevention and Control Program, Catalan Institute of Oncology, Avinguda de la Granvia de l'Hospitalet 199-203, 08908, Hospitalet de Llobregat, Barcelona, Spain

[10] Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Gran Via de les Corts Catalanes, 585, 08007, Barcelona, Spain

[11] Colorectal Cancer Group, Bellvitge Biomedical Research Institute (IDIBELL). Gran Via de l'Hospitalet, 199, 08908, Hospitalet de Llobregat, Barcelona, Spain

[12] Instituto de Salud Pública de Navarra, IdiSNA- Navarra Institute for Health Research, C/ Leyre, 15, 31003, Pamplona, Spain

[13] Oncology Institute, University of Oviedo, Fernando Bongera. Building "Santiago Gascón", 1st Floor. Campus of "El Cristo" B, 33006, Oviedo, Asturias, Spain

[14] University of Cantabria – IDIVAL-Instituto de Investigación Marqués de Valdecilla-Marqués de Valdecilla Research Institute (IDIVAL), C/ Cardenal Herrera Oria, s/n, 39011, Santander, Spain

[15] Centro de Investigación en Salud y Medio Ambiente (CYSMA), Universidad de Huelva-Research Center on Health and Environment (CYSMA), University of Huelva, Campus de El Carmen, 21007, Huelva, Spain.

[16] Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana FISABIO–Salud Pública -FISABIO-Public Health Research Center of the Valencian Health Department, Avda. de Catalunya, 21, 46020, Valencia, Spain

[17] Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA)-Granada Health Research Institute (ibs.GRANADA), 18012, Granada, Spain

[18] Department of Preventive Medicine and Public Health, University of Granada, Avda. de la Investigación, 11, Building A, 8th Floor, 18016, Granada, Spain

[19] Department of Epidemiology, Murcia Regional Health Council, Ronda de Levante, 11, 2ª planta, 30008, Murcia, Spain; IMIB-Arrixaca

[20] Department of Health and Social Sciences, University of Murcia, Avda. Teniente Flomesta, 5, 30003, Murcia, Spain

[21] ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Doctor Aiguader, 88, 08003, Barcelona, Spain

[22] IMIM (Hospital del Mar Medical Research Institute), Dr. Aiguader, 88, 08003, Barcelona, Spain

[23] Pompeu Fabra University (UPF), Plaça de la Mercè, 10-12, 08002, Barcelona, Spain

[24] Cancer Epidemiology Research Program, Catalan Institute of Oncology, Avinguda de la Granvia de l'Hospitalet 199-203, 08908, l'Hospitalet de Llobregat, Spain

Corresponding author: Nuria Aragonés. Instituto de Salud Carlos III - Centro Nacional de Epidemiología. Avda. Monforte de Lemos, 5 - Pabellón 12 – 1ª planta. 28029. Madrid. Spain. Phone: +34918222627. Fax: +34913877815. E-mail address: naragones@isciii.es

Running title: *H. pylori* gastric cancer risk seromarkers

Word count: 3903

Abstract

Background: *Helicobacter pylori* infection is one of the main risk factors for non-cardia gastric cancer. However, only a minority of infected persons develop the disease. This study aims at identifying bacterial risk markers in a multicase-control study in Spain (MCC-Spain).

Methods: Incident cases and population controls (age, sex and region frequency-matched) participated. Seroreactivities against 16 *H. pylori* proteins were determined using multiplex serology. Infection was defined as seropositivity against ≥four proteins. Relationship of serological results with non-cardia and cardia gastric cancer was assessed using multivariable mixed logistic regression models and principal components analysis.

Results: Infection prevalence was 95% among 202 non-cardia gastric cancer cases, 85% among 62 cardia cancer cases and 88% among 2071 controls (OR=1.9 (95% CI: 1.0-3.6) and OR=0.5 (95% CI: 0.3-1.1), respectively). Among infected subjects, seropositiviy to UreA, HP231, NapA and Cagδ was associated with lower non-cardia gastric cancer risk, while seropositivity to CagA and VacA was associated with higher risk. Seropositivity to CagA and seronegativity to Cagδ remained associated with non-cardia gastric cancer risk after additional adjustment by serostatus of significant proteins. We identified two antibody reactivity patterns, one related to higher and other to lower non-cardia gastric cancer risk.

Conclusions: In our population, people seropositive to *H. pylori* were characterized by two patterns of antibody reactivity against *H. pylori* proteins: one with high seroreactivity against several proteins simultaneously others than CagA and VacA, associated with a lower non-cardia gastric cancer risk, and another one with high seroreactivity against CagA and VacA, associated with an increased risk.

Mini-abstract: High antibody reactivity against several *H. pylori* proteins simultaneously was associated with a lower non-cardia gastric cancer risk, and high seroreactivity against CagA and VacA with an increased risk.

Keywords: Gastric neoplasm, *Helicobacter pylori* infection, multiplex serology, biomarkers, case-control studies.

Introduction

In spite of a marked decreasing trend in its incidence and mortality in the last decades, gastric cancer continues to be the third leading cause of cancer death in the world. Almost one million of new gastric cancer cases were estimated to occur in 2012, accounting for 7% of all incident cases of cancer worldwide, and 723,073 deaths were reported in the same year (9% of the total cancer deaths). Mortality rates are especially high in certain areas such as Eastern Asia, Central and Eastern Europe and South America [1,2]. In 2012, the estimated number of new gastric cancer cases in Spain was 7,810, and the number of registered deaths was 5,675; the corresponding age-standardized incidence and mortality rates (European Standard Population) were 16.4 per 100,000 population in men and 7.5 in women [3], and 11.3 per 100,000 population in men and 5.2 in women, respectively [4].

Geographic variations have also been observed within countries [5–7]. In Spain, gastric cancer mortality displays a singular geographical pattern, with higher mortality rates in Central and Northern regions [8,9]. This pattern, characterized by its persistence in time and its similarity in both sexes, has not changed substantially over the last decades [10,11]. Although differences in mortality among regions can be partially associated with different dietary habits or with territorially-related environmental exposures, the implication of *Helicobacter Pylori (H. pylori)* infection in the observed pattern is uncertain.

Chronic infection with *H. pylori* is the strongest established causal factor for the development of gastric cancer [12]. This association has been reported to be higher, and even limited to, cancer localized distal to the esophagogastric junction, the so called non-cardia gastric cancer, with risk estimates around 2.8 for non-cardia gastric cancer and 1.1 for cardia gastric cancer [13]. The prevalence of *H. pylori* infection is over 50% worldwide, with marked geographical variations [14]. Though higher prevalence is usually reported for developing countries [15–17], prevalences over 70%

have been reported in specific Western populations, including some Portuguese and Spanish regions [18–20]. Fortunately, only about 1% of the infected people develop a gastric cancer. Specific microbial strains and phenotypes have been linked to a higher risk of gastric cancer, and are now recognized virulence factors [21,22]. Also, different host and environmental factors have been shown to interact and modulate individual risk [23,24]. However, important issues such as how to identify infected people at higher risk of developing gastric cancer remain poorly understood.

Multiplex serology is a recently developed technique that allows the simultaneous and quantitative detection of antibodies directed against different antigens in a high throughput assay, which makes this technique appropriate for large epidemiological studies [25]. By means of multiplex serology we quantified antibody reactivities against a wide range of *H. pylori* proteins in a well characterized group of gastric cancer cases and healthy controls.

The aim of this study is to assess the relationship between serological reactivity to 16 *H. pylori* proteins and the development of gastric cancer, differentiating non-cardia and cardia gastric cancer, in order to identify serological markers of gastric cancer risk.

Methods

Study population

MCC-Spain study (www.mccspain.org) is a multicase-control study with population controls that was carried out in 12 Spanish provinces to study environmental and genetic risk factors for five types of tumours: colorectal, breast, oesophagogastric, prostate and chronic lymphocytic leukaemia [26]. Cases were recruited between 2008 and 2013 in 23 hospitals and controls were frequency-matched to cases by age, sex and region (province). In each province at least two types of cancer cases were recruited, and a common set of controls was selected for all their cases. To obtain an age and sex frequency-matched sample, an initial estimation of the expected distribution of this variables among

cases recruited in each province was made, taking into account data from the Spanish cancer registries. When the recruitment of cases finished, homogeneity in the age-sex distribution between cases and controls was checked, and additional controls were recruited when needed, to assure the availability, in each province, of at least one control of the same sex and age (+/- 5 years) for each case. Selection of controls was done through a random procedure from the General Practitioner's lists at primary healthcare centers of the hospitals' catchment areas. Potential controls were contacted by phone and, if they agree to participate, they were scheduled for a face-to-face interview. Cases and controls had to fulfill the following selection criteria: to have lived for at least 6 months in the study areas, to be 20-85 years old and to able to answer the epidemiological questionnaire. Gastric cancer cases were recruited in 15 hospitals from 9 Spanish provinces (Asturias, Barcelona, Cantabria, Granada, Huelva, León, Madrid, Navarra and Valencia). Eligible cases included histologically confirmed incident gastroesophageal cancer cases (International Classification of Diseases 10th Revision (ICD-10) codes C16, D00.2 or C15.5) diagnosed during the recruitment period and with no personal history of gastroesophageal cancer. We also recruited cases with tumour in the lower third of the oesophagus to be able to include all adenocarcinomas of the oesophagogastric junction. Cases were classified as cardia, non-cardia and others (tumors overlapping cardia and non-cardia or with non-specified gastric location). We identified 923 gastric cancer cases and 7734 controls potentially eligible. Response rates varied by centre and on average were 54.6% in cases and 53.0% in controls. Among the common set of controls, for this analysis we excluded those with personal history of gastroesophagic cancer, those from provinces that did not recruit gastric cancer cases and those younger than the youngest gastric cancer case in each province. Finally, 459 cases and 3 440 matched controls fulfilled selection criteria. All the participants answered an epidemiological questionnaire administered through personal interviews by trained personnel and were asked for donation of biological samples (blood, urine, hair and nails). The study protocol was approved by the Review Board of the participating institutions and all the included subjects gave written informed consent to participate.

Serological characterization

Blood samples were refrigerated locally until being processed and aliquoted in the first 48 hours. Then, they were stored at –80°C. Seroreactivities against 16 *H. pylori* proteins were determined using *H. pylori* multiplex serology (Online Resource 1). Multiplex serology is a glutathione S-transferase capture immunosorbent assay combined with fluorescent-bead technology, as described elsewhere [27]. This technique simultaneously quantifies antibodies directed against arrays of protein antigens. In brief, bacterially expressed recombinant glutathione S-transferase-*H. pylori* fusion proteins were used as antigens. The fusion proteins were loaded and affinity-purified directly on individual sets of spectrally distinct glutathione-casein-coupled fluorescence-labelled polystyrene beads (SeroMap, Luminex, Austin, TX). Bead sorts, each carrying a different antigen, were mixed and incubated with human sera at 1:100 dilutions. Antibodies bound to the beads via the bacterial antigens were stained by biotinylated anti-human-IgA, IgM, IgG (Dianova, Hamburg, Germany) and streptavidin-R-phycoerythrin. Beads were examined in a Luminex 200 analyser that quantifies the antibody bound to bacterial antigen via the median R-phycoerythrin fluorescence intensity of at least 100 beads of the same internal colour. Net (bead and glutathione S-transferase background subtracted) Median reporter Fluorescence Intensity values were calculated and negative values were set to +1.

For *H. pylori* proteins, serostatus cut-offs were calculated (mean of the median reporter fluorescence intensity + 3 standard deviations, excluding positive outliers) in 17 *H. pylori* negative sera previously classified for *H. pylori* status run within the same experiment. According to these cut-offs (Online Resource 1), each participant was classified as seropositive or seronegative to each protein. Following previously published criteria, *H. pylori* seropositivity was defined as positivity for at least 4 of the 15 *H. pylori* proteins (excluding HomB, a protein recently added to *H. pylori* multiplex serology), and subjects fulfilling this criteria were considered infected [25].

Statistical analyses

Association of proposed gastric cancer risk factors with both, case-control status and *H. pylori* serostatus in controls, was analysed by means of Chi-squared test.

To study the association between serological results and gastric cancer, odds ratios (OR) and 95% confidence intervals (95% CI) were estimated using multilevel logistic regression mixed models, with province as a random effect term and the following variables as fixed effects terms: age, sex, education, smoking status and gastric cancer family history. This analysis was performed in the overall sample to assess the effect of *H. pylori* seropositivity.

Then, to identify infected people at higher risk of developing gastric cancer, analyses assessing serostatus of individual proteins were performed in the infected subsample. Dose-response relationship was additionally explored using restricted cubic spline functions to assess potential departures from linearity. For this purpose seroreactivity values were log-transformed and subjects with values under percentile 1 or over percentile 99 were excluded. Subsequently, to study the independent effect of seropositiviy against proteins found to be associated with non-cardia cancer risk at a p-value<0.10, a multivariable logistic regression model was fitted simultaneously including serostatus of all of them in addition to potential confounding factors.

Finally, to account for a possibly high collinearity among antibody reactivities against analysed proteins, and in order to identify antibody response patterns, a principal components analysis was carried out. The principal components analysis was performed using the control population, and reactivity values of each tested antibody were standardized to a mean of 0 and a standard deviation of 1. Selection of relevant components was done taking into account obtained eigenvalues and clinical interpretability of the components arising from the analysis. To assess the relationship of the selected components with non-cardia gastric cancer risk, the score for each component was

calculated for each case and control. Then scores were categorized, according to quartiles of their distribution among controls, and included in a multivariable logistic regression mixed model, to estimate the association between gastric cancer risk and the obtained components. These models were adjusted by the same covariates as the models described above (with province as a random effect term and age, sex, education, smoking status and gastric cancer family history as fixed effects terms).

Results

Among participants fulfilling selection criteria, *H. pylori* serostatus could be determined in 281 cases (202 non-cardia gastric cancer, 62 cardia gastric cancer, 9 located in the esophageal lower third and 8 with overlapping or not classifiable tumor location) and 2 071 matched controls, and were therefore included in the present analyses. The other participants had no multiplex serology results (no consent for blood sample collection or sample not processed) or their serological results were considered as non-valid (glutathione S-transferase>300 median reporter fluorescence intensity) (two cases).

Table 1 shows the characteristics of studied cases and controls. Compared to controls, cases were more likely to be men, aged 75 or above, no Caucasian, current smoker, have lower education, lower socioeconomic level at birth, have first degree relatives with gastric cancer and self-reported gastritis or heartburn. Among controls, *H. pylori* infection was associated with sex, age, education, socioeconomic level at birth and body mass index (Online Resource 2).

According to the established definition of seropositivity to 4 or more *H. pylori* proteins, 88% (95% CI: 87-90) of controls, 95% (95% CI: 91-98) of non-cardia gastric cancer cases and 85% (95% CI: 76-95) of cardia gastric cancer cases were deemed to be infected. Infection was associated with a 90% increased risk of non-cardia gastric cancer (*p*-value=0.047) and with a non-statistically significant

decrease of cardia gastric cancer risk (OR=0.54; 95% CI: 0.25-1.14). No association was observed between the number of proteins against which a seropositive result was obtained and non-cardia gastric cancer risk; meanwhile, for cardia cancer, a 9% lower risk was estimated for each seropositive protein (Table 2).

In controls, seropositivity to individual *H. pylori* proteins ranged from 31% (Cad) to 83% (GroEL) (Online Resource 3). Only 1.5% of controls were seronegative to all of the analysed proteins. The median number of proteins against which controls were seropositive was 9 (Interquartile range: 7-11) in infected controls (those seropositive to 4 or more proteins) and 2 (Interquartile range: 1-3) in non-infected controls. Among cases, seropositivity ranged from 29% for HomB to 89% for GroEL, with a median of 9 proteins seropositive (Interquartile range: 7-11) in infected cases and 2 (Interquartile range: 2-3) in non-infected cases.

Correlation among seroreactivity, measured as median reporter fluorescence intensity, of the 16 proteins was mild for the majority of them (Online Resource 4). When measured as dichotomized variables, concordance of serostatus between proteins was low (*kappa* coefficient <0.4 for all pairs) among infected participants.

Among infected subjects, seropositivity to eight of the analysed *H. pylori* proteins showed a statistically significant or almost significant association with non-cardia gastric cancer (Figure 1): six were associated with a lower risk (UreA, HP231, NapA, Cagδ, Catalase and HcpC) and two (CagA and VacA) with a higher risk. With respect to cardia gastric cancer, HP305 and Cagδ appeared associated with a lower risk (in the limit of the statistical significance). Dose-response association with non-cardia cancer for these nine proteins can be visualized in Figure 2, again including only participants classified as *H. pylori* seropositive. A decreasing OR from the seropositivity cut-off value upward was estimated for UreA, NapA, Cagδ, Catalase and HcpC, while the opposite effect was estimated for

CagA and, to a lesser extent, for VacA. For HP305 and HP231, higher risks were estimated for median reporter fluorescence intensity levels under the cut-off. For UreA, NapA, HP305, CagA, HcpC and GroEL (the latter not included in Figure 2) relationship with non-cardia cancer departed from linearity ($p<0.05$).

Results were qualitatively similar between men and women, and between those under and over 65 years old. No differences in the direction of the effects were present and, where differences in the magnitude of the ORs were observed, confidence intervals overlapped (data not shown).

When mutually adjusting by serostatus against those proteins associated at a p-value<0.10 with non-cardia gastric cancer risk in the individual analyses, infected subjects seropositive to CagA presented a threefold increased risk of non-cardia gastric cancer (OR=3.65; 95% CI: 2.44-5.46), meanwhile seropositivity to Cagδ was associated with a 34% lower risk (OR=0.66; 95% CI: 0.47-0.92). The association of serostatus against VacA, UreA, HP231, NapA, Catalase and HcpC with NCGC was not longer statistically significant.

Based on the eigenvalues and interpretability of the patterns, the principal components analysis revealed two serological reactivity patterns. The two factors explained the 25% and 11% of total variation in the serological profiles. The first component was characterized by moderate weights in several of the analysed antibodies, mainly those against GroEL, NapA, UreA, HyuA, Catalase and HP231 (loadings>0.5). For the second component, main contributors were CagA and VacA antibody reactivities (loadings >0.5). Figure 3 depicts the correlation among each component and seroreactivity against the 16 analysed proteins. Table 3 shows the associations of each component with non-cardia gastric cancer risk among *H. pylori* infected subjets. High scores in the first component were associated with a 60% lower risk of non-cardia gastric cancer (fourth vs. first

quartile). On the other hand, high scores in the second component were associated with a threefold increased risk.

Discussion

The high prevalence of *H. pylori* infection and the relatively low incidence of non-cardia gastric cancer even among those infected, render the identification of markers of increased or decreased cancer risk a crucial issue to decide how to manage the infection, both at individual and at population level. In this study we have identified two patterns of antibody reactivity against 16 *H. pylori* proteins related to different risk of non-cardia gastric cancer: a pattern characterized by the concomitant presence of high antibody reactivities against a range of different proteins, which was associated with a lower risk, and another one characterized by high antibody reactivities against the main recognized *H. pylori* virulence factors, CagA and VacA, which was associated with a higher risk. Interestingly, in our control population these two patterns show some parallelism with the type of *H. pylori* strain causing the infection [28]. Those infected by a type I *H. pylori* strain (defined by a seroreactivity against CagA higher than 9000 median reporter fluorescence intensity, following published criteria for multiplex serology [25]) presented lower scores in the first component and higher scores in the second one than those infected by a type II strain (data not shown). Therefore, the first pattern could represent an *H. pylori* type II infection ("non-virulent pattern") and the second one a type I infection ("virulent pattern").

The analyses of serostatus of proteins individually, showed that, after adjusting by main gastric cancer risk factors, infected people seropositive to CagA or VacA had and increased risk of non-cardia gastric cancer compared to those infected but seronegative to these proteins. On the other hand, infected people seropositive to Cagδ, UreA, HP231 or NapA had a lower risk. When including in the same model serostatus against these proteins, only CagA and Cagδ remained associated with non-cardia gastric cancer in a statistically significant manner. The dose-response analysis showed a

statistically significant trend with intensity of antibody reactivity for both of them, giving some support for a real association. CagA is a well-recognized *H. pylori* virulence factor that has been related to atrophic gastritis and non-cardia gastric cancer. On the other hand, the exact function of Cagδ is still not well understood [29]. This protein is a component of the T4SS (type 4 secretion system) of *H. pylori*. The major role of this T4SS is thought to be the translocation of the CagA effector protein into the host cell, but CagA-independent pathogenicity pathways, such as the induction of IL-8 production from the host, have also been suggested [30]. There is not consistency among the results from different studies in the possible role of serostatus to Cagδ as a marker of lower non-cardia gastric cancer risk. Therefore, this finding should be taken cautiously until more evidence becomes available [31–35].

Recent studies using *H. pylori* multiplex serology have also found statistically significant inverse relationships between seropositivity against some *H. pylori* proteins, such as CagM [35], GroEL and NapA [33], and gastric cancer risk, but others have not [31,32,34]. In a case-control study carried out in Sweden [32], a non-statistically significant lower risk of non-cardia gastric cancer was associated with seropositivity to the same four proteins identified in our study (UreA, HP231, NapA and Cagδ) and to nine additional proteins (HP305, HpaA, CagM, VacA, HcpC, Cad, Omp, HomB and BabA), but only in the subgroup with serologic evidence of atrophic gastritis. In contrast, other studies have found higher gastric cancer risk associated with seropositivity to GroEL [31,32,34,35], HcpC [31,32,35], HP305 [31,34,35], Catalase [31,32], HyuA [31,32,34], Omp [32,34,35] or HpaA [32,34].

Some methodological differences among studies could help to explain the differences with our findings. First, statistical analyses performed to identify serological virulence markers in our study have been restricted to subjects classified as infected (positive to 4 or more *H. pylori* proteins), in contrast to other studies, where the overall sample (infected and non-infected) was considered. However, we repeated our analyses in the overall sample, and not noteworthy differences appeared

between the two approaches (data not shown). Second, differences in patient selection criteria, such as the exclusion of patients with atrophic gastritis [31], may lead to increased risk estimations. Third, the prevalence of *H. pylori* infection and of seropositivity to each protein differ among studies. High prevalences of exposure among controls, like in our sample, tend to reduce the magnitude of ORs in case-control studies [36,37]. Nevertheless, real differences among countries in the predominant bacterial strains and in host genetics and environmental factors cannot be ruled out as an explanation for different results among studies. This would be in accordance with the importance generally attributed to geographical variations in the epidemiology of both, *H. pylori* infection and gastric cancer [38–40].

Limitations of our study include the case-control design, which limits the possibility to demonstrate a cause to effect direction of the associations. Therefore, a possible impact on our results of changes in *H. pylori* serostatus secondary to the disease or its treatment cannot be completely ruled out. To explore the possibility of reverse causation in our data, we performed stratified analyses by cancer stage and by treatment status at the moment of blood sample collection. No remarkable differences appeared in the estimated ORs between tumoral stage subgroups. Meanwhile, we observed some indications of a higher OR for the association between *H. pylori* infection and non-cardia gastric cancer risk among those not having received cancer therapy, compared to those that had received some treatment (data not shown).

With regard to the criteria used to define infection (seropositivity for ≥4 proteins by multiplex serology), it should be noted that it has not been validated in our sample. However, this definition has been contrasted with standard ELISA and Western blot assays in other populations, showing acceptable validity [25]. Another limitation is the lack of information about the presence of atrophic gastritis, a factor that can exert a confounding effect on the association between serological status and non-cardia gastric cancer risk [31–33,41–43]. Nonetheless, the high sensitivity of the multiplex

serology could minimize the impact that factors related to the persistence of the serological response, such as those mentioned, may have on the assessment of associations between infection diagnosed by serological techniques, and disease. Lastly, both the relatively low sample size and the mild-moderate correlation among antibody reactivities against different proteins, reduce statistical power, especially to perform subgroup analyses.

In spite of these limitations, our study included an acceptable number of incident gastric cancer patients including cardia as well as non-cardia gastric cancer, and a broad sample of population controls, and was conducted following a meticulous methodology which allowed gathering precise information for the most important recognized risk factors for this disease. Besides, estimating risks for the infected subjects instead of the whole sample, represents a novel approach that enriches previous knowledge, once the carcinogenic role of *H. pylori* in non-cardia gastric cancer has been generally accepted. This approach, applied to an *H. pylori* infection defined as seropositivity to four or more bacterial proteins, also contributes to reduce the possible classification error derived from eventual false positive results in serostatus against a single protein. In addition, our results complement data derived from other regions, which is an important issue for a disease characterized by high geographical variations. In this respect, due to geographical and historical characteristics, the Spanish population has coexisted with a diversity of populations from other origins, which may have influenced its genetic background, and its response to *H. pylori* colonization. The high prevalence of infection identified in this study does not have a reflection in the gastric cancer incidence or mortality rates in Spain, which are similar to or lower than the European average [2]. This could suggest that there are differences with respect to other populations either in the characteristics of the circulating *H. pylori* strains, in the host response to the infection or in environmental factors. Investigating whether the antibody reactivity patterns identified in this study are also present in other populations, and whether they are also associated with the risk of non-cardia gastric cancer would be of value to contribute to the understanding of geographical differences.

In summary, according to our results people with *H. pylori* infection that have high reactivities against CagA and VacA would have an increased risk of non-cardia gastric cancer. Additionally, we have identified a group that, in spite of have being infected by the bacteria, would experience a low risk of non-cardia gastric cancer. This group would be characterized by high concomitant seroreactivities against many *H. pylori* proteins, particularly against UreA, HP231, NapA, HyuA and Catalase.

Compliance with ethical standards

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (Comité de Ética de la Investigación y de Bienestar Animal del Instituto de Salud Carlos III) and with the Helsinki Declaration of 1964 and later versions. Informed consent was obtained from all participants for being included in the study.

Acknowledgements

The authors thank the participants and staff of the collaborating hospitals and primary care centers for their valuable contributions as well as the other investigators of the Project MCC-Spain for their contributions to this study. They are also grateful to Adela Castelló and Roberto Pastor for their technical support in the principal component analysis. Also, the authors acknowledge the continuous support received from the Public Health and Epidemiology Biomedical Research Networking Centre (CIBERESP).

Funding: This study was partially funded by the “Acción Transversal del Cáncer", approved on the Spanish Ministry Council on the 11th October 2007, by the Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), by the Instituto de Salud Carlos III grants, co-funded by FEDER funds -a way to build Europe- (grants PI08/1770 to M. Kogevinas, PI09/0773 to J. Llorca, PI09/1286 to V. Martín, PI09/1903 to R. Peiró, PI09/2078 to F.J. Caballero, PI09/1662 to J.J. Jiménez-

Moleón, PI11/01403 to N. Aragonés, PI14/01219 and 2014 SGR 756 to S. de Sanjose), by the Fundación Marqués de Valdecilla (grant API 10/09 to J. Llorca), by Catalan Government DURSI (grant 2014SGR647 to V. Moreno), by the Junta de Castilla y León (grant LE22A10-2 to V. Martín), by the Consejería de Salud of the Junta de Andalucía (grant 2009-S0143 to J. Alguacil), and by the Conselleria de Sanitat of the Generalitat Valenciana (grant AP061/10 to R. Peiró). The funders had no role in the study design and data analysis.

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

References

1. Forman D, Bray F, Brewster DH, Gombe Mbalawa C, Kohler B, Piñeros M, Steliarova-Foucher E, Swaminathan R and Ferlay J., eds. Cancer Incidence in Five Continents, Vol. X (electronic version). [Internet]. Lyon, France: International Agency for Research on Cancer (IARC); 2013. Report No.: Vol X. Available from: http://ci5.iarc.fr

2. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013 [cited 2015 Oct 9]. Available from: http://globocan.iarc.fr

3. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur. J. Cancer Oxf. Engl. 1990. 2013;49:1374–403.

4. Cancer and Environmental Epidemiology Unit. Interactive Epidemiological Information System (Ariadna) [Internet]. Madrid: Carlos III Institute of Health; 2015 [cited 2015 Sep 14]. Available from: http://ariadna.cne.isciii.es/

5. Verma Y, Pradhan PK, Gurung N, Sapkota SD, Giri P, Sundas P, et al. Population-based cancer incidence in Sikkim, India: report on ethnic variation. Br. J. Cancer. 2012;106:962–5.

6. Ang TL, Fock KM. Clinical epidemiology of gastric cancer. Singapore Med. J. 2014;55:621–8.

7. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric Cancer: Descriptive Epidemiology, Risk Factors, Screening, and Prevention. Cancer Epidemiol. Biomarkers Prev. 2014;23:700–13.

8. García-Esquinas E, Pérez-Gómez B, Pollán M, Boldo E, Fernández-Navarro P, Lope V, et al. Gastric cancer mortality trends in Spain, 1976-2005, differences by autonomous region and sex. BMC Cancer. 2009;9:346.

9. Aragonés N, Pérez-Gómez B, Pollán M, Ramis R, Vidal E, Lope V, et al. The striking geographical pattern of gastric cancer mortality in Spain: environmental hypotheses revisited. BMC Cancer. 2009;9:316.

10. Aragonés N, Goicoa T, Pollán M, Militino AF, Pérez-Gómez B, López-Abente G, et al. Spatio-temporal trends in gastric cancer mortality in Spain: 1975-2008. Cancer Epidemiol. 2013;37:360–9.

11. López-Abente G, Aragonés N, Pérez-Gómez B, Pollán M, García-Pérez J, Ramis R, et al. Time trends in municipal distribution patterns of cancer mortality in Spain. BMC Cancer. 2014;14:535.

12. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Schistosomes, liver flukes and Helicobacter pylori. Lyon, France: International Agency for Research on Cancer (IARC); 1994 p. 1–241. Report No.: 61.

13. Cavaleiro-Pinto M, Peleteiro B, Lunet N, Barros H. Helicobacter pylori infection and gastric cardia cancer: systematic review and meta-analysis. Cancer Causes Control CCC. 2011;22:375–87.

14. IARC Working Group on the Evaluation of, Carcinogenic Risks to Humans. A review of human carcinogens. Part B: Biological agents. Lyon, France: International Agency for Research on Cancer (IARC); 2009. Report No.: 100B.

15. Brown LM. Helicobacter pylori: epidemiology and routes of transmission. Epidemiol. Rev. 2000;22:283–97.

16. Eusebi LH, Zagari RM, Bazzoli F. Epidemiology of *Helicobacter pylori* Infection. Helicobacter. 2014;19:1–5.

17. Peleteiro B, Bastos A, Ferro A, Lunet N. Prevalence of Helicobacter pylori Infection Worldwide: A Systematic Review of Studies with National Coverage. Dig. Dis. Sci. 2014;59:1698–709.

18. Bastos J, Peleteiro B, Barros R, Alves L, Severo M, de Fátima Pina M, et al. Sociodemographic determinants of prevalence and incidence of Helicobacter pylori infection in Portuguese adults. Helicobacter. 2013;18:413–22.

19. Sánchez Ceballos F, Taxonera Samsó C, García Alonso C, Alba López C, Sainz de Los Terreros Soler L, Díaz-Rubio M. [Prevalence of Helicobacter pylori infection in the healthy population of Madrid (Spain)]. Rev. Esp. Enfermedades Dig. Organo Of. Soc. Esp. Patol. Dig. 2007;99:497–501.

20. Macenlle García R, Gayoso Diz P, Sueiro Benavides RA, Fernández Seara J. Prevalence of Helicobacter pylori infection in the general adult population of the province of Ourense. Rev. Esp. Enfermedades Dig. Organo Of. Soc. Esp. Patol. Dig. 2006;98:241–8.

21. Kusters JG, van Vliet AHM, Kuipers EJ. Pathogenesis of Helicobacter pylori Infection. Clin. Microbiol. Rev. 2006;19:449–90.

22. Sgouras DN, Trang TTH, Yamaoka Y. Pathogenesis of *Helicobacter pylori* Infection. Helicobacter. 2015;20:8–16.

23. Wroblewski LE, Peek RM, Wilson KT. Helicobacter pylori and Gastric Cancer: Factors That Modulate Disease Risk. Clin. Microbiol. Rev. 2010;23:713–39.

24. Zhang W, Lu H, Graham DY. An Update on Helicobacter pylori as the Cause of Gastric Cancer. Gastrointest. Tumors. 2014;1:155–65.

25. Michel A, Waterboer T, Kist M, Pawlita M. Helicobacter pylori multiplex serology. Helicobacter. 2009;14:525–35.

26. Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, Martín V, Llorca J, Moreno V, et al. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. Gac. Sanit. SESPAS. 2015;29:308–15.

27. Waterboer T, Sehr P, Michael KM, Franceschi S, Nieland JD, Joos TO, et al. Multiplex human papillomavirus serology based on in situ-purified glutathione s-transferase fusion proteins. Clin. Chem. 2005;51:1845–53.

28. Xiang Z, Censini S, Bayeli PF, Telford JL, Figura N, Rappuoli R, et al. Analysis of expression of CagA and VacA virulence factors in 43 strains of Helicobacter pylori reveals that clinical isolates can be divided into two major types and that CagA is not necessary for expression of the vacuolating cytotoxin. Infect. Immun. 1995;63:94–8.

29. Olbermann P, Josenhans C, Moodley Y, Uhr M, Stamer C, Vauterin M, et al. A global overview of the genetic and functional diversity in the Helicobacter pylori cag pathogenicity island. PLoS Genet. 2010;6:e1001069.

30. Zanotti G, Cendron L. Structural and functional aspects of the Helicobacter pylori secretome. World J. Gastroenterol. WJG. 2014;20:1402–23.

31. Gao L, Michel A, Weck MN, Arndt V, Pawlita M, Brenner H. Helicobacter pylori infection and gastric cancer risk: evaluation of 15 H. pylori proteins determined by novel multiplex serology. Cancer Res. 2009;69:6164–70.

32. Song H, Michel A, Nyrén O, Ekström A-M, Pawlita M, Ye W. A CagA-independent cluster of antigens related to the risk of noncardia gastric cancer: associations between Helicobacter pylori antibodies and gastric adenocarcinoma explored by multiplex serology. Int. J. Cancer J. Int. Cancer. 2014;134:2942–50.

33. Shakeri R, Malekzadeh R, Nasrollahzadeh D, Pawlita M, Murphy G, Islami F, et al. Multiplex H. pylori serology and risk of gastric cardia and non-cardia adenocarcinomas. Cancer Res. [Internet]. 2015 [cited 2015 Oct 22]; Available from: http://cancerres.aacrjournals.org/cgi/doi/10.1158/0008-5472.CAN-15-0556

34. Epplein M, Zheng W, Xiang Y-B, Peek RM, Li H, Correa P, et al. Prospective study of Helicobacter pylori biomarkers for gastric cancer risk among Chinese men. Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol. 2012;21:2185–92.

35. Murphy G, Freedman ND, Michel A, Fan J-H, Taylor PR, Pawlita M, et al. Prospective study of Helicobacter pylori antigens and gastric noncardia cancer risk in the nutrition intervention trial cohort. Int. J. Cancer J. Int. Cancer. 2015;137:1938–46.

36. Huang JQ, Sridhar S, Chen Y, Hunt RH. Meta-analysis of the relationship between Helicobacter pylori seropositivity and gastric cancer. Gastroenterology. 1998;114:1169–79.

37. Vrijheid M, Deltour I, Krewski D, Sanchez M, Cardis E. The effects of recall errors and of selection bias in epidemiologic studies of mobile phone use and cancer risk. J. Expo. Sci. Environ. Epidemiol. 2006;16:371–84.

38. Yamaoka Y, Kato M, Asaka M. Geographic differences in gastric cancer incidence can be explained by differences between Helicobacter pylori strains. Intern. Med. Tokyo Jpn. 2008;47:1077–83.

39. Kang J, Jones KR, Jang S, Olsen CH, Yoo Y-J, Merrell DS, et al. The geographic origin of Helicobacter pylori influences the association of the homB gene with gastric cancer. J. Clin. Microbiol. 2012;50:1082–5.

40. Sugimoto M, Yamaoka Y. The association of vacA genotype and Helicobacter pylori-related disease in Latin American and African populations. Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. 2009;15:835–42.

41. Ekström AM, Held M, Hansson LE, Engstrand L, Nyrén O. Helicobacter pylori in gastric cancer established by CagA immunoblot as a marker of past infection. Gastroenterology. 2001;121:784–91.

42. Wang C, Yuan Y, Hunt RH. The association between Helicobacter pylori infection and early gastric cancer: a meta-analysis. Am. J. Gastroenterol. 2007;102:1789–98.

43. Camorlinga-Ponce M, Flores-Luna L, Lazcano-Ponce E, Herrero R, Bernal-Sahagún F, Abdo-Francis JM, et al. Age and severity of mucosal lesions influence the performance of serologic markers in Helicobacter pylori-associated gastroduodenal pathologies. Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol. 2008;17:2498–504.

Tables

Table 1. Characteristics of gastric cancer cases and controls

Variable	CONTROLS n(%) (N=2071)	GC CASES n(%)[a]		p-value cardia	p-value non-cardia
		Cardia (N=62)	Non-cardia (N=202)		
Sex					
Male	1183 (57%)	54 (87%)	126 (62%)	<0.001	0.149
Female	888 (43%)	8 (13%)	76 (38%)		
Age (years)					
<55	391 (19%)	13 (21%)	36 (18%)	0.023	0.001
55-64	492 (24%)	17 (27%)	30 (15%)		
65-74	731 (35%)	11 (18%)	70 (35%)		
>=75	457 (22%)	21 (34%)	66 (33%)		
Race					
White/Caucasian	2036 (98%)	61 (98%)	193 (96%)	0.991	0.004
Other	33 (2%)	1 (2%)	9 (4%)		
Education					
No/incomplete primary school	427 (21%)	18 (29%)	61 (30%)	0.335	0.003
Primary school	784 (38%)	24 (39%)	79 (39%)		
Secondary school	528 (25%)	13 (21%)	42 (21%)		
University degree	332 (16%)	7 (11%)	20 (10%)		

Variable	CONTROLS n(%) (N=2071)	GC CASES n(%)[a]		p-value cardia	p-value non-cardia
		Cardia (N=62)	Non-cardia (N=202)		
Socioeconomic level at birth					
Low	902 (47%)	33 (53%)	110 (55%)	0.451	0.133
Intermediate	943 (49%)	26 (42%)	85 (42%)		
High	62 (3%)	3 (5%)	6 (3%)		
BMI (Kg/m2)					
<25	599 (34%)	16 (27%)	66 (35%)	0.410	0.894
25-29.9	816 (46%)	27 (46%)	83 (44%)		
≥30	372 (21%)	16 (27%)	40 (21%)		
Smoking status					
Never smoker	907 (44%)	18 (29%)	88 (44%)	0.004	0.595
Former smoker	734 (36%)	21 (34%)	67 (33%)		
Current smoker	421 (20%)	23 (37%)	47 (23%)		
GC family history					
No GC family history	1817 (88%)	50 (82%)	155 (78%)	0.031	<0.001
Only 2nd degree relatives	111 (5%)	2 (3%)	13 (7%)		
≥1 first degree relative	132 (6%)	9 (15%)	32 (16%)		
History of gastritis					
No	1844 (96%)	56 (92%)	187 (93%)	0.063	0.020
Yes	69 (4%)	5 (8%)	14 (7%)		

Variable	CONTROLS n(%) (N=2071)	GC CASES n(%)[a]		p-value cardia	p-value non-cardia
		Cardia (N=62)	Non-cardia (N=202)		
History of heartburn					
No	1232 (65%)	35 (57%)	103 (52%)	0.248	<0.001
Yes	676 (35%)	26 (43%)	97 (49%)		
Histological type					
AC	-	60 (97%)	186 (92%)		
Other	-	2 (3%)	16 (8%)		
Laurén classification[b]					
Intestinal	-	21 (35%)	81 (44%)		
Diffuse	-	7 (12%)	51 (27%)		
Mixed	-	32 (53%)	9 (5%)		
Not available	-	0 (0%)	45 (24%)		
WHO classification[b]					
Papillary/Tubular	-	19 (32%)	76 (41%)		
Mucinous	-	2 (3%)	5 (3%)		
Poorly cohesive	-	8 (13%)	54 (29%)		
Mixed	-	1 (2%)	10 (5%)		
Not available	-	30 (50%)	41 (22%)		
Tumor stage					
Localized (TNM stages 0-II)	-	9 (17%)	69 (39%)		
Advanced (TNM stages III-IV)	-	45 (83%)	106 (61%)		

Variable	CONTROLS n(%) (N=2071)	GC CASES n(%)[a]		p-value cardia	p-value non-cardia
		Cardia (N=62)	Non-cardia (N=202)		
Blood collection moment					
Prior/concomitant to treatment	-	28 (50%)	86 (45%)		
First 2 months after treatment	-	14 (25%)	65 (34%)		
>2 months after treatment	-	14 (25%)	41 (21%)		
Initial treatment					
Surgery	-	28 (50%)	164 (82%)		
Chemotherapy	-	25 (45%)	36 (18%)		
Chemo-radiotherapy	-	3 (5%)	0 (0%)		

[a] Seventeen cases excluded from site-specific analyses (8 not classifiable as cardia or non-cardia gastric cancer and 9 located in the esophageal lower third). [b] Only applicable to adenocarcinomas.
Sum of cases in some variables does not coincide with overall number of cases because of missing information. Cases and controls are not balanced by age and sex in spite of the frequency-matching, because matching was made based on the distribution of all the types of cancer cases recruited in each province, and gastric cancer cases were in general older and more frequently men than cases of other tumors. AC: Adenocarcinoma; BMI: Body mass index; GC: Gastric cancer; WHO: World Health Organization.

Table 2. Association between gastric cancer and *H. pylori* infection according to multiplex serology results

	Controls	Cases										
		All		Non-cardia		Cardia	All			Non-cardia			Cardia		
H. pylori serostatus	N (%)	N (%)		N (%)		N (%)	OR	(95% CI)[a]	p-value	OR	(95% CI)[a]	p-value	OR	(95% CI)[a]	p-value
H.pylori+	1814 (88%)	257 (92%)		189 (95%)		52 (85%)	1.31	(0.82-2.11)	0.263	1.90	(1.01-3.56)	0.047	0.54	(0.25-1.14)	0.105
Number of proteins+ [b]	8.2 [3.5]	8.3 [3.1]		8.5 [3.0]		7.6 [3.5]	0.99	(0.95-1.03)	0.705	1.02	(0.97-1.06)	0.485	0.91	(0.84-0.98)	0.018
Number of proteins+ (grouped):															
<4	238 (12%)	21 (8%)		11 (6%)		9 (15%)	1.00			1.00			1.00		
4-6	394 (19%)	60 (22%)		44 (22%)		14 (23%)	1.47	(0.86-2.50)	0.157	2.05	(1.03-4.08)	0.040	0.75	(0.31-1.81)	0.528
>6	1420 (69%)	197 (71%)		145 (73%)		38 (62%)	1.27	(0.78-2.05)	0.336	1.85	(0.98-3.50)	0.058	0.48	(0.22-1.04)	0.063
Trend							1.04	(0.85-1.27)	0.733	1.17	(0.92-1.49)	0.207	0.68	(0.48-0.97)	0.036

[a] Adjusted by age, sex, education, family history of gastric cancer and smoking status; province included as a random-effect term. [b] Mean [standard deviation] of the number of proteins against which antibodies were present. In this case, ORs represent the risk for each additional protein against which antibody reactivity was positive.

Table 3. Risk of non-cardia gastric cancer among *H. pylori* infected subjects, according to the scores in the components arising from principal component analysis.

		Controls N(%)	Non-cardia GC N(%)	OR (95% CI)	P value
Component 1	Q1	457 (25%)	74 (39%)	Reference	
	Q2	459 (25%)	51 (27%)	0.61 (0.41–0.90)	0.014
	Q3	458 (25%)	35 (18%)	0.40 (0.25–0.62)	<0.001
	Q4	457 (25%)	31 (16%)	0.38 (0.24–0.61)	<0.001
Component 2	Q1	458 (25%)	23 (12%)	Reference	
	Q2	457 (25%)	36 (19%)	1.48 (0.86–2.53)	0.154
	Q3	459 (25%)	65 (34%)	2.55 (1.55–4.20)	<0.001
	Q4	457 (25%)	67 (35%)	3.07 (1.88–5.00)	<0.001

ORs from logistic regression analysis adjusted by age, sex, education, family history of gastric cancer, smoking status and province, and mutually adjusted by principal component scores. GC: Gastric cancer. Q1: Lowest quartile; Q2: Second quartile; Q3: Third quartile; Q4: Highest quartile.

Figure legends

Figure 1

Title: Association between gastric cancer and positivity for antibody reactivity to *H. pylori* proteins among infected subjects.

Legend: OR adjusted by age, sex, education, family history of gastric cancer and smoking status; province included as a random-effect term. OR (95% CI): Odds ratio (95% confidence interval).

Figure 2

Title: Dose-response association between antibody reactivity against selected *H. pylori* proteins and non-cardia gastric cancer risk among infected subjects

Legend: Median reporter fluorescence intensity values were log transformed and cubic restricted spline functions were estimated using 3 knots (at 10, 50 and 90^{th} percentiles). Ln(OR) were estimated by multilevel logistic regression mixed analysis adjusted for age, sex, education, smoking status and family history of gastric cancer; province was included as a random-effect term. Cut-off value for each *H. pylori* protein seropositivity was taken as the reference value. Departure from linearity was assessed using Wald test. Vertical grey lines represent the seroreactitivy level used as reference value for each protein. Dashed lines show the 95% confidence interval for the estimated ln(OR). Bars represent the distribution of participants according to their median reporter fluorescence intensity levels, quantified by multiplex serology. MFI: Median reporter Fluorescence Intensity.

Figure 3

Title: Principal component analysis results

Legend: Correlations between antibody reactivity (MFI) against each *H. pylori* protein and the two selected components. MFI: Median reporter Fluorescence Intensity.

Supplementary material captions

Online Resource 1

Title: Analysed *H. pylori* proteins and cut-offs used for serostatus classification

Caption: All the proteins were expressed from *H. pylori* strain 26695, except GroEL, from strain G27 and HomB from strain J99. Criteria for choosing proteins for the assay were: known surface exposure and immunogenicity in two-dimensional immunoblot analyses (UreA, HP231, NapA, HpaA, CagA, Catalase and VacA), serologic association with gastric cancer (GroEL, HyuA, Cad, HcpC and Omp) and/or with gastric ulcer (HP305 and CagM), and specific recognition in *H. pylori*-positive sera (Cagδ and CagM).

MFI: Median reporter fluorescence intensity.

Online Resource 2

Title: Association between potential confounding factors and *H. pylori* seropositivity in controls

Caption: HP -: Serology positive to less than 4 *H. pylori* proteins; HP +: Serology positive to at least 4 *H. pylori* proteins; BMI: Body mass index; GC: Gastric cancer.

Online Resource 3

Title: Seropositivity to each *H. pylori* protein in controls.

Caption: Seropositivity to each *H. pylori* protein in controls, according to their status of infection: All controls, non-infected (no seropositive to more than 3 *H. pylori* proteins) and infected (seropositive to 4 or more *H. pylori* proteins) controls.

Online Resource 4

Title: Correlation of seroreactivity to the 16 studied *H. pylori* proteins in infected participants

Caption: Numbers represent Spearman rho coefficient.