H. pylori is the strongest known risk factor for gastric adenocarcinoma, yet only a fraction of infected persons ever develop cancer. The extensive genetic diversity inherent to this pathogen has precluded comprehensive analyses of constituents that mediate carcinogenesis. We previously reported that in vivo adaptation of a non-carcinogenic H. pylori strain endowed the output derivative with the ability to induce adenocarcinoma, providing a unique opportunity to identify proteins selectively expressed by an oncogenic H. pylori strain. Using a global proteomics DIGE/MS approach, a novel missense mutation of the flagellar protein FlaA was identified that affects structure and function of this virulence-related organelle. Among 25 additional differentially abundant proteins, this approach also identified new proteins previously unassociated with gastric cancer, generating a profile of H. pylori proteins to use in vaccine development and for screening persons infected with strains most likely to induce severe disease. Molecular & Cellular Proteomics 8:1947–1958, 2009.

Helicobacter pylori is a Gram-negative bacterial species that selectively colonizes gastric epithelium and induces an inflammatory response within the stomach that persists for decades (1, 2). Biological costs incurred by the long term relationship between H. pylori and humans include an increased risk for distal gastric adenocarcinoma (3–8), and eradication of this pathogen significantly decreases cancer risk among infected individuals without premalignant lesions (9). However, only a fraction of colonized persons ever develop neoplasia, and enhanced cancer risk is related to H. pylori strain differences, inflammatory responses governed by host genetic diversity, and/or specific interactions between host and microbial determinants (10).

From the ‡Division of Gastroenterology, Departments of Medicine and Cancer Biology, §Mass Spectrometry Research Center, Department of Biochemistry, and Departments of ¶Biological Sciences and **Pathology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232 and ¶¶School of Biomolecular and Biomedical Science, University College Dublin, Belfield Campus, Ardmore House, Dublin 4, Ireland

Received, March 13, 2009, and in revised form, May 5, 2009


H. pylori strains are remarkably diverse (11–15), and the genetic composition of strains can change over time within an individual colonized stomach (16, 17). Despite this diversity, several genetic loci have been identified that augment disease risk. The cag pathogenicity island encodes a type IV bacterial secretion system, and the product of the terminal gene in this island, CagA, is translocated into host epithelial cells by the cag secretion system following adherence (18–20). Within the host cell, CagA undergoes Src- and Abl-dependent tyrosine phosphorylation (21) and activates the eukaryotic phosphatase SHP-2, leading to dephosphorylation of host cell proteins and cellular morphological changes (19–21). CagA also dysregulates β-catenin signaling (22, 23) and apical-junctional complexes (24), events linked to increased cell motility and oncogenic transformation in several models (25, 26). Another H. pylori constituent linked to gastric cancer is the cytotoxin VacA, encoded by the gene vacA, which is present in virtually all H. pylori strains (27). In vitro, VacA induces the formation of intracellular vacuoles (27) and can induce apoptosis (28), and vacuolating activity is significantly associated with the presence of the cag pathogenicity island (3).

Approximately 20% of H. pylori bind to gastric epithelial cells in vivo (29), and sequence analysis has revealed that the H. pylori genome contains an unusually high number of ORFs relative to its genome size that are predicted to encode outer membrane proteins (15). BabA, a member of a family of highly conserved outer membrane proteins and encoded by the strain-specific gene babA2, binds the Lewis histo-blood group antigen on gastric epithelial cells (30, 31), and H. pylori babA2+ strains are associated with an increased risk for gastric cancer (30). However, not all persons infected with cag− babA2+ toxigenic strains develop gastric cancer, indicating that additional H. pylori constituents are important in carcinogenesis.

We recently identified a strain of H. pylori, 7.13, that reproducibly induces gastric cancer in two rodent models of gastritis, Mongolian gerbils and hypergastrinemic INS-GAS mice (22). This strain was derived via in vivo adaptation of a clinical H. pylori strain, B128, which induces inflammation, but not cancer, in rodent gastric mucosa. The oncogenic 7.13 phenotype is not due to an enhanced ability of strain 7.13 to colonize as there were no significant differences in gastric colonization density or efficiency between strains B128 and 7.13 as assessed by either quantitative culture or histology.
However, carcinogenic strain 7.13 binds more avidly to gastric epithelial cells in vitro than does strain B128, suggesting that the two strains may variably express different outer membrane proteins.

To define proteins that may mediate the development of H. pylori-induced gastric cancer, we performed two-dimensional DIGE coupled with MS to identify differentially abundant membrane-associated and cytosolic proteins from non-carcinogenic H. pylori strain B128 and its carcinogenic derivative, strain 7.13. DIGE/MS is a well established proteomics technology based on conventional 2D gel protein separations whereby prelabeling samples with spectrally resolvable fluorescent dyes and multiplexing samples onto a series of gels that contain a mixture of all experimental samples (internal standard) provide quantitative data on abundance changes for thousands of intact proteins from multiple experimental conditions, each measured in replicate for statistical confidence (32–36). Techniques including DIGE/MS have recently been utilized to robustly define differences in protein abundance profiles between bacterial strains and to compare expression patterns of proteins harvested from bacteria maintained under different growth conditions (37, 38).

Utilizing DIGE/MS, we detected and identified 26 proteins with statistically significant differences between strains B128 and 7.13, including a novel cysteine-to-arginine mutation in the H. pylori flagellar protein FlaA. We demonstrate that this FlaA mutation results in structural and functional aberrations. Application of this technique to two genetically related bacterial strains that induce distinct phenotypes also identified several novel candidate H. pylori virulence factors, providing a framework for studies targeting the pathogenesis of microbiologically induced cancer.

**EXPERIMENTAL PROCEDURES**

**Bacterial Strains and Protein Preparation—**H. pylori strains B128 and 7.13 were grown in Brucella broth with 10% fetal bovine serum (FBS) for 18 h at 37 °C and 5% CO₂ with shaking and harvested by centrifugation at 6,000 × g for 10 min as described previously (22). Bacterial pellets were washed once with PBS and then resuspended in PBS with protease inhibitors (Roche Applied Science). To disrupt bacteria, samples were subjected to three rounds of French press-mediated lysis at 20,000 p.s.i. Cellular debris were removed by centrifugation at 5000 × g for 10 min. Membrane-associated proteins were collected by ultracentrifugation at 100,000 × g for 45 min at 4 °C. The supernatant containing cytosol-associated proteins also was collected for subsequent analysis. The pellet, containing membrane-associated proteins, was resuspended in PBS, subjected to two rounds of rapid freeze-thaw cycles and sonication, and used for subsequent analysis. Levels of total protein in both fractions were quantified by the Bradford assay (Pierce) (17).

**DIGE/MS—**Quadruplicate samples from each bacterial preparation were collected by ultracentrifugation at 100,000 × g for 10 min. Membrane-associated proteins were subjected to three rounds of French press-mediated lysis at 20,000 p.s.i. Cellular debris were removed by centrifugation at 5000 × g for 10 min. Membrane-associated proteins were collected by ultracentrifugation at 100,000 × g for 45 min at 4 °C. The supernatant containing cytosol-associated proteins also was collected for subsequent analysis. The pellet, containing membrane-associated proteins, was resuspended in PBS, subjected to two rounds of rapid freeze-thaw cycles and sonication, and used for subsequent analysis. Levels of total protein in both fractions were quantified by the Bradford assay (Pierce) (17).

**DIGE/MS—**Quadruplicate samples from each bacterial preparation were collected by ultracentrifugation at 100,000 × g for 10 min. Membrane-associated proteins were subjected to three rounds of French press-mediated lysis at 20,000 p.s.i. Cellular debris were removed by centrifugation at 5000 × g for 10 min. Membrane-associated proteins were collected by ultracentrifugation at 100,000 × g for 45 min at 4 °C. The supernatant containing cytosol-associated proteins also was collected for subsequent analysis. The pellet, containing membrane-associated proteins, was resuspended in PBS, subjected to two rounds of rapid freeze-thaw cycles and sonication, and used for subsequent analysis. Levels of total protein in both fractions were quantified by the Bradford assay (Pierce) (17).

**DIGE/MS—**Quadruplicate samples from each bacterial preparation were collected by ultracentrifugation at 100,000 × g for 10 min. Membrane-associated proteins were subjected to three rounds of French press-mediated lysis at 20,000 p.s.i. Cellular debris were removed by centrifugation at 5000 × g for 10 min. Membrane-associated proteins were collected by ultracentrifugation at 100,000 × g for 45 min at 4 °C. The supernatant containing cytosol-associated proteins also was collected for subsequent analysis. The pellet, containing membrane-associated proteins, was resuspended in PBS, subjected to two rounds of rapid freeze-thaw cycles and sonication, and used for subsequent analysis. Levels of total protein in both fractions were quantified by the Bradford assay (Pierce) (17).

**DIGE/MS—**Quadruplicate samples from each bacterial preparation were collected by ultracentrifugation at 100,000 × g for 10 min. Membrane-associated proteins were subjected to three rounds of French press-mediated lysis at 20,000 p.s.i. Cellular debris were removed by centrifugation at 5000 × g for 10 min. Membrane-associated proteins were collected by ultracentrifugation at 100,000 × g for 45 min at 4 °C. The supernatant containing cytosol-associated proteins also was collected for subsequent analysis. The pellet, containing membrane-associated proteins, was resuspended in PBS, subjected to two rounds of rapid freeze-thaw cycles and sonication, and used for subsequent analysis. Levels of total protein in both fractions were quantified by the Bradford assay (Pierce) (17).

**DIGE/MS—**Quadruplicate samples from each bacterial preparation were collected by ultracentrifugation at 100,000 × g for 10 min. Membrane-associated proteins were subjected to three rounds of French press-mediated lysis at 20,000 p.s.i. Cellular debris were removed by centrifugation at 5000 × g for 10 min. Membrane-associated proteins were collected by ultracentrifugation at 100,000 × g for 45 min at 4 °C. The supernatant containing cytosol-associated proteins also was collected for subsequent analysis. The pellet, containing membrane-associated proteins, was resuspended in PBS, subjected to two rounds of rapid freeze-thaw cycles and sonication, and used for subsequent analysis. Levels of total protein in both fractions were quantified by the Bradford assay (Pierce) (17).

1 The abbreviations used are: 2D, two-dimensional; PCA, principal component analysis; MOWSE, molecular weight search; FBS, fetal bovine serum; ANOVA, analysis of variance; TCF, T cell factor; LEF, lymphoid enhancer factor; LPS, lipopolysaccharide.
performed on a Voyager 4700 (Applied Biosystems). The resulting peptide mass maps and the associated fragmentation spectra were collectively used to interrogate \textit{H. pylori} sequences (NCBI non-redundant databases) to generate statistically significant candidate identifications using GPS Explorer software (Applied Biosystems) running the MASCOT v1.9 search algorithm (Matrix Science). Database searches were performed allowing for complete carbamidomethylation of cysteine, partial oxidation of methionine residues, and one missed cleavage. Molecular weight search (MOWSE) scores, number of matched ions (and number of unmatched ions), number of matching ions with independent MS/MS matches, percent protein sequence coverage, and correlation of gel region with predicted molecular weight and pl were collectively considered for each protein identification (Table I and supplemental Table 1). In every case reported, the identified \textit{H. pylori} protein was unequivocally matched and distinct from family members/isofoms. Only proteins identifications with MOWSE scores above 77 (threshold for the 95th percentile confidence interval) are reported. Individual ion scores (for MS/MS) above 41 were similarly within the 95th percentile confidence interval, but ion scores below this threshold were still considered for those cases where MOWSE scores were above 77 (Table I and supplemental Table 1).

**Cell Culture, Plasmids, and Reagents**—AGS human gastric epithelial cells were grown in RPMI 1640 medium (Invitrogen) with 10% FBS (Sigma) and 20 μg/ml gentamicin in 5% CO\textsubscript{2} at 37 °C. Topflash and Fopflash reporter plasmids and TCF were kind gifts from K. Kinzler and B. Vogelstein (Johns Hopkins University). \textit{H. pylori} strains were grown in \textit{Brucella} broth with 5% FBS for 18 h, harvested by centrifugation, and added to gastric cells at a bacteria:cell concentration of 100:1 (22).

**DNA Sequencing**—\textit{H. pylori} genomic DNA from strains B128 and 7.13 was extracted as described previously (40). \textit{flaA} was amplified from genomic DNA using the primer pair 5’-CACAACATCTTTGGT-GAGC-3’ and 5’-GGGCGTGAAATCATT-3’ and 5’-GGCGGTCAATGAAATCAGA-3’. A single product was amplified and purified (MinElute PCR purification kit, Qiagen), and sequence determination was performed by the Vanderbilt University DNA sequencing facility with the same primers using BigDye Terminator chemistry and analyzed on an ABI 3730xl DNA Analyzer (GenBank accession numbers EU400215 and EU400216). HP0310 was amplified with primers 5’-ACAAACTCCCAAAGG-3’ and 5’-CTACTGCTCAATGAAAGGAC-3’ and purified as above. Sequencing was performed using the same primer pair and the primer pair 5’-TCAATACGACCCACTCGG-3’ and 5’-GGCGGTCAATGAAATCATT-GGTG-3’ (GenBank accession numbers EU410066 and EU410067). Cloning and Genetic Exchange of \textit{flaA} Genes between \textit{H. pylori} Strains B128 and 7.13—To generate the \textit{flaA}-complemented strain, \textit{flaA} with its own promoter was amplified by PCR from strain 7.13 (\textit{flaA}\textsubscript{7.13}) and cloned into pGEM-Teasy (Promega). A chloramphenicol resistance cassette was inserted within the \textit{flaA} gene to create a suicide knock-out vector as described previously (18). This vector was then used to generate an \textit{flaA} knock-out mutant in strain B128. The successful integration of the chloramphenicol cassette into the \textit{flaA}\textsubscript{B128} gene was confirmed by PCR, and the absence of \textit{flaA} expression was demonstrated by Western blotting using flagellin-specific rabbit antibodies (data not shown). To transfer the intact \textit{flaA}\textsubscript{7.13} gene into the B128 isogenic \textit{flaA} mutant, "flaA\textsubscript{7.13} was fused with another selectable marker, the kanamycin resistance \textit{aphA} cassette (41). The intergenic region between open reading frames HP0999 and HP1000 in the \textit{H. pylori} chromosome was selected as the integration site for complementation of \textit{flaA}\textsubscript{7.13}. This region is in one of two plasticity zones and has no known function in \textit{H. pylori} colonization or virulence. Specifically the \textit{flaA}\textsubscript{7.13}-\textit{aphA} construct was cloned into pBluescript containing the \textit{H. pylori} genes HP0998-1001, and this construct was then used to transform the B128 \textit{flaA} isogenic mutant using natural transformation as described previously (41). Double resistant clones were selected on agar plates containing both kanamycin and chloramphenicol, and re-expression of \textit{flaA}\textsubscript{7.13} in the B128 isogenic \textit{flaA} mutant was confirmed by Western blotting using flagellin-specific rabbit antibodies (data not shown).

**Motility Assays**—Soft agar motility assays for \textit{H. pylori} were performed on medium containing Brucella broth, 0.35% agar, and 10% (v/v) FBS. A stab of \textit{H. pylori} was spotted into soft agar plates and allowed to incubate for 3–5 days under microaerophilic conditions. Motility was assessed by measuring the diameter of outward migration as swarm halos apparent in the soft agar.

**Transmission Electron Microscopy**—All \textit{H. pylori} samples for transmission electron microscopy were grown for 24 h on tryptic soy agar plates under microaerobic conditions. \textit{H. pylori} were transferred to carbon-coated copper grids (Electron Microscopy Sciences) by placing grids directly onto plates for 30 s. Samples were fixed with 1% glutaraldehyde for 1 min, rinsed, and negatively stained with 0.05% uranyl formate and 1% uranyl acetate. Samples were visualized and photographed using a Philips CM12 transmission electron microscope (Vanderbilt Cell Imaging Core) operating at 80 keV.

**Western Analysis**—Isogenic HP0310, HP0391, HP0743, and HP1286 null mutants were constructed within strain 7.13 by insertion of mutagenesis using \textit{apfA} and were selected with kanamycin (25 μg/ml) (22). \textit{H. pylori} were added to AGS cells at a bacteria:cell ratio of 100:1; gastric cell lysates were harvested, and protein concentrations were quantified by the Bradford assay (Pierce). Proteins (20 μg) were separated by SDS-PAGE and transferred to PVDF membranes (Pall Corp.). Levels of intracellular phosphorylated CagA were determined using an anti-pY99 antibody (1:300; Santa Cruz Biotechnology), and actin levels were determined using an anti-actin (C-11) antibody (1:500; Santa Cruz Biotechnology). Primary antibodies were detected using goat anti-rabbit (1:20,000, Sigma) or donkey anti-goat (1:5000; Sigma) horseradish peroxidase-conjugated secondary antibodies, visualized by the ECL detection system (Cell Signaling Technology), and quantified using the ChemiGenius system (Syngene). Densitometric analysis of multiple Western blots was then performed as described previously (22).

**Transfections and Luciferase Assays**—AGS cells (2 × 10\textsuperscript{4}) were transfected with 3 μl of Lipofectamine 2000 (Invitrogen), 1 μg/ml Topflash or 1 μg/ml Fopflash, 0.5 μg/ml TCF/LEF, and 0.5 μg/ml Renilla luciferase reporter in Opti-MEM I (Invitrogen) for 5 h. Transfection mixtures were then replaced with complete medium containing \textit{H. pylori} or medium alone. After 24 h, cells were harvested in 1× Passive Lysis Buffer (Promega). Luciferase activity was determined using a luminometer and was normalized to Renilla luciferase activity using a Dual-Luciferase reporter assay system (Promega).

**Immunofluorescence**—Gastric epithelial cells were cultured on glass cover slides, and cells treated with or without \textit{H. pylori} were washed twice with PBS, permeabilized, and fixed with ice-cold methanol at −20 °C as described previously (22). Slides were incubated in 3% BSA (Sigma) for 10 min and then incubated with rabbit anti-β-catenin antibody (1:100; Sigma) overnight at 4 °C. Washed slides were then incubated with goat anti-rabbit IgG-Cy2 (1:100; Molecular Probes) at room temperature for 30 min. For each sample, at least 100 cells were evaluated by an independent observer unaware of experimental conditions.

**Animals and \textit{H. pylori} Challenge**—All procedures were approved by the Institutional Animal Care Committee of Vanderbilt University. Male Mongolian gerbils (Harlan, Indianapolis, IN) 4–8 weeks of age were ovariectomically challenged with sterile \textit{Brucella} broth or \textit{H. pylori} strains 7.13, an isogenic B128 \textit{flaA} mutant, or an isogenic B128 \textit{flaA} mutant complemented with \textit{flaA} from strain 7.13 and were sacrificed 12–14 weeks postinoculation. One half of the glandular stomach was fixed for histologic examination, and the other half was homogenized.
in sterile PBS, plated on selective Trypticase soy agar plates, and incubated for 3–5 days for \( H. \) pylori culture (22). Linear strips extending from the squamocolumnar junction through proximal duodenum were fixed in 10% neutral buffered formalin. Tissues were paraffin-embedded and stained with hematoxylin and eosin, and indices of inflammation and injury were scored on a 0–3 scale by a single pathologist blinded to treatment groups. Dysplasia/carcinoma were diagnosed using morphologic criteria previously established for gastrointestinal neoplasia in rodent models of disease (22).

Statistical Analysis—Gerbils were considered to be successfully infected if results from histological examination of tissue or culture were positive. ANOVA and Student’s \( t \) tests were used for statistical analyses of intergroup comparisons. Significance was defined as \( p < 0.05 \).

RESULTS

\( H. \) pylori Strains 7.13 and B128 Exhibit Distinct Protein Expression Profiles—We previously demonstrated that \( in \) vivo adapted \( H. \) pylori strain 7.13 reproducibly induces premalignant and malignant gastric lesions in two different rodent models, whereas cancer was not induced by the progenitor strain B128 (22). An initial study using DNA array technology failed to detect any significant differences between the strains (data not shown). This result may be based on the fact that differences are post-translational in nature or that point mutations, gene inversions, or small deletions are difficult to detect with this technology. In addition, \( H. \) pylori DNA arrays are based on sequences derived from two unrelated strains (26695 and J99) and do not contain ORFs that may be unique to strain B128 and/or 7.13. Therefore, we focused on identification of proteomic differences between these two genetically related strains by performing two-dimensional DIGE/MS.

To increase the depth of analysis, we first fractionated minimally passaged whole cell bacterial lysates into membrane and cytoplasmic fractions and then analyzed these proteins independently by DIGE with isoelectric points in the pH 4–7 and pH 7–11 range.

Cytoplasmic and membrane fractions were prepared from each strain independently in quadruplicate, and the 16 resulting extracts were co-resolved across eight DIGE gels coordinated by a Cy2-labeled 16-mixture pooled sample internal standard as described under “Experimental Procedures” (Fig. 1). For the pH 4–7 gels, a total of 842 resolved protein features, including many isoforms that result from charged post-translational modification and/or processing, were matched across all eight gels (Fig. 2A). Within each gel, direct measurements of expression levels for each resolved protein were obtained between Cy3 and Cy5 channels relative to the Cy2 channel independently without the interference of technical variation because the tripartite labeled proteins co-migrate. This methodology enabled normalization of the Cy3/Cy2 and Cy5/Cy2 intragel ratios between gels by virtue of the signal from the Cy2 internal standard, which was present in each gel. This analysis was performed independently for each of the 842 resolved features.

We anticipated that the greatest source of variation among the four sample categories would be between cytoplasmic and membrane fractions with smaller variations resulting from differences between strains. PCA was used to assess the major sources of variation between samples and to distinguish biologically relevant changes from technical artifacts that may have resulted from either sample preparation, gel
electrophoresis, or image acquisition. PCA reduces the complexity of the multidimensional space into two principle components, PC1 and PC2, which orthogonally divide the samples based upon the two largest sources of variation in the data set (Fig. 2B).

The protein expression characteristics from all 842 features from each individual sample are represented by each of the 16 data points in the PCA plot (Fig. 2B). This analysis demonstrated that 80.3% of the variance (PC1) separated cytoplasmic from membrane samples as expected. An additional 5.2% of variance (PC2) separated the carcinogenic strain 7.13 from the non-carcinogenic strain B128. Similar results were obtained for an additional 300 protein features resolved in the pH 7–11 range (data not shown; PC1 = 84.5%, membrane versus cytoplasmic; PC2 = 5%, strain 7.13 versus strain B128).

These results were confirmed using unsupervised hierarchical clustering analysis of the 16 independent 842-member proteome maps whereby samples were clustered based on similarities between global expression patterns (represented as an expression matrix heat diagram; Fig. 2C). These unbiased analyses demonstrated reproducibility between independent replicate samples and an extremely low level of technical variation and highlighted a discrete subset of protein expression changes representing candidate proteins potentially involved in carcinogenesis.

We next proceeded to identify the subset of proteins that were differentially expressed between the two strains using the univariate Student’s t test and ANOVAs (p ≤ 0.05 with many < 0.01; Table I and supplemental Table 1). Twenty proteins (often present as multiple charged isoforms) were identified from the pH 4–7 separations as being present at significantly higher or lower levels in carcinogenic strain 7.13 compared with the parental derivative B128 strain (Table I, supplemental Tables 1 and 2, and supplemental Figs. 1 and 2). An additional six proteins were identified from the pH 7–11 separations (Table I, supplemental Tables 1 and 2, and supplemental Figs. 1 and 2). More proteins were observed as decreased versus increased in abundance in the carcinogenic versus the non-carcinogenic strain.

A C296R Mutation in the FlaA Protein Identified by Proteomics Analysis Results in Altered Bacterial Motility and Morphology—Identification of differentially regulated proteins in a carcinogenic H. pylori strain provides a foundation for targeted and rational investigations into the role of specific bacterial factors in virulence. Of particular interest in this regard were apparent changes in levels of the H. pylori flagellar protein FlaA (Fig. 3A), which is required for a phenotype (motility) implicated in pathogenesis. A detailed analysis indicated that FlaA levels did not differ between strains B128 and 7.13 but that each strain expressed a different series of four FlaA charged isoforms with the two series being slightly offset in pi (Fig. 3B). This pi shift for a specific series of FlaA charged isoforms is consistent with a new charged modification that is present on all isoforms or a missense mutation involving a charged amino acid. Because FlaA was found to resolve as at least four distinct charged isoforms, it is likely that this alteration would have been missed using conventional 2D gel analysis where the pi shift for this series of isoforms could easily be mistaken for gel-to-gel variation.

The source of this difference was investigated using MALDI-TOF MS and TOF/TOF tandem MS on gel-resolved FlaA isoforms that were run on separate gels (because the DIGE gels each contained a mixture of samples from both strains). Peptide mass mapping of the resulting in-gel digestion with trypsin protease clearly demonstrated a unique peptide ion at m/z 732.38 that was specific to FlaA from strain B128 (Fig. 4, A and B). The fragmentation pattern that resulted from tandem TOF/TOF MS on this peptide was consistent with the sequence GC296LNLR (Fig. 4C). Direct genomic sequencing of flaA in both strains confirmed the presence of a single T→C nucleotide substitution that specified a cysteine in strain B128 and an arginine in strain 7.13 at position 296 of the coding sequence (Fig. 5A).
Carcinogenic *H. pylori* Proteome

Proteins identified as present in differing amounts in membrane and cytosolic fractions of strains 7.13 and B128

<table>
<thead>
<tr>
<th>Gel position numbera</th>
<th>NCBI accession number</th>
<th>Gene numberb</th>
<th>Name</th>
<th>Protein</th>
<th>-Fold difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Membrane fraction: proteins lower in 7.13 vs. B128</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>gi</td>
<td>15644938</td>
<td>HP0310c</td>
<td>Hypothetical ORF</td>
<td>Putative protein</td>
<td>−1.87</td>
</tr>
<tr>
<td>14 (15)</td>
<td>gi</td>
<td>15645139</td>
<td>HP0512</td>
<td>glIA</td>
<td>Glutamine synthetase</td>
<td>−1.62</td>
</tr>
<tr>
<td>31</td>
<td>gi</td>
<td>25808199</td>
<td>HP0545</td>
<td>cag24/cagD</td>
<td>Cag pathogenicity island protein</td>
<td>−1.3</td>
</tr>
<tr>
<td>11</td>
<td>gi</td>
<td>7994574</td>
<td>HP0658</td>
<td>gatB</td>
<td>Aspartyl/glutamyl-tRNA amidotransferase</td>
<td>−3.58</td>
</tr>
<tr>
<td>18 (17)</td>
<td>gi</td>
<td>15345313</td>
<td>HP0690</td>
<td>fadA</td>
<td>Acetyl-CoA acetyltransferase</td>
<td>−1.69</td>
</tr>
<tr>
<td>38 (37, 39, 40)</td>
<td>gi</td>
<td>84626117</td>
<td>HP0875</td>
<td>katA</td>
<td>Catalase</td>
<td>−4.04</td>
</tr>
<tr>
<td>29</td>
<td>gi</td>
<td>4155406</td>
<td>HP0900</td>
<td>hypB</td>
<td>Urease nickel incorporation protein</td>
<td>−1.54</td>
</tr>
<tr>
<td>27</td>
<td>gi</td>
<td>15645574</td>
<td>HP0958</td>
<td>Hypothetical ORF</td>
<td>Putative</td>
<td>−2.93</td>
</tr>
<tr>
<td>16</td>
<td>gi</td>
<td>15645793</td>
<td>HP1179</td>
<td>deoB</td>
<td>Phosphopentomutase</td>
<td>−1.58</td>
</tr>
<tr>
<td>44</td>
<td>gi</td>
<td>15645899</td>
<td>HP1286</td>
<td>Hypothetical ORF</td>
<td>Conserved hypothetical secreted protein</td>
<td>−2.02</td>
</tr>
<tr>
<td>1 (2)</td>
<td>gi</td>
<td>15646062</td>
<td>HP1453</td>
<td>Hypothetical ORF</td>
<td>Putative outer membrane protein</td>
<td>−5.95</td>
</tr>
<tr>
<td>42</td>
<td>gi</td>
<td>15646171</td>
<td>HP1564</td>
<td>Outer membrane protein</td>
<td>Outer membrane protein</td>
<td>−2.2</td>
</tr>
<tr>
<td>36</td>
<td>gi</td>
<td>4155432</td>
<td>JHP0870</td>
<td>Hypothetical ORF</td>
<td>Hypothetical outer membrane protein</td>
<td>−2.86</td>
</tr>
<tr>
<td><strong>Membrane fraction: proteins higher in 7.13 vs. B128</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>gi</td>
<td>15644938</td>
<td>HP0310c</td>
<td>Hypothetical ORF</td>
<td>Putative protein</td>
<td>3.2</td>
</tr>
<tr>
<td>28</td>
<td>gi</td>
<td>15644959</td>
<td>HP0331</td>
<td>minD</td>
<td>Cell division inhibitor</td>
<td>1.33</td>
</tr>
<tr>
<td>20</td>
<td>gi</td>
<td>2920363</td>
<td>HP0743</td>
<td>mreB</td>
<td>Rod shape-determining protein</td>
<td>1.27</td>
</tr>
<tr>
<td>26</td>
<td>gi</td>
<td>15645444</td>
<td>HP0825</td>
<td>trxB</td>
<td>Thioredoxin reductase</td>
<td>3.63</td>
</tr>
<tr>
<td>9</td>
<td>gi</td>
<td>18075724</td>
<td>HP1134</td>
<td>atpA</td>
<td>ATP synthase α chain</td>
<td>1.65</td>
</tr>
<tr>
<td><strong>Cytosolic fraction: proteins lower in 7.13 vs. B128</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>gi</td>
<td>4154609</td>
<td>HP0110</td>
<td>grpE</td>
<td>Co-chaperone and heat shock protein</td>
<td>−1.24</td>
</tr>
<tr>
<td>35</td>
<td>gi</td>
<td>7188722</td>
<td>HP0243</td>
<td>napA</td>
<td>Neutrophil-activating protein</td>
<td>−1.33</td>
</tr>
<tr>
<td>21</td>
<td>gi</td>
<td>15644938</td>
<td>HP0310c</td>
<td>Hypothetical ORF</td>
<td>Putative; homology to polysaccharide deacetylase</td>
<td>−2</td>
</tr>
<tr>
<td>33</td>
<td>gi</td>
<td>15645018</td>
<td>HP0390</td>
<td>tagD</td>
<td>Adhesin-thioli peroxidase</td>
<td>−1.22</td>
</tr>
<tr>
<td>32</td>
<td>gi</td>
<td>15645019</td>
<td>HP0391</td>
<td>cheW</td>
<td>Chemotaxis protein</td>
<td>−1.24</td>
</tr>
<tr>
<td>14 (15)</td>
<td>gi</td>
<td>15645139</td>
<td>HP0512</td>
<td>glIA</td>
<td>Glutamine synthetase</td>
<td>−1.49</td>
</tr>
<tr>
<td>31</td>
<td>gi</td>
<td>25808199</td>
<td>HP0545</td>
<td>cag24/cagD</td>
<td>Cag pathogenicity island protein</td>
<td>−1.37</td>
</tr>
<tr>
<td>18 (17)</td>
<td>gi</td>
<td>15345313</td>
<td>HP0690</td>
<td>fadA</td>
<td>Acetyl-CoA acetyltransferase</td>
<td>−1.26</td>
</tr>
<tr>
<td>24</td>
<td>gi</td>
<td>15645444</td>
<td>HP0825</td>
<td>trxB</td>
<td>Thioredoxin reductase</td>
<td>−1.66</td>
</tr>
<tr>
<td>38 (37, 39, 40)</td>
<td>gi</td>
<td>84626117</td>
<td>HP0875</td>
<td>katA</td>
<td>Catalase</td>
<td>−2.32</td>
</tr>
<tr>
<td>29</td>
<td>gi</td>
<td>4155406</td>
<td>HP0900</td>
<td>hypB</td>
<td>Urease nickel incorporation protein</td>
<td>−1.73</td>
</tr>
<tr>
<td>27</td>
<td>gi</td>
<td>15645574</td>
<td>HP0958</td>
<td>Hypothetical ORF</td>
<td>Putative protein</td>
<td>−4.76</td>
</tr>
<tr>
<td>16</td>
<td>gi</td>
<td>15645793</td>
<td>HP1179</td>
<td>deoB</td>
<td>Phosphopentomutase</td>
<td>−1.31</td>
</tr>
<tr>
<td>44</td>
<td>gi</td>
<td>15645899</td>
<td>HP1286</td>
<td>Hypothetical ORF</td>
<td>Conserved hypothetical secreted protein</td>
<td>−1.59</td>
</tr>
<tr>
<td>42</td>
<td>gi</td>
<td>15646171</td>
<td>HP1564</td>
<td>Outer membrane protein</td>
<td>Outer membrane protein</td>
<td>−1.45</td>
</tr>
<tr>
<td>23</td>
<td>gi</td>
<td>4155056</td>
<td>JHP0504</td>
<td>accA</td>
<td>Acetyl-CoA carboxylase subunit A</td>
<td>−2.94</td>
</tr>
<tr>
<td><strong>Cytosolic fraction: proteins higher in 7.13 vs. B128</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>gi</td>
<td>15644859</td>
<td>HP0231</td>
<td>Hypothetical ORF</td>
<td>Putative protein</td>
<td>1.29</td>
</tr>
<tr>
<td>22</td>
<td>gi</td>
<td>15644938</td>
<td>HP0310c</td>
<td>Hypothetical ORF</td>
<td>Putative protein</td>
<td>18.8</td>
</tr>
<tr>
<td>28</td>
<td>gi</td>
<td>15644959</td>
<td>HP0331</td>
<td>minD</td>
<td>Cell division inhibitor</td>
<td>1.25</td>
</tr>
<tr>
<td>43</td>
<td>gi</td>
<td>15645898</td>
<td>HP1285</td>
<td>hypB</td>
<td>Acid phosphatase</td>
<td>1.35</td>
</tr>
<tr>
<td>19</td>
<td>gi</td>
<td>4154912</td>
<td>JHP0387</td>
<td>pepQ</td>
<td>Putative proline peptidase</td>
<td>1.32</td>
</tr>
</tbody>
</table>

a Corresponds to gel position number in supplemental Table 1 and supplemental Fig. 1. Gel position numbers in parentheses indicate redundant excisions from charged isoforms that exhibited the same expression pattern.

b Corresponds with gene number from the sequenced genome of *H. pylori* 26695 (HP number) or J99 (JHP number).

c The apparent abundance change in HP0310 was subsequently determined to be due to differences in isoforms between strains (see text).
Strains B128 and 7.13 Exhibit Differences in Motility and Ultrastructure—The C296R FlaA amino acid substitution identified by DIGE/MS localized to a predicted helical region. Because differences in flagellar structure can alter motility and motility is essential for colonization and persistence of *H. pylori* (42, 43), we next determined whether differences in motility were present between the two strains. Samples of both *H. pylori* strains were spotted onto soft agar plates and incubated for 3 days, and the resulting diameters of swarm halos were quantified. Strain 7.13 consistently migrated farthest, demonstrating that disruption of flagella abrogates the ability of *H. pylori* to colonize animals (44). Of interest, complementation of flaA null B128 background completely restored motility in soft agar (Fig. 6A) but only to levels seen with parental strain B128, suggesting that additional factors mediate the increased motility phenotype observed in cancer strain 7.13. We next examined the ability of these mutants to alter CagA translocation and β-catenin activation, phenotypes previously shown to differ between strains B128 and 7.13 (22). In contrast to motility, there were no differences in CagA translocation or β-catenin activation when either of the mutant flaA strains were compared with the wild-type strain (data not shown).

**Complementation of FlaA from H. pylori Strain 7.13 into Strain B128 per se Does Not Reconstitute an Oncogenic Phenotype in Vivo**—We next sought to define the role of FlaA in a model of *H. pylori*-induced carcinogenesis by infecting Mongolian gerbils with *Brucella* broth alone (*n* = 10), strain 7.13 (*n* = 10), an isogenic flaA null strain (*n* = 10), or the complemented B128 flaA null mutant (*n* = 10). Gerbils were challenged with *H. pylori* and sacrificed 12–14 weeks postinoculation. There was no evidence of gastric inflammation or injury in uninfected control animals. The majority of gerbils (90%) challenged with strain 7.13 were successfully infected and developed acute and chronic gastric inflammation as well as dysplasia as described previously (22). As expected, inactivation of flaA in strain B128 led to an inability of this mutant to colonize gerbils, which is consistent with previous reports demonstrating that disruption of flagella abrogates the ability of *H. pylori* to colonize animals (44). Of interest, complementation of flaA from strain 7.13 into the isogenic flaA null mutant B128 strain did not restore the ability to colonize gerbils or induce injury. These results indicate that complementation of FlaA from strain 7.13 into strain B128 alone is not sufficient to induce lesions with premalignant potential within the gastric niche.

**Additional Proteins Differentially Abundant in Strains B128 and 7.13**—In light of the above results, we next sought to...
investigate the roles of several additional loci of pathogenic interest identified in the DIGE/MS analysis in CagA translocation and β-catenin activation. These loci included HP0743 (up-regulated in strain 7.13), which is predicted to encode a rod shape-determining protein; HP0391 and HP1286 (each up-regulated in strain B128), which are predicted to encode proteins that regulate chemotaxis or that possess hypothetical functions, respectively; and HP0310, an N-acetylglucosamine deacetylase (45, 46). These loci were selected based upon differences in bacterial morphology that we observed between strains B128 and 7.13 (Fig. 5C), our interest to determine whether certain microbial proteins may exert suppressive roles in H. pylori-induced cancer, and recent data indicating that HP0310 may function to modulate the murein layer of H. pylori by deacetylating peptidoglycan, a substrate that is translocated by the cag secretion system, and indicate that multiple H. pylori disease-associated constituents can alter CagA translocation.

Our global survey by DIGE/MS also revealed several other differentially abundant proteins of interest. These include AtpA and NapA (Table I). Compared with B128, membrane fractions from strain 7.13 expressed significantly more AtpA (1.65-fold increase; \( p = 0.019 \)), a protein homologous to the ATP synthase unit F1 subunit, that plays a key role in the synthesis of ATP (49). NapA, a neutrophil-activating protein, was present at significantly lower levels in strain 7.13 versus B128 (1.3-fold lower in 7.13 cytoplasm; \( p = 0.009 \)). Thus,DIGE/MS identified several additional candidate loci that can be targeted in future studies.
DISCUSSION

H. pylori isolates obtained from different individuals exhibit substantial genetic diversity (11–15) consistent with extensive recombination and a panmictic population structure. Differences among strains include point mutations in highly conserved genes, the presence of nonconserved and/or mosaic forms of genes, and chromosomal organization (11–15). Because of this remarkable genetic diversity, identification of H. pylori factors associated with virulence has been difficult; this prompted our focus on two genetically related prototype strains that differ in carcinogenic potential. By performing 2D DIGE of both cytoplasmic and membrane extracts from four independent experiments across a series of gels with an internal standard, we were able to determine statistical significance for the abundance differences identified, and PCA indicated ~5% variance between the non-carcinogenic strain B128 and the carcinogenic strain 7.13 (Fig. 2B). Unsupervised hierarchical clustering analysis confirmed reproducibility between replicates with minimal technical variation, further underscoring the utility of this method for the analysis of closely related samples. Our results comparing these two prototype strains identified a limited number of differentially represented proteins that we can test in future functional assays using an easily transformable carcinogenic H. pylori strain (7.13) and a susceptible host (Mongolian gerbils). These findings must ultimately be extended into larger populations of H. pylori strains that have been isolated from persons with or without gastric cancer to rigorously confirm the importance of candidate loci identified by our DIGE/MS analysis in the carcinogenic process. However, our current results have provided a foundation for these future corroborative studies, which will bring important new insights into mechanisms that mediate H. pylori-induced gastric carcinogenesis.

One of the most important differences detected was expression of H. pylori strain-specific FlaA isoforms. This analysis demonstrated that FlaA expressed by strain B128 possessed a cysteine at residue 296, whereas an arginine is normally specified at this position in strain 7.13 as well as other strains. Consistent with an arginine present at this position in the carcinogenic strain, the overall pI change for the four major FlaA isoforms is shifted to the right (more basic) in strain 7.13 as expected (Fig. 3B), and DNA sequencing confirmed the strain-specific presence of cysteine or arginine as expected. It is important to note in this regard that this missense mutation may have been overlooked if the complementary peptide-based ESI-LC/MS/MS (“shotgun”) proteomics analysis was used unless alterations in FlaA were specifically targeted and searched. Further analysis into phenotypic differences related to this substitution revealed variations in flagellar adhesive properties in conjunction with morphological differences (Fig. 5C), which are consistent with the disparity in motility observed between these two strains (Fig. 5B). Complementation of flaA from strain 7.13 into the isogenic flaA^- mutant B128 strain restored motility in vitro; however, this did not restore the ability of the isogenic flaA^- mutant to colonize gerbils for which there are several potential explanations.

One possibility is that H. pylori flagellin exerts other functions in vivo in addition to regulating motility. For example,
Carcinogenic *H. pylori* Proteome

Reeves et al. (50) recently reported that inactivation of *flaA* significantly reduced binding of *H. pylori* to trefoil factor 1 (TFF1), a cysteine-rich protein found within gastric mucus. In that study, disruption of *flaA* also altered the production of *H. pylori* LPS as wild-type *H. pylori* was able to express both rough- and smooth-type LPS, but inactivation of *flaA* abolished the ability to express rough-type LPS (50). Flagella from mucosal Gram-negative pathogens such as *Salmonella* can induce other phenotypes, such as production of proinflammatory cytokines (51). Thus, it is possible that complementation of *flaA* from strain 7.13 into strain B128 may have only partially restored functions attributable to *H. pylori* flagellin.

Unlike most prokaryotes, *H. pylori* lacks the requisite enzymes to generate glutathione and other thiol reductants. However, this pathogen does possess a thioredoxin system, including TrxB, that catalyzes the reduction of specific proteins by NADPH (52), which is critical for survival of *H. pylori* in an environment containing DNA-damaging agents (53). Infection with *H. pylori* strain 7.13 induces a severe inflammatory response in gerbils (22) that is associated with increased levels of reactive oxygen and nitrogen radicals at sites juxtaposed to colonizing organisms. Proteomics analysis demonstrated that the membrane, but not the cytoplasm, of strain 7.13 contained higher levels of TrxB than did the corresponding fractions in strain B128. Increased expression of TrxB and altered localization of thioredoxin to the membrane versus the cytoplasm places this defense molecule in an optimal position to scavenge extracellular oxygen and nitrogen radicals, potentially protecting strain 7.13 from lethal mutagenesis and enhancing its survival within inflamed mucosa. The neutrophil-activating protein NapA was also differentially abundant in strain B128 versus 7.13. Because NapA plays an important role in recruiting human neutrophils and monocytes to sites of infection (54), down-regulation of this protein may facilitate immune evasion, thus conferring a selective advantage to strain 7.13 for inducing pathogenic responses.

Another candidate virulence protein, HP0310, was identified to be present at higher levels in both the membrane and cytosolic fractions of strain 7.13 versus strain B128. Upon detailed inspection, however, these differences were found to result from the presence of different charged isoforms rather than an abundance change between the two strains. This difference was confirmed by genomic sequencing of HP0310 in both B128 and 7.13 that demonstrated a single point mutation that converted a cysteine at amino acid position 34 in HP0310 into an arginine in strain 7.13. The protein encoded by the HP0310 gene is homologous to polysaccharide deacetylase, an enzyme that catalyzes the hydrolysis of N-linked acetylglucosamine residues or O-linked acetylglucosamine residues in *H. pylori* peptidoglycan (45), a substrate for the cag secretion system. Our current results indicate that loss of HP0310 augments CagA translocation, which may play a role in pathogenesis. Of interest, HP0310 was abundant in both cytoplasmic and membrane fractions of strain...
7.13, suggesting that this protein may function not only to modify bacterial proteins and residues such as peptidoglycan prior to translocation into host cells but also to modify surface proteins present on the surfaces of neighboring epithelial cells.

In conclusion, H. pylori has evolved numerous strategies to facilitate its persistence within the human stomach, and one consequence of long term colonization is an increased risk of gastric cancer. By utilizing proteomics technology to contrast consequence of long term colonization is an increased risk of facilitate its persistence within the human stomach, and one optimize prevention of subsequent neoplastic events.

Such results would ultimately permit physicians to more appropriately focus diagnostic and eradication strategies on high risk populations to eventually permit physicians to more appropriately focus diagnostic and eradication strategies on high risk populations to approximately permit physicians to more appropriately focus diagnostic and eradication strategies on high risk populations to facilitate its persistence within the human stomach, and one optimize prevention of subsequent neoplastic events.

H. pylori-induced carcinogenesis on a deeper level. Such results would ultimately permit physicians to more appropriately focus diagnostic and eradication strategies on high risk populations to facilitate its persistence within the human stomach, and one optimize prevention of subsequent neoplastic events.

References


* This work was supported, in whole or in part, by National Institutes of Health Grants CA116087, DK58587, and CA77955 (to R. M. P.) and DK058404 (to the Vanderbilt Digestive Diseases Research Center). This work was also supported by the Vanderbilt Academic Venture Capital Fund.

†‡ To whom correspondence should be addressed: Division of Gastroenterology, Vanderbilt University School of Medicine, 1030C MRB IV, 2215B Garland Ave., Nashville, TN 37232-2279; Tel.: 615-322-5200; Fax: 615-343-6229; E-mail: richard.peek@vanderbilt.edu.

†‡‡ The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.


