Published in final edited form as:

J Chromatogr B Analyt Technol Biomed Life Sci. 2009 April 15; 877(0): 1193–1199. doi:10.1016/j.jchromb.2009.03.003.

Proteomic *Helicobacter pylori* biomarkers discriminating between duodenal ulcer and gastric cancer

Ghalia Khoder^a, Yoshio Yamaoka^b, Jean-Louis Fauchère^a, Christophe Burucoa^a, and Christo Atanassov^{a,*}

^aEA 4331 LITEC, Université de Poitiers, France

^bVeteran Affairs Medical Center and Baylor College of Medicine, Houston, TX, USA

Abstract

Protein patterns of 129 *Helicobacter pylori* strains isolated from Korean and Colombian patients suffering from duodenal ulcer or gastric cancer were analyzed by the high-throughput methodology SELDI-TOF-MS. Eighteen statistically significant candidate biomarkers discriminating between the two clinical outcomes were selected by using the Mann–Whitney test. Three biomarker proteins were purified and identified as a neutrophil-activating protein NapA (HU HPAG1_0821), a RNA-binding protein (HPAG1_0813), and a DNA-binding histone-like protein HU, respectively (jhp0228). These novel biomarkers can be used for development of diagnostic assays predicting the evolution to gastric cancer in *H. pylori*-infected patients.

Keywords

Helicobacter pylori; SELDI-TOF-MS; Biomarker; Gastric cancer; Duodenal ulcer

1. Introduction

The bacterium *Helicobacter pylori* is a human pathogen which colonizes the gastric mucosa and causes inflammation resulting in different clinical outcomes such as chronic and atrophic gastritis, gastro-duodenal ulcers, gastric MALT lymphoma and gastric cancer. To date, there exist no reliable diagnostics to predict, among the very large and clinically heterogeneous group of *H. pylori*-infected patients, the small group at risk for developing gastric cancer. Efforts for developing better predictive diagnostics have focused mostly on disease-related biomarkers of *H. pylori* that are explored at gene level [1–3]. However, specific gene patterns associated with a particular clinical outcome have proven with limited clinical relevance. For example, the combination of the extensively studied factors of pathogenicity *cagA* and *babA* cannot help to segregate a particular group of *H. pylori* virulent strains because of the wide distribution of these factors over strains [4]. With the complete sequencing of four *H. pylori* strains available on the net, (for strains 26695, J99

and HPAG1: <http://cmr.jcvi.org/tigr-scripts/CMR/shared/Genomes.cgi>; for strain Shi470: NCBI, accession number NC 010698), proteomic technologies hold promise for better disease-specific classification of the *H. pylori* strains. The underlying approaches are more or less straightforward. By the first approach, direct proteomics comparison is made between 2D protein maps of *H. pylori* strains [5, 6]. Through using the second, indirect immunoproteomics approach, the *H. pylori* proteins are separated by 2D-E, transferred to membranes, and screened by a panel of sera collected from *H. pylori*-infected patients with different gastric pathologies. The most frequently immunorecognized spots are used for precise localisation of the protein antigens on the preparative gel [7–12]. For example, Krah et al. compared by immunoproteomics two groups of 30 sera each from patients with duodenal ulcer or gastric cancer. They used 2D-E gels of the completely sequenced *H. pylori* strain 26695 for immunostaining with patient's sera, and successfully identified 14 antigenic protein species that are discriminating between the two clinical outcomes [10]. Because of methodological limitations, most studies are based on very small number *H. pylori* strains, which cannot allow to estimate statistical significance. Although some protein biomarkers have been associated with severe gastric pathologies, a more in-depth investigation is required.

This issue can be addressed by studying subproteomes of *H. pylori* by means of the high-throughput technology, SELDI-TOF-MS, which allows the generation of protein patterns from hundreds of strains within a same experiment [13]. In the present study, we used more than hundred *H. pylori* strains from two distinct geographic origins, Colombia and South Korea. The *H. pylori* strains were obtained from patients with gastric cancer (GC) or duodenal ulcer (DU). The latter serves as a control group, since it is well documented that patients with duodenal ulcer practically never develop GC [14, 15]. Eighteen biomarkers were selected in this way and three of them were purified and identified.

2. Materials and methods

2.1. Bacterial strains and cell growth conditions

Characterization [1, 3, 16] and cell growth conditions [17] of the 129 *H. pylori* strains used in this study have been described previously. Bacterial cells, controlled for the presence of more than 95% bacillary forms by Gram staining, were harvested and washed in 10mL of PBS, pH 7.4 containing 1mM of PMSF. After centrifugation at 2000×g for 15min, the cell pellet was frozen and stored at –80 °C.

2.2. Protein extract

The bacterial cell pellets of three Petri plates (9cm in diameter) were thawed and resuspended in 1mL of lysis buffer (16mM Na₂HPO₄, 4mM NaH₂PO₄, 150mM NaCl, 1% Triton X-100, 1mM PMSF). The cell suspension was transferred into a FastProtein Blue tube and homogenized in a FastPrep apparatus (MP Biomedicals) according to the following protocol: four cycles, 40 s each at power setting 6.5; between each cycle—cooling of the tubes in ice for 5min. After centrifugation at 15,000×g for 15 min, supernatants of each sample were divided into several aliquots and stored at –80 °C.

2.3. ProteinChip array processing

Two types of ProteinChip ion-exchange arrays, Q10 and CM10, were assembled into a 96-well bioprocessor (Bio-Rad) and preactivated for 30 min with their respective buffers (100mM Tris/HCl, pH 9.0 and 100mM sodium acetate, pH 4.0). In the next step, 180 μ L of the binding buffer for the respective array were mixed with 20 μ L of the protein extract at 0.5 mg/mL, and were incubated for 60 min. All protein samples were tested in triplicate. After two washes with binding buffers and one quick rinse with HPLC grade water, the spots were processed two times with 1 μ L of saturated solution of sinapinic acid in 50% ACN/0.5% TFA (v/v). All steps were carried out at room temperature (18–20 °C), using a Micromix-5 platform shaker and a robot-pipetting workstation Biomek 3000 (Beckman-Coulter). The processed spots were analyzed using a PCS 4000 ProteinChip Reader (Bio-Rad) in a positive ion mode, at ion acceleration potential of 20 kV.

2.4. Spectra processing and statistics using ProteinChip Data Manager 3.0.7 software (Bio-Rad)

After calibration and normalization of all spectra using the total ion current method, clusters of peaks of the same mass were defined at the following settings: signal-to-noise ratio (S/N) (first pass) 5, minimum peak threshold: 20%, mass error: 0.3% and S/N (second pass) 2. Only clusters containing peaks with signal-to-noise ratio 5 were subjected to the Mann–Whitney test.

The data dealing with the clustered peaks (strain name, clinical group, m/z intensity, protocol used, array type, etc.) can be provided on request as an Excel file (.csv).

2.5. Purification and identification

The selected candidate biomarkers were initially purified by a combination of size-exclusion chromatography using Chroma Spin™ 30 (Clontech, ref. 636087) and reversed phase high-performance liquid chromatography (RP-HPLC) using the PerkinElmer HPLC instrument, series 200. Samples from SEC were applied on the RP-HPLC Prosphere C4 column, 300 Å (Alltech-Grace, USA), under the following elution conditions: buffer A—98.9% H₂O, 1% ACN, 0.1% TFA; buffer B—9.9% H₂O, 90% ACN and 0.1% TFA; linear gradient between buffer A and B—20 min; flow rate—0.8 ml/min; 40 °C temperature of the column oven; absorbance—220 nm; fraction size—1 mL. All RP-HPLC fractions were concentrated 20-fold in a miVac (Genevac) and 3 μ L of the concentrate was spotted on gold arrays and tested by SELDI for the presence of a target protein. The fraction containing the target protein was further concentrated in a miVac to a final volume of ca. 10 μ L, and mixed with an equal volume of tricine SDS sample buffer (2 $\times$) and 2 μ L of NuPAGE reducing agent (10 $\times$) (Invitrogen). This was then divided into two samples that were heated at 85 °C for 2 min, and were separated in parallel by 1D SDS-PAGE in 10–20% Novex Tris-tricine gels (Invitrogen), for 2 h at constant 125V. The gels were stained with colloidal Coomassie G-250, and one of the two protein bands whose location corresponded to the Mr of the target protein was excised and submitted to passive elution [18]. In this way, the mass of the passively eluted protein from the first protein band was definitely confirmed by SELDI using gold arrays. The proteins in the second band were trypsinized and identified by means

of LC–MS/MS, as routinely performed on the proteomics platform of the University of Bordeaux (<http://www.pgfb.u-bordeaux2.fr/proteome/index.html>).

3. Results

3.1. Reproducibility assays

Because of the diversity of *H. pylori* depending partially from environment, the first issue to address was to check how selected cell growth conditions affect the stability of phenotypic protein profiles obtained by SELDI. For this purpose, two randomly selected strains (one Korean and one Colombian) were cultured independently eight times under these same conditions. Respective protein patterns of the eight samples, tested in duplicate, were obtained using two types of arrays, CM10 and Q10. In this regard, Fig. 1 shows an example of the spectra of the Korean strain Dual 45 which display minimal variations in peak intensity (Q10 surface). Similar results were obtained with the second CM10 surface and the Colombian strain CA 66 (not shown).

3.2. EDM and candidate biomarker selection

The protein patterns of 129 *H. pylori* strains were analyzed on two types ProteinChip arrays, CM10 and Q10. Seventy-two strains were from South Korea (GC, $n=41$ et DU, $n=31$), and 57 from Colombia (GC, $n=27$ et DU, $n=30$). Univariate analysis of the EDM data (Mann–Whitney test) generated nine *H. pylori* proteins per country that are discriminating between DU and GC (Table 1). These discriminating proteins are hereafter referred to as candidate biomarkers. They were detected in the measuring range of 2.5–30 kDa.

Among the nine Korean candidate biomarkers (Table 1), the protein levels of five of them were higher in the GC strains (9.3 kDa, 7.4 kDa, 9.5 kDa, 13.2 kDa and 3.7 kDa), and other four—in the DU strains (10.4 kDa, 12.8 kDa, 10.6 kDa and 13.0 kDa). Fig. 2 shows an example of the spectral view and the scatter plot of b10.4 (the prefix “b” stands for biomarker, the numerals—for Mr in kDa), one of the most significant candidate biomarkers in the Korean strains ($P=0.00001$).

In the Colombian strains, the protein levels of six candidate biomarkers were higher in the GC strains (10.3 kDa, 8.5 kDa, 4.5 kDa, 17.5 kDa, 8.4 kDa and 25.0 kDa) and those of three other candidate biomarkers—in the DU strains (4.8 kDa, 7.5 kDa and 9.4 kDa).

3.3. Purification and identification of biomarkers

Three candidate biomarker proteins, b8.4, b9.3 and b10.4 were selected for further purification in view of identification by LC–MS/MS. For each protein, we applied a specific combination of subsequent purification steps including SEC, RP-HPLC and 1D SDS-PAGE. Throughout purification, all obtained fractions were systematically tested by SELDI for the presence of the protein target.

In this regard Fig. 3 concerns not only b10.4, but also illustrates the general purification scheme used. The crude extract (Fig. 3A) was initially separated by SEC in two fractions, above and below 30 kDa. The fraction of proteins with Mr less than 30 kDa was used for further separation and enrichment of b10.4 by RP-HPLC (Fig. 3B). The next step was

separation of the fraction enriched with b10.4 by 1D SDS-PAGE, where equal amounts of the sample were applied on two adjacent lanes (lanes 2 and 3) (Fig. 3C). After final quality control, on gold arrays, of the purity of the protein passively eluted from one of the two polyacrylamide gel bands (Fig. 3D), the protein from the other band was subjected to LC-MS/MS (Fig. 3E). Data search showed that the masses of five peptides matched those of a DNA-binding protein HU HPAG1 0821 of the sequenced *H. pylori* strain HPAG1. This represented 40% sequence coverage of the protein, thereby confirming its identification. The DNA-binding protein HU HPAG1 0821 is composed of 94 amino acids, its Mr is 10.4 kDa, its pI—9.69. The protein is highly conserved among the three other sequenced *H. pylori* strains 26695, J99 and Shi470 with 96–98% identity (the respective annotations for these strains are HP0835, jhp0744 and HPSH_02625); it also shares homology with *H. acinonychis* str. *Sheeba* (88% identity) and, to a lesser extent—with *H. hepaticus* ATCC 51449 (62% identity).

The purification of b9.3 was performed by using the same approach. Identification by LC-MS/MS revealed the identity of the *H. pylori* target protein as being a RNA-binding protein. Two main peptides matched those of RNA-binding protein HPAG1 0813. They cover 43% of the whole protein sequence: the first was K.NIYVGNLVYSATSEQVK.E (experimental mass = 1886.09 Da), and the second was K.GFGFVEM*QEEGVSEAIK.L (experimental mass = 1945.13 Da). The RNA-binding protein HPAG1_0813 is composed of 82 amino acids, its theoretical Mr is 9327.6, the pI—9.61. Homology search in non-redundant database showed that this protein is identical to HPSH_02670 of the recently sequenced strain Shi470. It is also highly conserved among the two other sequenced *H. pylori* strains 26,695 and J99 (identity 97–98%, respective annotations—HP0827 and jhp 0766), and shares homology with *H. acinonychis* str. *Sheeba* (92% identity) and, to a lesser extent—with *H. hepaticus* ATCC 51449 (60% identity).

Three main peptides of the third target protein b8.4, identified by LC-MS/MS, matched those of the neutrophil-activating protein NapA, jhp0228 (strain J99): the first was E.LSNTAEKE.G (experimental mass = 890.4822 Da), the second was E.ILGDYKHLE.K (experimental mass = 1086.6428 Da), and the third was E.FADMFDLAE.R + Oxidation (M) (experimental mass = 1188.5496 Da). NapA jhp0228 is composed of 144 amino acids; its theoretical Mr and pI are, respectively, 16,836.2 Da and 5.69. Homology search showed that this protein is also highly conserved (97% identity) among the proteins of the three other sequenced human *H. pylori* strains 26695 (HP0243), HPAG1 (HPAG1 0246) and Shi470 (HPS 01265); it also shares homology with *H. acinonychis* str. *Sheeba* (95% identity).

The stable protein levels of the identified biomarkers b8.4, b9.3 and b10.4 was confirmed after eight consecutive *in vitro* passages of one and a same *H. pylori* strain. In this regard, Fig. 4 shows the protein abundance of b9.3 and b10.4. The stable protein level of identified biomarker b8.4 was also obtained through using two Colombian *H. pylori* strains (data not shown).

4. Discussion

As observed on several *Helicobacter* species, although common proteins are detected between initial clinical and culture collection strains, greater variability occurs between the clinical and culture collections strains than among strains of the culture collection alone [19]. A possible explanation to that phenomenon is that after continuous passages under identical conditions *in vitro* adaptation of the *H. pylori* strains results in stabilization of the pattern of expressed proteins. In our investigation we used strains that have been cultured several times which would account for a relatively stable phenotype as shown in Fig. 1A and B.

Univariate analysis highlights 18 biomarker proteins defined by their masses (Table 1), whose discriminative potential in terms of clinical outcome merits to be explored in larger series of clinically and histologically documented *H. pylori* strains originating from geographically homogenous populations of patients.

The results of our analysis reveal *H. pylori* protein biomarkers discriminating between duodenal ulcer and gastric cancer for each of the countries. This is not unexpected since it is known that *H. pylori* populations of East Asia and South America show high level of genotypical diversity [20].

The identified DNA-binding protein of *H. pylori*, a histone-like HU protein, is presumed to contribute to inflammation and damage of gastric epithelial cells [21, 22].

The RNA-binding protein of *H. pylori* shares 44% identity with a human protein, a cleavage stimulation factor, 3' pre-RNA subunit. The common region for both RNA-binding proteins is part of the RNA recognition motif (RRM) [23]. Interestingly, the mammalian RNA-binding protein Musashi-1 (Msi-1) contains two RRM in its N-terminal part, which are the same as those of its bacterial homolog. While Msi-1 is underexpressed in human intestinal meta-plasia [24], the *H. pylori* infection strongly induces its expression in gastric progenitor cells [25].

The NapA protein has been shown to promote neutrophil adhesion to endothelial cells [24]. It is considered as an important virulence factor that is overexpressed in *H. pylori* strains infecting polymorphonuclear leukocytes and the human gastric epithelial cell line AGS [27]. NapA is considered as a major antigen in the human immune response to *H. pylori* [28].

The three identified candidate biomarkers of *H. pylori* in our study have not been reported before to be associated to any particular gastric pathology. Interestingly, the NapA was not found by immunoproteomics as a candidate antigen discriminating between German patients with duodenal ulcer or gastric carcinoma [10], while in a recent Chinese study the recombinant protein of NapA has been detected by sandwich ELISA in almost 90% of 151 patients infected by *H. pylori* [29]. This may be explained by conformational changes in the critical epitopes of the NapA protein blotted on the nitrocellulose membrane in the study of Krah et al. [10], which cannot be recognized by corresponding antibodies, or some post-translational modifications of the NapA which are affecting its antigenicity. In this letter regard NapA separated in several spots has been documented with MS data in the 2-DE

database of the Max Planck Institute, Berlin, which is based on the completely sequenced *H. pylori* strain 26695 (<http://www.mpiib-berlin.mpg.de/2D-PAGE>).

From a methodological point of view, we demonstrated that SELDI can be used successfully for generating and analysing 100–200 spectra within 3 days. In this latter context, SELDI has already been applied for searching predictive biomarkers in serum of patients with gastric cancer [30]. SELDI is not only useful for generation of protein biomarkers defined by their masses, but also in purification of the most significant ones in view of structural identification. Throughout purification steps, all obtained fractions are systematically tested by SELDI for the presence of target protein, which is very practical. The SELDI technology presents some limitations. Firstly, irrespective of the nature of the chromatography surface of the ProteinChip used, the most abundant protein patterns are obtained in the range of 2.5–30 kDa. However, the number of proteins in the Mr range up to 30 kDa is ca. 50 % among the theoretically predicted proteomes of the completely sequenced *H. pylori* strains 26695, J99 and HPAG1, which indicates that SELDI is an appropriate approach for studying this particular subproteome. Another limitation of SELDI is that even in the range of 2.5–30 kD, the number of all detectable proteins with S/N > 5 is about 100–200. In this latter regard, the number of the detectable proteins may be increased by using ProteinChip arrays with different surface chemistries, and/or modifying the binding conditions (pH, ionic strength of used buffers). In practice, increasing the initial conditions of protein binding in four of the most frequently used ProteinChip arrays (ion-exchange CM10 and Q10, hydrophobic H50, and metal-affinity IMAC30 with chelated copper or zinc) augments the number of the detectable proteins, but it still remains far below of those of the 2D-E (usually 1500–2000 spots per gel). For comparative proteomic analysis of low number of samples, the SELDI methodology turns more restrictive than the conventional 2D-E approach. The advantage of the SELDI in terms of throughput becomes apparent for analysis of large series of samples. In this study, we tested 139 strains, while testing in triplicate the same number of samples by 2D-E would imply 417 gels. SELDI is also helpful in analysing on appropriate arrays some “difficult” proteins (acidic, basic, hydrophobic, membrane, etc.), as well as proteins of Mr below 3 kDa, which are not detectable by 2D-E. Compared to 2D-E, the starting material of a crude protein extract is at least 15 times less for running a SELDI experiment. In our hands, it was 30µg for a sample tested in triplicate (10µg per spot), while running the same experiment by 2D-E usually requires 150–300µg for analytical gel or 500µg for preparative gel [10]. SELDI allows comparisons of clustered peaks with the same mass, which excludes the possibility to detect post-translational modifications and/or partly degraded fragments of one and a same protein. The latter can be done by conventional proteomics and immunoproteomics [9, 12]. Protein identification by LC–MS/MS may be performed in enzymatically treated proteins from the 2D-E spots, while identification of the proteins selected by using SELDI requires additional steps of purification [13].

In summary, despite the above-mentioned concerns, the SELDI technology allows to carry out rapidly high-throughput comparative proteomics of hundreds and thousands of samples, which would be cumbersome and technically difficult to be performed by using 2D-E in most of the academic laboratories. The gain in rapidity and throughput in the initial screening performed by SELDI, as well its capacity to detect differentially “difficult”

proteins (small Mr, hydrophobic, basic, acidic, etc.) makes this technology an useful complement to the widely used 2D-E approach in the field of the comparative proteomics.

In conclusion, by using SELDI for high-throughput screening and SELDI-assisted purification, we propose three novel candidate biomarkers discriminating duodenal ulcer from gastric cancer. The review of current literature shows that their possible involvement in gastric cancerogenesis mechanisms necessitates further elucidation. The identification of their primary structure opens possibilities of producing them as recombinant proteins, and to raise specific antibodies to their whole molecule or to synthetic peptides mimicking their antigenic regions. These could be used in functional studies to elucidate their mechanism of action on gastric epithelial cells, as well as for the development of novel immunochemical tests for better evaluation of the risk for developing gastric cancer.

Acknowledgments

Prof. David Y. Graham (Veteran Affairs Medical Center and Baylor College of Medicine, Houston, TX, USA), Dr. Sabine Jourdain (Bio-Rad, France), Dr. Jacques d'Alayer (Institut Pasteur, Paris), Dr. Stephane Claverol (University of Bordeaux, France) and Dr. Maya Belghazi (University of Marseille, France). Special thanks to Mrs. Corinne Lacombe and Mrs. Marylise Perivier (CHU de Poitiers, France) for the technical assistance.

Grant support: Ligue contre le cancer, Comité des Deux-Sèvres, France (C.A.), Agence Universitaire de la Francophonie (G.K.).

Abbreviations

2D-E	two-dimensional electrophoresis
ACN	acetonitrile
DU	duodenal ulcer
EDM	expression difference mapping
GC	gastric cancer
<i>H. pylori</i>	<i>Helicobacter pylori</i>
HPLC	high-performance liquid chromatography
kDa	kilodaltons
ELISA	enzyme-linked immunosorbent assay
LC-MS/MS	liquid chromatography/tandem mass spectrometry
Mr	molecular weight
S/N	signal-tonoise ratio
PMSF	phenylmethylsulphonyl fluoride
pI	isoelectric point
RP-HPLC	reversed-phase high-performance liquid chromatography
SEC	size-exclusion chromatography

SELDI-TOF-MS	surface-enhanced laser desorption/ionization time-of-flight mass spectrometry
TFA	trifluoroacetic acid

References

1. Yamaoka Y, Soucek J, Odenbreit S, Haas R, Arnqvist A, Borén T, Kodama T, Osato MS, Gutierrez O, Kim JG, Graham DY. *J Clin. Microbiol.* 2002; 40:2244. [PubMed: 12037098]
2. Lu H, Hsu PI, Graham DY, Yamaoka Y. *Gastroenterology.* 2005; 128:833. [PubMed: 15825067]
3. Yamaoka Y, Ojo O, Fujimoto S, Odenbreit S, Haas R, Gutierrez O, El-Zimaity HM, Reddy R, Arnqvist A, Graham DY. *Gut.* 2006; 55:775. [PubMed: 16322107]
4. Höcker M, Hohenberger P. *Lancet.* 2003; 362:1231. [PubMed: 14568748]
5. Jungblut PR, Bumann D, Haas G, Zimny-Arndt U, Holland P, Lamer S, Siejak F, Aebischer A, Meyer TF. *Mol. Microbiol.* 2000; 36:710. [PubMed: 10844659]
6. Park JW, Song JY, Lee SG, Jun JS, Park JU, Chung MJ, Ju JS, Nizamutdinov D, Chang MW, Youn HS, Kang HL, Baik SC, Lee WK, Cho MJ, Rhee KH. *Helicobacter.* 2006; 11:533. [PubMed: 17083375]
7. Kimmel B, Bosserhoff A, Frank R, Gross R, Goebel W, Beier D. *Infect. Immun.* 2000; 68:915. [PubMed: 10639463]
8. Haas G, Karaali G, Ebermayer K, Metzger WG, Lamer S, Zimny-Arndt U, Diescher S, Goebel UB, Vogt K, Roznowski AB, Wiedenmann BJ, Meyer TF, Aebischer T, Jungblut PR. *Proteomics.* 2002; 2:313. [PubMed: 11921447]
9. Krah A, Jungblut PR. *Methods Mol. Med.* 2004; 94:19. [PubMed: 14959820]
10. Krah A, Miehle S, Pleissner KP, Zimny-Arndt U, Kirsch C, Lehn N, Meyer TF, Jungblut PR, Aebischer T. *Int. J. Cancer.* 2004; 108:456. [PubMed: 14648714]
11. Lin YF, Wu MS, Chang CC, Lin SW, Lin JT, Sun YJ, Chen DS, Chow LP. *Mol. Cell Proteomics.* 2006; 5:1484. [PubMed: 16769709]
12. Mini R, Bernardini G, Salzano AM, Renzone G, Scaloni A, Figura N, Santucci A. *J Chromatogr. B: Anal. Technol. Biomed. Life Sci.* 2006; J833:63.
13. Issaq HJ, Veenstra TD, Conrads TP, Felschow D. *Biochem. Biophys. Res. Commun.* 2002; 292:587. [PubMed: 11922607]
14. Hansson LE, Nyrén O, Hsing AW, Bergström R, Josefsson S, Chow WH, Fraumeni JF, Adami HO Jr. *N Engl. J. Med.* 1996; 335:242.
15. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. *N Engl. J. Med.* 2001; 345:784.
16. Yamaoka Y, Kodama T, Gutierrez O, Kim JG, Kashima K, Graham DY. *J Clin. Microbiol.* 1999; 37:2274. [PubMed: 10364597]
17. Atanassov C, Pézenne L, d'Alayer J, Grollier G, Picard B, Fauchère JL. *J Clin. Microbiol.* 2002; 40:547. [PubMed: 11825970]
18. Le Bihan MC, Hou Y, Harris N, Tarelli E, Coulton GR. *Proteomics.* 2006; 6:4646. [PubMed: 16858738]
19. Hynes SO, McGuire J, Wadström T. *FEMS Immunol. Med. Microbiol.* 2003; 36:151. [PubMed: 12738385]
20. Linz B, Schuster SC. *Res. Microbiol.* 2007; 158:737. [PubMed: 18023330]
21. Chen C, Ghosh S, Grove A. *Biochem. J.* 2004; 383(Pt 2):343. [PubMed: 15255779]
22. Kim N, Weeks DL, Shin JM, Scott DR, Young MK, Sachs G. *J Bacteriol.* 2002; 184:6155. [PubMed: 12399485]
23. Maris C, Dominguez C, Allain FH. *FEBS J.* 2005; 272:2118. [PubMed: 15853797]
24. Akasaka Y, Saikawa Y, Fujita K, Kubota T, Ishikawa Y, Fujimoto A, Ishii T, Okano H, Kitajima M. *Histopathology.* 2005; 47:348. [PubMed: 16178889]

25. Murata H, Tsuji S, Tsujii M, Nakamura T, Fu HY, Eguchi H, Asahi K, Okano H, Kawano S, Hayashi N. *Dig. Dis. Sci.* 2008; 53:363. [PubMed: 17549630]
26. Evans DJ Jr, Evans DG, Takemura T, Nakano H, Lampert HC, Graham DY, Granger DN, Kvietys PR. *Infect. Immun.* 1995; 63:2213. [PubMed: 7768601]
27. Gieseler S, König B, König W, Backert S. *Microbes Infect.* 2005; 7:437. [PubMed: 15788154]
28. Satin B, Del Giudice G, Della Bianca V, Dusi S, Laudanna C, Tonello F, Kelleher D, Rappuoli R, Montecucco C, Rossi F. *J Exp. Med.* 2000; 191:1467. [PubMed: 10790422]
29. Tang RX, Luo DJ, Sun AH, Yan J. *World J. Gastroenterol.* 2008; 14:4816. [PubMed: 18720546]
30. Ebert MP, Meuer J, Wiemer JC, Schulz HU, Reymond MA, Traugott U, Malfertheiner P, Röcken C. *J Proteome Res.* 2004; 3:1261. [PubMed: 15595736]

(B)

Chip array	Mr Range	Nb of Peaks	CVp
Q10	3-30 kDa	15	20%
CM10	3-30 kDa	10	22%

(C)

Chip array	Mr Range	Nb of Peaks	CVp type	CVp
CM10	3-30 kDa	21	Inter-array	16%
CM10	3-30 kDa	17	Intra-array	16%
Q10	5-25 kDa	21	Inter-array	16%
Q10	5-25 kDa	17	Intra-array	19%

Fig. 1.

Reproducibility of protein profiles of bacterial cultures: (A) spectra of eight independent cultures (i.e. cultured on different times, each culture—started from a separate frozen aliquote) of one and a same *H. pylori* strain (Dual 45) tested on the Q10 surface. Four from the eight obtained spectra are represented; (B) 10–15 peaks (mass range: 3–30 kDa, $S/N \geq 5$) are used from the eight spectra of strain Dual 45 to calculate the pooled coefficients of variance (CVp) of the Q10 and CM10 surfaces. Reproducibility of the SELDI methodology: (C) intra- and inter-array reproducibility of the Q10 and CM10 surfaces as calculated by

CVp of 17–21 peaks from one and a same aliquot applied 8 times on 8 spots of the same array (intra-array), as well as on one spot per array of 8 different arrays (inter-array).

Fig. 2.

Relative protein abundance of biomarker b10.4 in *H. pylori* Korean strains. (A) Gastric cancer strains: (a), (b) and (c); duodenal ulcer strains: (d), (e) and (f). X-axis: m/z ($z = 1$); Y-axis: peak intensity. (B) Scatter plot (Mann–Whitney test) of b10.4 (Q10 surface): (red) GC strains and (blue) DU strains.

Fig. 3.

Purification and identification of biomarker protein b10.4. (A) Spectrum of the crude protein extract (Korean *H. pylori* strain “Dual 45”). (B) Spectrum of the b10.4 fraction obtained after HPLC separation of the same crude protein extract. (C) Separation by 1D SDS-PAGE (Novex tris-tricine gel, 10–20%) of the HPLC fraction containing b10.4: (lane 1) Mark12 protein standards; (lanes 2 and 3) bands containing b10.4. (D) Quality control of purity of the passively eluted b10.4 from Novex tris-tricine 10–20% gel. (E) LC–MS/MS

identification of the b10.4 biomarker shows a DNA-binding protein (94 amino acids) sequenced through 4 peptides marked (in red), with a percent coverage equal to 40%.

Fig. 4. Stable protein levels of identified biomarkers b9.3 and b10.4 after eight consecutive *in vitro* passages of one and a same *H. pylori* strain ("Dual 45").

Table 1

Protein biomarkers discriminating *H. pylori* strains isolated from patients suffering from gastric cancer or duodenal ulcer.

Country/ProteinChip	Biomarkers ^a	P-Value ^b	ROC-area ^c	Up-regulated
Korea/CM10	9.3 kDa	0.001	0.277	GC
	7.4 kDa	0.017	0.329	GC
	9.5 kDa	0.021	0.310	GC
	13.2 kDa	0.027	0.329	GC
Korea/Q10	10.4 kDa	0.00001	0.647	DU
	3.7 kDa	0.0003	0.218	GC
	12.8 kDa	0.021	0.659	DU
	10.6 kDa	0.026	0.634	DU
	13.0 kDa	0.044	0.659	DU
Colombia/CM10	10.3 kDa	0.014	0.299	GC
	8.5 kDa	0.035	0.323	GC
	4.5 kDa	0.044	0.323	GC
	4.8 kDa	0.048	0.660	DU
Colombia/Q10	17.5 kDa	0.0008	0.210	GC
	7.5 kDa	0.009	0.701	DU
	8.4 kDa	0.024	0.328	GC
	9.4 kDa	0.025	0.671	DU
	25.0 kDa	0.039	0.309	GC

^a In each ProteinChip array series, biomarkers are ordered by increasing *P*-value.

^b *P*-Value determined by Mann–Whitney test.

^c Duodenal ulcer is used as a positive group in the ROC plot.