GASTROENTEROLOGY 2002;123:1637–1648

Global Analysis of *Helicobacter pylori* Gene Expression in Human Gastric Mucosa

JAMES E. GRAHAM,* RICHARD M. PEEK, Jr,*,† UMA KRISHNA,* and TIMOTHY L. COVER*,†,§

*Department of Medicine, Vanderbilt University School of Medicine; †Veterans Affairs Medical Center; and §Department of Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, Tennessee

Background & Aims: *Helicobacter pylori* inhabits a highly restricted ecological niche in the human gastric mucosa. Microbial gene expression in the context of persistent infection remains largely uncharacterized. Methods: An RNA analysis method, selective capture of transcribed sequences, was used in conjunction with genomic array hybridization to characterize *H. pylori* complementary DNAs (cDNAs) obtained from both human and experimentally infected gerbil gastric tissue specimens. Results: Bacterial cDNAs obtained by selective capture of transcribed sequences from tissues hybridized to arrayed DNA fragments representing approximately 70% of open reading frames in the *H. pylori* genome. RNAs for most of these open reading frames were also detected by array hybridization analyses of total RNA prepared from the isolated *H. pylori* strains cultured in vitro. However, a subset of *H. pylori* RNAs detected in gastric tissue specimens was consistently undetectable in bacteria grown in vitro. The majority of these RNAs encode factors unique to *H. pylori* that are potentially produced in response to interactions with mammalian gastric mucosa. Conclusions: The combination of selective capture of transcribed sequences with array hybridization has allowed a global analysis of bacterial gene expression occurring in human tissues during a natural infection.

Helicobacter pylori is a Gram-negative spiral-shaped bacterium found almost exclusively in the human gastric mucosa. *H. pylori* infection is widespread, and in most developing countries, the majority of the population is infected. Although colonization by *H. pylori* consistently results in gastric inflammation, most infected individuals do not develop any symptoms or clinically apparent adverse affects. However, the presence of *H. pylori* and coexisting gastritis is associated with an increased risk for development of peptic ulcer disease, distal gastric adenocarcinoma, and gastric lymphoma. At present, the bacterial determinants contributing to gastritis or these more severe forms of disease remain incompletely characterized.[1,2]

The acidity of the mammalian stomach is an effective barrier to colonization by essentially all bacteria other than *Helicobacter* species. The unique ability of *H. pylori* to colonize and persist for decades within the human gastric mucosa, despite development of a mucosal inflammatory and immune response, is of considerable interest. Analyses of the complete genome sequences of 2 different *H. pylori* isolates have provided important insights into the metabolic capacities and physiological characteristics that underlie this remarkable ability. However, relatively little is known about regulation of gene expression in *H. pylori*, particularly within the context of host infection.

Bacteria adapt to different environments by regulating the expression of genes in response to specific environmental signals. For bacterial pathogens, complex patterns of environmental cues signal changes in gene expression during the progressive stages of host infection.[3–5] Powerful genetic methods for the identification of these genes are now available[6] and include in vivo expression technologies[7,8] and other promoter trapping methods.[9] Application of these approaches to the study of human infectious diseases is limited by a requirement for disease models in which microbes encounter specific local environments and host factors similar to those present in naturally occurring human infections. This becomes a primary concern when suitable cell culture and animal models are unavailable or fail to show important features of human infections.[10]

One approach to the identification of bacterial factors produced during human infections involves examination of bacterial RNA isolated from naturally infected human tissues. RNA-based analysis of bacterial gene expression

Abbreviations used in this paper: DNase, deoxyribonuclease I; OD, optical density; ORF, open reading frame; PCR, polymerase chain reaction; RT, reverse transcriptase; SCOTS, selective capture of transcribed sequences.

© 2002 by the American Gastroenterological Association
0016-5085/02/$35.00
doi:10.1053/gast.2002.36589

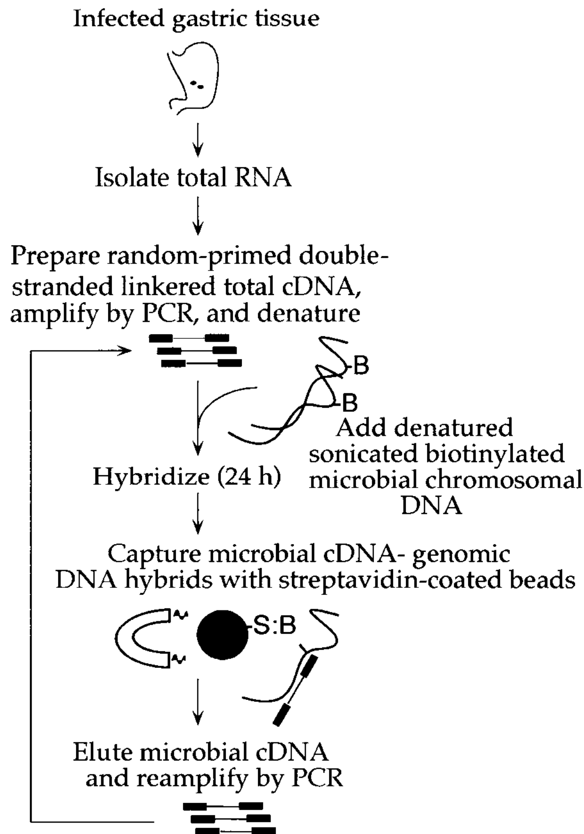

Figure 1. Outline of the basic method used for amplification of *H. pylori* cDNA from gastric tissue specimens. Total RNA was isolated from infected gastric tissue and converted to double-stranded cDNA containing defined terminal sequences. After amplification, *H. pylori* cDNA was selectively captured by hybridization to sonicated biotinylated *H. pylori* genomic DNA fragments. The cDNA/genomic DNA hybrids were then obtained by binding to streptavidin-coated beads, and total bacterial cDNA was eluted and amplified by PCR. Three rounds of capture hybridization and amplification were then used to prepare cDNA probes for array hybridization. S, streptavidin; B, biotin.

is technically challenging because of the general instability and limited polyadenylation of prokaryotic messenger RNA (mRNA). In naturally occurring infections, these challenges are often compounded by low numbers of bacteria present and by a requirement for appropriately preserved human tissue specimens. A complementary DNA (cDNA) analysis method designated SCOTS (for selective capture of transcribed sequences) was recently designed to overcome many of these obstacles,[11] successfully identifying several *Mycobacterium tuberculosis* RNAs expressed in response to phagocytosis by cultured primary human macrophages. SCOTS uses hybridization to biotinylated bacterial genomic DNA to obtain the bacterial component of total cDNA prepared directly from infected tissues (Figure 1).

In this study, we sought to characterize *H. pylori* gene expression occurring during natural colonization of the human stomach. We report that whole-genome array hybridization with cDNAs obtained by SCOTS from single gastric biopsy specimens detected expression of RNAs corresponding to more than 1100 *H. pylori* genomic open reading frames (ORFs). By comparing these RNA expression patterns with those of the isolated *H. pylori* strains growing in laboratory cultures, we identified a group of largely *H. pylori*–specific genes, including those within putative species-specific operons, that are potentially expressed in response to environmental signals and host factors encountered in *H. pylori*–infected gastric mucosa.

Materials and Methods

Bacterial Cultures

H. pylori strains B128 and B213 were isolated from human gastric biopsy specimens as previously described.[12] Both strains were $cagA^+$ and contained type s1a/m2 and s1b/m1 *vacA* alleles, respectively, as determined by previously described methods.[13,14] Strain B128 has also recently been shown to possess an intact *cag* pathogenicity island and is capable of reproducibly colonizing Mongolian gerbils and inducing a robust inflammatory response.[15] *H. pylori* were initially grown on Trypticase Soy Agar (Difco, Detroit, MI) plates containing 5% sheep's blood agar (BA) incubated at 37°C in a 5% CO_2 incubator for 96 hours before being frozen as seed stocks for in vitro cultures. For analysis of in vitro RNA expression, 3 mL of sulfite-free brucella broth (containing 10% fetal bovine serum and 15 μg/mL of vancomycin) was inoculated to an optical density (OD) of 0.1 (OD at 600 nm) with *H. pylori* collected in this media after 24 hours of growth on BA plates. Cultures were grown at 37°C for 16 hours with shaking in a 5% CO_2 incubator before transfer into 30 mL of brucella broth in a 300-mL baffled Erlenmeyer flask (starting OD of 0.1). When culture OD reached approximately 0.6, cells were reinoculated into 300-mL flasks as described previously. *H. pylori* cultures reached the midlogarithmic phase (OD of 0.3) after approximately 8 hours (doubling time of 5–7 hours), at which time bacteria were collected for RNA isolation.

Isolation of RNA From *H. pylori* Growing In Vitro

H. pylori cultures were decanted into 2 chilled 50-mL Oak Ridge centrifuge tubes on wet ice before centrifugation at 4°C and 1000 × *g*. Bacterial pellets were briefly resuspended in 3 mL of ice-cold phosphate-buffered saline containing $MgCl_2$ and $CaCl_2$ (10 mmol/L) on ice, centrifuged, and resuspended in 3 mL of ice-cold 10 mmol/L EDTA containing 0.3% sodium dodecyl sulfate. RNA was extracted with hot phenol as previously described[16] and was precipitated twice with isopropanol. Resuspended nucleic acids were then treated with deoxyribonuclease I (DNase) as described by the manu-

facturer (Ambion, Austin, TX). DNase was heat-inactivated (75°C for 5 minutes after the addition of EDTA to 10 mmol/L), and RNAs were stored at −70°C as ethanol precipitates.

Characteristics of Gastric Biopsy Specimens B128C and B213A

Biopsies of gastric tissue (~5 mg) were obtained from 2 consenting patients as part of a study approved by the Health Sciences Committee of the Vanderbilt University Institutional Review Board. Patient B128 was a 55-year-old white man in whom a gastric ulcer was visualized during upper gastrointestinal tract endoscopy. Patient B213 was a 66-year-old black man in whom erosive esophagitis was visualized during upper gastrointestinal tract endoscopy. An average of 16 *H. pylori* per high-power field were present in tissue sections from patient B128, and, on the basis of previous studies, this density corresponds to approximately 5×10^3 to 1×10^4 cultivatable *H. pylori* per 5-mg biopsy specimen.[17] An average of 70 *H. pylori* per high-power field were present in tissue sections from patient B213. Biopsy specimens from both patients were immediately frozen at −70°C for analysis of *H. pylori* in vivo gene expression as described below.

Experimental Infection of Mongolian Gerbils

With the approval of the Institutional Animal Care Committee of Vanderbilt University, a 6-week-old Mongolian gerbil was infected via oral inoculation with *H. pylori* B128 as previously described.[15] Two weeks after inoculation, the stomach was removed from the dead animal; one half was fixed in 10% formalin for histological examination, one fourth was used for quantitative culture,[15] and one fourth was quick-frozen at −70°C for RNA analysis. Approximately 50 mg of frozen tissue was later processed as described below to analyze bacterial RNA content by SCOTS. Quantitative culture indicated that approximately 50 viable bacteria were present per milligram of infected tissue.

Isolation of Total RNA From *H. pylori*–Infected Tissues

Gastric tissue specimens were added directly to 0.5 mL of heated (60°C) cell lysis solution (1% sodium dodecyl sulfate, 0.5 mmol/L EDTA, and 0.1 mol/L β-mercaptoethanol) and triturated repeatedly by micropipette to lyse exposed adherent bacteria before transfer to a QIAshredder (Qiagen, Valencia, CA) tissue homogenizing spin column (B128C) or microcentrifuge tube (B213A and gerbil specimen) for grinding with a fitted pestle (Kimble-Kontes, Vineland, NJ) on ice. The resulting lysates were extracted with an equal volume of phenol containing 0.1 mol/L sodium acetate (pH 4.0) and then with a phenol/chloroform mixture (5:1). After an additional extraction with a chloroform/isoamyl alcohol mixture (24:1), total nucleic acids were precipitated from each aqueous phase by the addition of an equal volume of isopropanol. Precipitates were resuspended, precipitated again, and treated with DNase as described previously.

Selective Capture of Transcribed Sequences

Each of the RNA preparations from gastric tissue was converted to double-stranded cDNA containing defined terminal sequences (5′-GACACTCTCGAGACATCACCGG-TACC-3′), as previously described.[11,18] Three replicate polymerase chain reaction (PCR) amplifications of each cDNA were then performed in parallel, and replicate PCR products were pooled before selective amplification of *H. pylori* cDNAs by SCOTS. Total tissue-derived random-primed cDNA fragments containing defined terminal sequences (prepared as described previously) were denatured, and bacterial cDNAs selected by hybridization to denatured sonicated photobiotinylated *H. pylori* genomic DNA fragments from the relevant strain, as previously described.[11] Microbial cDNA/genomic DNA hybrids were then collected by binding to streptavidin-coated beads, and bacterial cDNAs were eluted by alkaline denaturation (Figure 1). Eluted bacterial cDNAs were then amplified by PCR. Noncompetitive PCR amplification[19] of short random-primed DNA fragments with a single primer has previously been shown to yield a generally unbiased population of amplicons, even when applied to far more complex nucleic acid pools.[20] Three rounds of selective capture of *H. pylori* cDNAs and PCR amplification were used for each tissue specimen. The cDNAs were then used as templates to prepare radiolabeled probes used in hybridizations with *H. pylori* genomic arrays on nylon.

A previously described differential hybridization strategy was then used in an effort to enrich bacterial cDNA obtained from biopsy B128C for sequences not obtained by SCOTS from bacteria growing in vitro.[11] The resulting cDNA was then used as a template for synthesis of array hybridization probes (see below) and was also ligated into the *E. coli* plasmid vector pTOPO-TA (InVitrogen, Carlsbad, CA). Sequences were determined for 95 of these plasmid clones by using standard dye-terminator methodology at the Vanderbilt–Ingram Cancer Center DNA Sequencing Laboratory. BLAST[21] was used to search GenBank and *H. pylori* databases for previously described *H. pylori* ORFs corresponding to these cDNAs and to identify similarities between the encoded peptide sequences and those reported for other organisms.

Analysis of *H. pylori* In Vivo Gene Expression by Array Hybridization

H. pylori genome arrays on nylon (Sigma-Genosys, The Woodlands, TX) were used to analyze the composition of *H. pylori* cDNAs obtained by SCOTS from human gastric biopsy samples B128C and B213A and the B128-infected gerbil gastric tissue specimen. These arrays consist of PCR-amplified DNA fragments corresponding to the largest sequence for each of the 1590 ORFs of the *H. pylori* 26695 genome for which a suitably unique genomic region could be identified,[22] as well as unique regions for an additional 91 ORFs that are present only in the *H. pylori* J99 genome.[23]

Radiolabeled cDNA probes were prepared from 0.5–2 μg of amplified cDNA template by incorporation of $\alpha P^{[33]}$deoxyadenosine triphosphate and Prime-It (Stratagene, La Jolla, CA) reagents according to manufacturer's instructions, with the modification of including cold deoxyadenosine triphosphate at 1:100 the level of the other 3 deoxyribonucleotides. (We retain the traditional usage of the term *probe* for these radiolabeled cDNAs used in array hybridization throughout this article.) After removal of unincorporated nucleotides by Sephadex G50 column chromatography, probes were denatured by the addition of NaOH to 0.2 mol/L, and approximately 5×10^6 counts per minute (50 μL) were added to roller bottles containing 6 mL of hybridization solution. Nylon arrays were placed in bottles as described by Fawcett et al.[24] and prehybridized for 2 hours. After 24 hours of hybridization, nylon membranes were washed, covered in a single sheet of plastic film (Saran Wrap; 3M), and exposed for 24 hours to a high-resolution phosphorimager screen (Fuji BAS FLA-2000; Fuji Scientific Imaging).

Semiquantitative Analysis of *H. pylori* Gene Expression During In Vitro Culture

RNA was isolated from *H. pylori* growing in vitro as described previously. Radiolabeled first-strand cDNA probes for array hybridization were prepared by primer extension with a modified murine Maloney leukemia virus reverse transcriptase (RT) (Superscript II; Gibco) by using 3 μg of random nonamer primers, 10–25 μg of template RNA, and 125 pmol of cold carrier deoxynucleotide with 50 μCi of $P^{[33]}$ radiolabeled deoxynucleotide, as described by the manufacturer. Template RNA was then hydrolyzed by the addition of NaOH to 0.4 mol/L and incubation at 55°C for 15 minutes before purification of high-molecular-weight radiolabeled probe by Sephadex G50 column chromatography. Hybridization of *H. pylori* genome arrays on nylon membranes and imaging of the resulting patterns were performed as described previously.

Quantification of Array Hybridization Signals

Hybridization patterns obtained as described previously were analyzed with Array Vision software (Imaging Research Inc., Ontario, Canada). The average pixel intensity for each spot on each array was determined, resulting in numerical data on relative hybridization signal intensity. Background levels for each hybridization were determined as the mean intensity of 45 blank features positioned throughout that array. Signals that were greater than 2-fold higher than background levels were defined as significant hybridization for that array and were considered detectable expression of the corresponding ORFs. To identify array features showing differential hybridization, we selected those features showing hybridization signals at least 5-fold over background with bacterial cDNA from tissue specimens that produced signals less than 2-fold over background with cDNA probes from bacteria grown in vitro. Visual inspection was then used to confirm differential hybridization of specific array features (Tables 1 and 2).

RT-PCR Analyses

RT-PCR was used as previously described by Peek et al.[25] to amplify target regions within selected *H. pylori* ORFs. Total RNA was isolated from 6 gastric biopsy specimens obtained from consenting patients. Each RNA was then treated with DNase and reverse-transcribed with random hexamer primers before use as a template for amplification of regions within *H. pylori* ORFs. For biopsy specimens B128C and B213A and the B128-infected gerbil tissue specimen, cDNAs obtained by SCOTS were used as templates for amplification of target regions. *H. pylori* gene-specific deoxyoligonucleotide primers used were 5′CAGTGATGTATGGCTTTGTGAACG and 5′AAACCCCTTGCTCAATGCAATGCTCCC for HP0228, 5′AAAGCGGTGCGGTTGTTAGC and 5′GCTGGGCGTAAATAAGGTTGAGTG for JHP0945, 5′GAAAACATACCACCAAAGCC and 5′TCGCCGCTGATTCTAATGC for HP0061, 5′GAGGAAATGCTTCGCCAAGTG and 5′AAAAAACAAACCCCCACATCATC for HP0343, 5′AGCCATACCCAACCTTCG and 5′TATCAATCCCCTTATCGC for HP0667, and 5′CCTTTTTATGAATCCTTTCTCTTAG and 5′AATACTGATGGCTTTATGCTG for HP0833.

Results

Global Analysis of *H. pylori* Transcription During Growth In Vitro

Before analysis of *H. pylori* gene expression in human gastric mucosa, patterns of transcription in the corresponding bacterial isolates were analyzed during growth in vitro. Radiolabeled first-strand cDNA probes were prepared from total RNA isolated from *H. pylori* B128 midlogarithmic phase cultures and hybridized to *H. pylori* genomic arrays. Despite the presence of rRNA, probes were capable of producing hybridization signals significantly above background for most array features (Figure 2). On the basis of analyses of RNA preparations from 3 independent cultures, a group of 40 array features consistently showed very strong hybridization signals (greater than 2 SD above the mean signal intensity; Table 3). This group included genes encoding several of the most abundant proteins previously identified in an *H. pylori* proteome analysis[26] (UreA, GroEL, GroES, TsaA, KatA, and TufB), as well as others for which high-level expression has not been previously described. These analyses also identified a set of 488 array features that failed to show significant hybridization, including 73 for genes previously shown to be absent from the *H. pylori* B128 genome.[15]

For comparison, total RNA was prepared from an *H. pylori* strain B213 midlogarithmic phase culture, and

Table 1. *H. pylori* Genomic ORFs Corresponding to Genes Potentially Expressed in Response to Interactions With Mammalian Gastric Mucosa[a]

Array features for 21 ORFs showing differential hybridization with cDNA obtained from *H. pylori* B128 in human gastric mucosa and cDNA from strain B128 cultured in vitro					
HP0061[b]	HP0315[c]	HP0441[c]	HP0718	HP1127	JHP0945[d]
HP0155[c]	HP0343[b]	HP0645[c]	HP0833[d]	HP1349[c]	
HP0228[b]	HP0347	HP0661	HP0966	HP1421	
HP0274	HP0358	HP0667[b]	HP0999[c]	HP1545	
Array features for 78 ORFs showing differential hybridization with cDNA obtained from *H. pylori* B128 in gerbil gastric mucosa and cDNA from strain B128 cultured in vitro					
HP0050	HP0315[c]	HP0667	HP1025	HP1365	JHP0827
HP0061[b]	HP0340	HP0684	HP1141	HP1433	JHP0829
HP0114	HP0343[b]	HP0685	HP1144	HP1475	JHP0917
HP0119	HP0370	HP0725	HP1176	HP1477	JHP0921
HP0120[c]	HP0441[c]	HP0765	HP1187	HP1528	JHP0936
HP0146	HP0454	HP0767	HP1188[e]	HP1536	JHP0940
HP0155[c]	HP0476	HP0780	HP1250	JHP0044	JHP0956
HP0167	HP0489	HP0800	HP1258	JHP0164	JHP0957
HP0168	HP0562	HP0882	HP1259	JHP0165	JHP0958[e]
HP0188	HP0579	HP0972	HP1264	JHP0318	JHP0959
HP0228[b]	HP0592	HP0999[c]	HP1288	JHP0616	JHP0960
HP0245	HP0602	HP1002	HP1349[c]	JHP0756	JHP1306
HP0291	HP0645[c]	HP1015[c]	HP1358	JHP0814	JHP1462
Array features for 44 ORFs showing differential hybridization with cDNA obtained from *H. pylori* B213 in human gastric mucosa and cDNA from strain B213 cultured in vitro					
HP0015	HP0228[b]	HP0463	HP0812	HP1015[c]	JHP0045
HP0038	HP0256	HP0505	HP0833[d]	HP1096	JHP0945[d]
HP0040	HP0263	HP0583	HP0879	HP1146	JHP0958[e]
HP0048	HP0321	HP0667[b]	HP0893	HP1188[e]	JHP1132
HP0061[b]	HP0338	HP0689	HP0895	HP1207	
HP0113	HP0343[b]	HP0696	HP0901	HP1289	
HP0120[e]	HP0435	HP0732	HP0981	HP1389	
HP0161	HP0437	HP0801	HP1005	HP1492	

[a]ORFs for which array features showed hybridization with *H. pylori* cDNAs obtained from tissue specimens by SCOTS, but not first-strand cDNAs from the corresponding strains grown in laboratory cultures. ORF numbers are those of Tomb et al.[31] and Alm et al.[23]
[b]ORFs for array features that showed hybridization with cDNA obtained by SCOTS from all gastric tissue specimens analyzed, but not cDNA from the isolates cultured in broth.
[c]Array features that showed hybridization with bacterial cDNA obtained from both *H. pylori* B128-infected tissue specimens (human and gerbil), but not cDNA from strain B128 at midlogarithmic phase in broth cultures.
[d]Array features that showed hybridization with bacterial cDNA obtained from two human gastric tissue specimens, but not from the corresponding *H. pylori* strains cultured in vitro.
[e]ORFs for which array features showed hybridization with bacterial cDNAs obtained from both strain B213-infected human tissue and strain B128-infected gerbil tissue, but not the corresponding strains cultured in vitro.

first-strand cDNA was used as an array hybridization probe (not shown). A similar number of array features ($n = 49$) were identified with hybridization signals greater than 2 SD above the mean intensity, including many of those identified in the analysis of strain B128 (Table 3). Four hundred ten array features failed to show significant hybridization; as for strain B128, a subset of these likely represent ORFs absent from the genome of strain B213. A total of 244 ORFs previously shown to be present in strain B128[15] did not show hybridization to cDNA from either strain B128 or B213 grown in vitro. In summary, these experiments identified a small group of transcripts that are among the most abundant mRNAs in both strains, as well as a group of transcripts that are either absent or present at levels below the threshold for detection in these strains when grown under standardized culture conditions.

Genomic Array Analysis of cDNA Obtained From *H. pylori* Growing In Vivo

To examine *H. pylori* gene expression in the human stomach, we first used SCOTS and genomic array hybridization to analyze bacterial RNA present in a gastric tissue specimen obtained from the B128 source patient. Bacterial cDNA obtained by SCOTS (Figure 1) from biopsy B128C was used as a template to prepare a radiolabeled probe for array hybridization. In replicate experiments, bacterial cDNA probes derived from B128C reproducibly hybridized to genomic DNA fragments corresponding to 329 *H. pylori* ORFs (Figure 3*A*).

Table 2. *H. pylori* Genomic ORFs for RNAs Identified in Multiple Tissue Analyses as Expressed by Bacteria in Mammalian Gastric Mucosa but Not Laboratory Cultures[a]

ORF	Prevalence[b]	Predicted product mass (*daltons*)	Putative identification
HP0061	100	20,790	Hypothetical protein
HP0228	100	42,900	Conserved hypothetical protein, probable sulfate permease
HP0343	80	15,400	Hypothetical protein
HP0667	87	9350	Hypothetical protein
HP0120	100	43,780	Hypothetical protein
HP0315	64	10,340	Virulence-associated protein D (*vapD*)
HP0441	7	94,380	VirB4 homolog
HP0645	100	61,600	Soluble lytic murein transglycosylase (slt)
HP0833	100	32,120	Hypothetical protein
HP0999	53	6700	Hypothetical protein
HP1015	100	18,480	Hypothetical protein
HP1188	100	29,590	Hypothetical protein
JHP0945	60	7,2390	Hypothetical protein
JHP0958	53	5,600	Hypothetical protein

[a]*H. pylori* genomic array features showing hybridization to cDNA obtained by SCOTS from gastric tissues, but hybridization levels not significantly above background with first-strand cDNA probes from bacteria grown to exponential phase in modified brucella broth. Annotations are those of Tomb et al.[31] and Alm et al,[23] as updated in the GeneQuiz database.[28] The first 4 genomic regions showed hybridization signals with bacterial cDNA from all *H. pylori* tissue specimens examined, and the remainder showed hybridization with bacterial cDNA from 2 of these 3 specimens, as indicated in Table 1.

[b]Percentage of 15 *H. pylori* strains reported to contain this ORF.[22]

Comparison of this pattern with that obtained from triplicate array analyses of RNA obtained from strain B128 grown in vitro identified 21 array features that reproducibly hybridized to SCOTS-derived cDNA probes from biopsy B128C, but not to cDNA from broth-grown *H. pylori* B128 (Table 1).

To extend our analyses of in vivo bacterial gene expression to an animal model of *H. pylori*–induced gastritis and disease, we next used SCOTS to prepare an array hybridization probe from a gastric tissue specimen obtained from a Mongolian gerbil experimentally infected with strain B128. This probe produced a more complex hybridization pattern than that derived from the B128-infected human tissue specimen (data not shown), detecting transcription of 1175 *H. pylori* genomic ORFs. Comparison of this pattern with those previously obtained from analyses of strain B128 grown in vitro resulted in the identification of 78 differentially hybridizing array features (Table 1). These included 10 of the 21 features (48%) originally identified as hybridizing to bacterial cDNA obtained from human biopsy B128C, but not to cDNA prepared from the cultured bacterial isolate (Table 1).

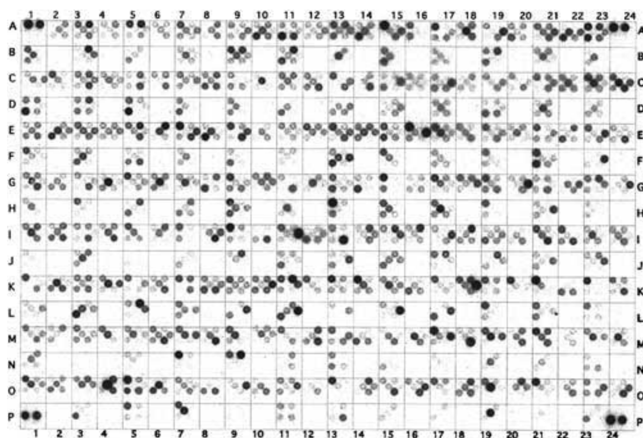

Figure 2. *H. pylori* genomic array hybridized with first-strand cDNA prepared from cultured *H. pylori* B128. Total RNA was extracted from bacteria grown to midlogarithmic phase in sulfite-free brucella broth and used to prepare radiolabeled cDNA probes for array hybridization. A representative array with a single feature for each predicted *H. pylori* genomic ORF is shown.

To further extend our investigation of *H. pylori* in vivo gene expression, we analyzed a second human gastric tissue specimen. Compared with biopsy specimen B128C, specimen B213A showed an approximately 4-fold higher *H. pylori* colonization density as determined by microscopic examination. Analysis of cDNA obtained by SCOTS from this specimen (B213A) showed a correspondingly more comprehensive qualitative profile of in vivo bacterial RNA expression (Figure 3*B*). A total of 1144 array features, representing 68% of the 1681 *H. pylori* ORFs analyzed, showed hybridization with this probe. Comparison of this hybridization pattern with that of cDNA obtained from strain B213 grown in vitro identified 44 potentially differentially expressed ORFs (Table 1). These included 6 of the 21 previously identified as potentially differentially expressed in *H. pylori*

Table 3. *H. pylori* Genomic ORFs Showing Maximum Hybridization With cDNA From *H. pylori* B128 Grown In Vitro[a]

ORF	Annotation	ORF	Annotation
HP0010[b]	Chaperone (*groEL*)	HP0980	Conserved secreted protein
HP0011[b]	Co-chaperone (*groES*)	HP1023	Hypothetical protein
HP0070[b]	Urease accessory protein (*ureE*)	HP1101	Glucose-6-P dehydrogenase (*g6pD*)
HP0071[b]	Urease accessory protein (*ureI*)	HP1132	ATP synthase F1, β subunit (*atpD*)
HP0073[b]	Urease α subunit (*ureA*)	HP1201[b]	Ribosomal protein L1 (*rpl1*)
HP0090	Acyl carrier transacylase (*fabD*)	HP1204	Ribosomal protein L33 (*rpL33*)
HP0190	Conserved secreted protein	HP1205[b]	Elongation factor EF-Tu (*tufB*)
HP0282	Hypothetical protein	HP1243[b]	Adhesin (*babA2*)
HP0289	Toxin-like outer membrane protein	HP1252	Oligopeptide transporter (*oppA*)
HP0294	Aliphatic amidase (*aimE*)	HP1339	Biopolymer transport protein (*exbB*)
HP0472[b]	Omp11 (*horE*)	HP1427[b]	Histidine-rich polypeptide (*hpn*)
HP0486[b]	Hypothetical protein	HP1432	Histidine- and glutamine-rich protein
HP0513	Hypothetical protein	HP1452	Thiophene and furan oxidizer (*tdhF*)
HP0601[b]	Flagellin A (*flaA*)	HP1468[b]	Amino acid aminotransferase (*ilvE*)
HP0655[b]	Protective surface antigen D15	HP1469	Omp31 (*horJ*)
HP0777	Uridine 5′-P kinase (*pyrH*)	HP1480	Seryl-tRNA synthetase (*serS*)
HP0784	Hypothetical protein	HP1512[b]	Iron-regulated omp (*frpB*)
HP0875[b]	Catalase (*katA*)	HP1563[b]	Alkyl hydroperoxide reductase (*tsaA*)
HP0912[b]	Adhesin (*alpA*)	HP1580	Hypothetical protein
HP0913[b]	Adhesin (*alpB*)		
HP0920	Conserved integral membrane protein		

ATP, adenosine triphosphate; tRNA, transfer RNA.
[a]Radiolabeled first-strand cDNA probes were prepared from total RNA isolated from 3 *H. pylori* B128 cultures grown to midlogarithmic phase in modified brucella broth as described in Materials and Methods. Signal intensities were determined for each feature on each of 3 arrays as a pixel intensity value relative to the sum of the pixel intensities for all features on that array (except those with genomic DNA). A group of 40 array features with average intensity values greater than 2 SD above the calculated mean pixel value are listed. ORF numbers and annotations are those of the *H. pylori* 26695 genome.[32]
[b]ORFs for 19 array features also showed hybridization signal intensities 2 SD over the mean with a first-strand cDNA probe from *H. pylori* strain B213.

B128-infected human tissue, and 8 of 78 identified as differentially expressed in B128-infected gerbil tissue (Table 1).

A total of 123 different *H. pylori* ORFs were identified as potentially differentially expressed in analyses of these tissue specimens and the corresponding isolated strains grown in broth cultures (Table 1). Table 2 lists characteristics of 14 ORFs for array features that showed hybridization with bacterial cDNAs obtained by SCOTS from at least 2 tissue specimens but that consistently failed to show hybridization with cDNA probes prepared from either *H. pylori* strain B128 or B213 grown in vitro. The repeated identification of RNAs for these genomic regions in gastric tissue specimens (Table 2), but not in bacteria grown in vitro, suggests that the encoded bacterial factors are expressed in response to environments and interactions encountered in mammalian gastric mucosa.

RT-PCR Analyses of *H. pylori* In Vivo Gene Expression

In the next series of experiments, we used an independent experimental technique (RT-PCR) to corroborate results obtained by the DNA array analyses for several genes listed in Tables 1 and 2. We chose to investigate 4 ORFs (HP0061, HP0228, HP0343, and HP0667) whose differential expression had been detected in analyses of all 3 *H. pylori*–infected tissue samples and 2 ORFs (HP0833 and JHP0945) for which differential expression had been detected in both analyses of human gastric tissue specimens (Table 1). With RT-PCR, regions within 4 of these 6 ORFs (HP0061, HP0228, HP0667, and JHP0945) were successfully amplified from the original human and gerbil biopsy samples (Figure 4). Selected primer pairs did not amplify targeted regions within ORF HP0833 or HP0343, even though we subsequently confirmed the presence of cDNA for HP0833 by analysis of cDNA clones obtained from human gastric tissue specimen B128C (see below). We then extended this analysis to determine whether expression of these 4 genes (HP0061, HP0228, HP0667, and JHP0945) could be detected in gastric biopsy samples from 6 additional *H. pylori*–infected patients. Although we were unable to detect selected target regions of HP0061 and HP0667 transcripts, as shown in Figure 4, transcripts for HP0228 (encoding a sulfate permease) and JHP0945 (encoding a hypothetical protein) were successfully detected by RT-PCR in 5 of 6 and 6 of 6 biopsy specimens, respectively.

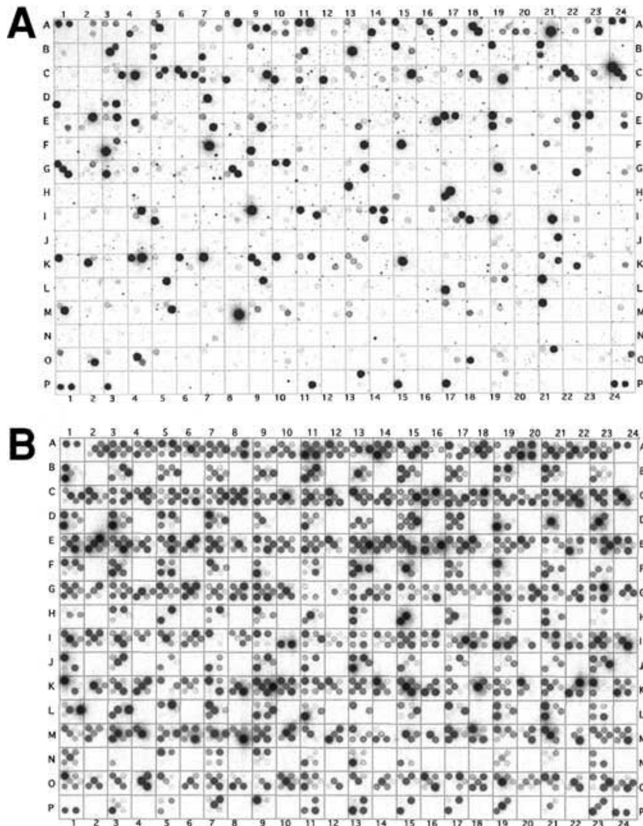

Figure 3. *H. pylori* genomic array hybridized with *H. pylori* cDNAs derived from human gastric tissues. Radiolabeled *H. pylori* cDNAs obtained by 3 rounds of SCOTS from gastric biopsy tissue specimens were used to probe 1681 arrayed genomic DNA fragments. Representative arrays are shown after hybridization with (*A*) bacterial cDNA from biopsy B128C and (*B*) bacterial cDNA from biopsy B213A.

Nucleotide Sequence Analysis of cDNA Obtained by SCOTS

As an additional method of identifying and confirming the in vivo expression of *H. pylori* RNAs, bacterial cDNA obtained by SCOTS from the human gastric biopsy specimen B128C was subjected to a differential hybridization procedure and additional PCR amplification, as outlined in Materials and Methods. The cDNA population obtained by 3 rounds of enrichment for differentially expressed sequences[11] was then cloned into an *E. coli* plasmid vector for sequence analysis. Nucleotide sequences of cDNA inserts, generally 200–600 base pairs in length, were determined and compared with the known genome sequences of *H. pylori* strains 26695 and J99. Among 95 randomly selected inserts, we identified 60 different cDNAs for in vivo–expressed RNAs (Table 4). Fifty-one of these corresponded to ORFs that were also identified by array hybridization with biopsy-derived bacterial cDNA probes (Figure 2). Several cDNAs identified by nucleotide sequence analysis were not detected by array hybridization (Table 4), indicating that they were present at low levels. Three additional inserts (*scs*HP1, *scs*HP2, and *scs*HP3) did not show any similarity to ORFs in the genome of either *H. pylori* strain 26695 or J99. However, all 3 showed significant similarity to *H. pylori* plasmid sequences (pHe14 and pHe15, GenBank accession numbers AF547991 and AF547992). These results are consistent with the observation that strain B128 possesses a native plasmid (T. Ando, personal communication, 2000).

Nucleotide sequence analysis of plasmid clones was particularly useful for definitively identifying cDNAs for members of paralogous *H. pylori* gene families. For example, by direct sequence analysis of cloned cDNAs, we were able to show the in vivo expression of mRNAs encoding 7 different *H. pylori* B128 outer-membrane proteins.[27] These showed greatest nucleotide similarity to the adhesins BabA (HP1243) and AlpB (HP0913), as well as HopE (HP0706), HopI (HP1156), HopL (HP1157), HopN (HP1342), and HorJ (HP1469).

Discussion

In this study, we used SCOTS in combination with genome array hybridization for the global analysis of RNA expression in *H. pylori* colonizing mammalian gastric mucosa. We then characterized RNA obtained from the isolated *H. pylori* strains cultured in vitro, to identify a group of ORFs that are potentially selectively

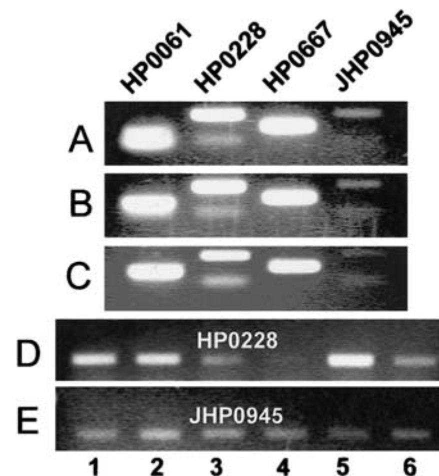

Figure 4. RT-PCR analyses of *H. pylori* in vivo gene expression. The cDNAs prepared from gastric tissue specimens from (*A*) patient B128, (*B*) patient B213, or (*C*) an *H. pylori* B128-infected gerbil were used as templates for amplification of target sequences within the indicated *H. pylori* genomic ORFs. (*D*) Subsequent RT-PCR analyses detected in vivo expression of transcripts for *H. pylori* HP0228, a putative sulfate permease, in 5 of 6 additional human tissue specimens. (*E*) Transcripts for JHP0945, an *H. pylori*–specific hypothetical protein, were detected in gastric tissues from all 6 *H. pylori*–infected patients.

Table 4. *H. pylori* Genomic ORFs Corresponding to Sequenced cDNA Clones Obtained by SCOTS From Human Gastric Biopsy Specimen B128C[a]

ORF	Predicted coding region	ORF	Predicted coding region
HP0010	Chaperone (*groEL*)	HP0887[b]	Vacuolating cytotoxin (*vacA*)
HP0033	ATP-dependent Clp protease (*clpA*)	HP0913	Adhesin (*alpB*)
HP0071	Urease accessory protein (*ureI*)	HP0965[b]	Hypothetical protein
HP0072	Urease β subunit (*ureB*)	HP0969	Cation efflux system protein (*czcA*)
HP0073	Urease α subunit (*ureA*)	HP0979	Cell divison protein (*ftsZ*)
HP0192	Fumarate reductase subunit (*frdA*)	HP1116	Hypothetical protein
HP0278	Guanosine phosphohydrolase (*gppA*)	HP1118	γ-glutamyltranspeptidase (*ggt*)
HP0305	Hypothetical protein	HP1156	Outer membrane protein (*hopI*)
HP0358	Hypothetical protein	HP1157[b]	Outer membrane protein (*hopL*)
HP0364	Nucleotide reductase subunit (*nrdB*)	HP1198	RNA polymerase subunit (*rpoB*)
HP0387	Primosomal replication factor (*priA*)	HP1205	Elongation factor EF-Tu (*tufB*)
HP0393[b]	Chemotaxis response regulator (*cheV*)	HP1243	Adhesin (*babA2*)
HP0397	Phosphoglycerate dehydrogenase (*serA*)	HP1257	Phosphoribosyltransferase (*pyrE*)
HP0417	Methionyl-tRNA synthetase (*metS*)	HP1329	Cation efflux system protein (*czcA*)
HP0440	DNA topoisomerase I (*topA*)	HP1342	Outer membrane protein (*hopN*)
HP0487	Hypothetical protein	HP1360	4-Hydroxybenzoate transferase (*ubiA*)
HP0488	Hypothetical protein	HP1379	ATP-dependent protease (*lon*)
HP0601	Flagellin A (*flaA*)	HP1398	Alanine dehydrogenase (*ald*)
HP0604	Uroporphyrinogen decarboxylase (*hemE*)	HP1422	Isoleucyl tRNA synthetase (*ileS*)
HP0605	Hypothetical protein	HP1440	Hypothetical protein
HP0693	Conserved membrane protein	HP1469	Outer membrane protein (*horJ*)
HP0706	Outer membrane protein (*hopE*)	HP1523	DNA recombinase (*recG*)
HP0723	*L*-asparaginase II (*ansB*)	HP1538	Ubiquinol cytochrome subunit c_1
HP0761	Hypothetical protein	HP1539	Ubiquinol cytochrome subunit *b*
HP0833	Hypothetical protein	HP1560	Cell division protein (*fts*II)
HP0870	Flagellar hook (*flgE*)	HP1565	Penicillin-binding protein 2 (*pbp2*)
HP0874[b]	Hypothetical protein	JHP0931	DNA topoisomerase I (*topA3*)
HP0875	Catalase (*katA*)	JHP0945	Hypothetical protein
HP0881[b]	Hypothetical protein		

ATP, adenosine triphosphate; tRNA, transfer RNA.
[a]*H. pylori* 26695 and J99 genomic ORFs with the greatest similarity to plasmid cDNA clones obtained from *H. pylori*–infected human gastric tissue by SCOTS as described in Materials and Methods.
[b]ORFs for array features not showing significant hybridization to biopsy-derived bacteria cDNA probes.

transcribed by bacteria in the stomach relative to bacteria grown under typical laboratory culture conditions.

To identify *H. pylori* genes not expressed (or expressed at only low levels) during growth in vitro, we analyzed RNA expression profiles of bacteria grown to the mid-logarithmic phase in standard broth culture media. On the basis of the analysis of total RNA prepared from 3 independent *H. pylori* B128 cultures, 488 of 1681 array features were identified that reproducibly failed to show hybridization signals above background levels. A similar number of array features also failed to hybridize cDNA prepared from *H. pylori* strain B213.

In an initial analysis of *H. pylori* in vivo gene expression, cDNA obtained by 3 rounds of SCOTS from human gastric tissue specimen B128C hybridized with 329 array features representing in vivo–expressed *H. pylori* RNAs. Most of these RNAs correspond to array features that showed strong hybridization signals when hybridized with first-strand cDNA probes from *H. pylori* B128 grown in vitro (data not shown). Subsequently, in analyses of an *H. pylori* B128-infected gerbil gastric tissue specimen and a second human gastric tissue biopsy (B213A), we were able to detect more than 1000 different bacterial RNAs, corresponding to approximately 70% of ORFs in the *H. pylori* genome. The more comprehensive representation of *H. pylori* transcripts in these analyses was likely the result of an improved method of homogenization of tissue samples for RNA extraction and, in the analysis of B213A, increased bacterial colonization density.

Most of the array features that failed to show hybridization with bacterial cDNA obtained directly from bacteria cultured in vitro also did not hybridize cDNA obtained by SCOTS from gastric tissue specimens; however, specific array features showed differential hybridization (Table 1). In each of 3 analyses of *H. pylori*–infected tissue specimens, the group of potentially differentially expressed genes identified comprised 4%–7% of the total number of bacterial genes for which RNAs were detected in infected gastric tissues (Table 1).

Failure to identify the same set of differentially expressed ORFs in the analysis of each tissue specimen was likely due to several factors. First, the SCOTS procedure did not yield cDNA preparations of equal complexity in all analyses. For example, in the analysis of gastric biopsy specimen B128C, it is clear that we identified only a subset of genes of interest. A second reason for the incomplete overlap between genes identified in different analyses of *H. pylori*–infected gastric tissue specimens likely relates to the fact that the gerbil model is unlikely to replicate all of the relevant aspects of a natural human gastric infection. Differences between the 2 human hosts and the well-recognized variability among *H. pylori* strains also may have been important factors. Considerable additional analyses will be necessary to investigate how variations in host, pathogen, disease state, localization, and colonization density affect detectable bacterial in vivo RNA expression patterns.

Despite these limitations, numerous genes were identified as potentially differentially expressed in multiple analyses. Four *H. pylori* genes were identified as potentially differentially expressed in all 3 tissue analyses, and 12 additional genes were identified as potentially differentially expressed in 2 of 3 tissue analyses (Table 1). To focus on genes that are most likely to be selectively expressed in response to bacteria–host interactions in the gastric mucosa, we then eliminated those genes for RNAs expressed to detectable levels by either strain B213 or B128 in broth cultures. Table 2 lists a group of 14 genes that we consider excellent candidates for further studies of *H. pylori* in vivo gene expression, including several whose expression can be studied in an animal model of gastric disease. By using RT-PCR, we confirmed the presence of transcripts for 4 of these genes in human gastric tissues (Figure 4). The presence of a fifth transcript (HP0833) was also confirmed by direct sequence analysis of cloned cDNA from the original human tissue specimen (Table 4). Moreover, transcripts for 2 of these genes (HP0228 and JHP0945) were readily detectable by RT-PCR in gastric biopsy specimens from multiple additional *H. pylori*–infected humans (Figure 4). HP0228 is predicted to encode a membrane protein involved in the transport of sulfate anion,[28] and sulfate acquisition is likely to be important for the in vivo survival of *H. pylori*, as well as for other bacterial pathogens.[29]

We were unable to confirm the presence of transcripts for all candidate differentially expressed genes analyzed by RT-PCR either in the original biopsy specimens or in tissues from additional patients. Further experimental work—for example, testing alternate target regions within bacterial mRNAs and increasing the sensitivity of our assays by nested PCR—will be required to confirm the expression of these genes by *H. pylori* in gastric mucosa.

H. pylori iceA1 transcript levels have previously been shown to increase after bacterial adherence to AGS gastric epithelial cells in vitro.[12] We detected substantial basal *iceA1* transcription in *H. pylori* cultures at midlogarithmic phase, confirming recent studies that focused on characterization of bacterial *iceA* transcripts.[30] Therefore, *iceA* did not meet the intentionally conservative criteria we used in this study to identify *H. pylori* genes potentially differentially expressed in vivo and in vitro.

Half of the ORFs listed in Table 2 have previously been shown to be present in the genomes of all 15 *H. pylori* strains examined by array hybridization.[22] Notably, this list is devoid of what are traditionally considered housekeeping genes, allowing us to hypothesize that they encode bacterial factors selectively produced in gastric mucosa that may contribute to the unique capacity of *H. pylori* to colonize the stomach. Many RNAs identified as differentially expressed (Table 1) are encoded by adjacent ORFs sharing the same chromosomal orientation, and therefore they potentially represent bacterial operons. Among all the ORFs identified, more than 10 contiguous clusters in the same chromosomal orientation can be identified (see http://genolist.pasteur.fr/PyloriGene/genome.cgi). The largest of these clusters is JHP0956–JHP0960, which comprises part of a putative 8-gene operon in the *H. pylori* J99 genome that shows no sequence similarity to genome sequences of other bacteria. HP0338, HP0340, and HP0343 are ORFs within a cluster of 12 *H. pylori*–specific ORFs sharing the same orientation on the *H. pylori* 26695 chromosome. HP0893 and HP0895 are the second and last of 4 tightly clustered collinear ORFs in the 26695 genome that were identified as differentially expressed in the analysis of human gastric tissue specimen B213. HP0893 and HP0895 are predicted to encode proteins lacking any significant sequence similarity to those encoded by other bacterial genomes, and HP0892 and HP0894 are annotated as conserved bacterial ORFs of unknown function.[31]

A striking feature of the genes listed in both Tables 1 and 2 is the preponderance of genes predicted to encode *H. pylori*–specific "hypothetical proteins." Of the 123 ORFs listed in Table 1, 65% are predicted to encode proteins specific to *H. pylori,* more than twice the proportion found in the genome as a whole. Of those iden-

tified as differentially expressed in more than 1 tissue analysis (Table 2), 75% are annotated as lacking homologs in other species. Bacterial species-specific genes are likely to encode factors relevant to survival in unique environments that are not inhabited by other species. On the basis of a comparison of the genome sequences of *E. coli*, *Haemophilus influenza,* and *Helicobacter pylori*, Huynen et al.[32] have noted that many *H. pylori* species-specific genes are known to be involved in interactions with human hosts, and they propose similar roles for other genes unique to *H. pylori.* Our functional analyses of *H. pylori* differential in vivo RNA expression support such a role for many of these previously uncharacterized genes.

Several aspects of the RNA-based approaches used for analysis of *H. pylori* gene expression in this study warrant discussion. Previous studies of bacterial gene expression have used SCOTS as a cDNA cloning strategy to identify differentially expressed bacterial RNAs.[11,33–35] In this study, the availability of *H. pylori* whole-genome arrays allowed us to characterize the entire bacterial cDNA population initially obtained by SCOTS from infected gastric tissue. Because SCOTS obtains bacterial cDNA by hybridization to bacterial genomic DNA, it is intended to obtain a normalized population of cDNA molecules. Use of this material as a probe for array hybridization therefore did not allow quantitative analysis of bacterial in vivo RNA expression levels. Although we used more conventional first-strand cDNA probes in the analysis of RNA expressed by bacteria growing in vitro, these analyses also provided only an approximation of the relative abundance of different transcripts. The size of arrayed genomic DNA fragments, allelic variation among *H. pylori* strains, and cross-hybridization can all potentially affect hybridization signal intensity. Further quantitative analyses of the transcripts identified in this study, as well as the encoded proteins, will be necessary to confirm the differential *H. pylori* gene expression suggested by our analyses.

Finally, we also attempted to obtain plasmid cDNA clones for differentially expressed *H. pylori* RNAs by using the originally described SCOTS differential cloning strategy.[11] However, array hybridization analyses of cDNA resulting from the enrichment procedure indicated that cDNAs for ORFs clearly transcribed during in vitro growth were not efficiently removed (data not shown). Array hybridization analyses of the cDNA used during the enrichment process showed that normalized cDNA obtained by SCOTS from large numbers of bacteria grown in vitro was very complex and was therefore capable of almost entirely blocking the genomic DNA fragments used in the enrichment procedure (data not shown). This likely contributed to our inability to obtain a small group of plasmid cDNA clones by using the originally described SCOTS differential cloning strategy. However, analysis of these cDNA clones confirmed the in vivo expression of several RNAs identified as potentially differentially expressed by array hybridization (HP0358, HP0833, and JHP0945), as well as others encoding products (including adhesins and transport proteins) that may be involved in colonization of the gastric mucosa (Table 4).

On the basis of genomic sequence analysis of 2 strains, *H. pylori* contains relatively few genes encoding components of previously described bacterial transcriptional regulatory networks compared with most other bacterial pathogens. Our current analysis has provided a list of candidate genes that are potentially expressed by *H. pylori* in response to environmental signals encountered in the human stomach, including several whose expression can be studied in an animal model of *H. pylori* gastritis. In future studies, it will be important to confirm and quantify differential expression by using other experimental approaches to identify the relevant environmental signals that lead to changes in *H. pylori* gene expression and to characterize the potentially novel regulatory mechanisms involved.

References

1. Dunn BE, Cohen H, Blaser MJ. *Helicobacter pylori.* Clin Microbiol Rev 1997;10:720–741.
2. Cover TL, Berg DE, Blaser MJ, Mobley HLT. *Helicobacter pylori* pathogenesis. In: Groisman EA, ed. Principles of bacterial pathogenesis. San Diego, CA: Academic Press, 2001:509–558.
3. Lee SH, Hava DL, Waldor MK, Camilli A. Regulation and temporal expression patterns of *Vibrio cholerae* virulence genes during infection. Cell 1999;99:625–634.
4. Lucas RL, Lee CA. Unravelling the mysteries of virulence gene regulation in *Salmonella typhimurium.* Mol Microbiol 2000;36: 1024–1033.
5. Mekalanos JJ. Environmental signals controlling expression of virulence determinants in bacteria. J Bacteriol 1992;174:1–7.
6. Handfield M, Levesque RC. Strategies for isolation of *in vivo* expressed genes from bacteria. FEMS Microbiol Rev 1999;23: 69–91.
7. Slauch JM, Mahan MJ, Mekalanos JJ. *In vivo* expression technology for selection of bacterial genes specifically induced in host tissues. Methods Enzymol 1994;235:481–492.
8. Merrell DS, Camilli A. Detection and analysis of gene expression during infection by *in vivo* expression technology. Philos Trans R Soc Lond B Biol Sci 2000;355:587–599.
9. Valdivia RH, Falkow S. Probing bacterial gene expression within host cells. Trends Microbiol 1997;5:360–363.
10. Moxon R, Tang C. Challenge of investigating biologically relevant functions of virulence factors in bacterial pathogens. Philos Trans R Soc Lond B Biol Sci 2000;355:643–656.
11. Graham JE, Clark-Curtiss JE. Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by

human macrophages by selective capture of transcribed sequences (SCOTS). Proc Natl Acad Sci U S A 1999;96:11554–11559.
12. Peek RM Jr, Thompson SA, Donahue JP, Tham KT, Atherton JC, Blaser MJ, Miller GG. Adherence to gastric epithelial cells induces expression of a *Helicobacter pylori* gene, *iceA*, that is associated with clinical outcome. Proc Assoc Am Physicians 1998;110:531–544.
13. Atherton JC, Cao P, Peek RM Jr, Tummuru MK, Blaser MJ, Cover TL. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. J Biol Chem 1995;270:17771–17777.
14. Peek RM Jr, Moss SF, Tham KT, Perez-Perez GI, Wang S, Miller GG, Atherton JC, Holt PR, Blaser MJ. *Helicobacter pylori cagA*+ strains and dissociation of gastric epithelial cell proliferation from apoptosis. J Natl Cancer Inst 1997;89:863–868.
15. Israel DA, Salama N, Arnold CN, Moss SF, Ando T, Wirth HP, Tham KT, Camorlinga M, Blaser MJ, Falkow S, Peek RM Jr. *Helicobacter pylori* strain-specific differences in genetic content, identified by microarray, influence host inflammatory responses. J Clin Invest 2001;107:611–620.
16. Emory SA, Belasco JG. The ompA 5′ untranslated RNA segment functions in *Escherichia coli* as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translational efficiency. J Bacteriol 1990;172:4472–4481.
17. Atherton JC, Tham KT, Peek RM Jr, Cover TL, Blaser MJ. Density of *Helicobacter pylori* infection *in vivo* as assessed by quantitative culture and histology. J Infect Dis 1996;174:552–556.
18. Froussard P. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. Nucleic Acids Res 1992; 20:2900.
19. Teng DH, Hsu F, Peterson I, Cardon KE, Caponigro G, Kamb A. Template selection during manipulation of complex mixtures by PCR. Biotechniques 2001;30:868–872, 874, 876–877.
20. Ko MS, Ko SB, Takahashi N, Nishiguchi K, Abe K. Unbiased amplification of a highly complex mixture of DNA fragments by 'lone linker'-tagged PCR. Nucleic Acids Res 1990;18:4293–4294.
21. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25: 3389–3402.
22. Salama N, Guillemin K, McDaniel TK, Sherlock G, Tompkins L, Falkow S. A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. Proc Natl Acad Sci USA 2000; 97:14668–14673.
23. Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, Caruso A, Uria-Nickelsen M, Mills DM, Ives C, Gibson R, Merberg D, Mills SD, Jiang Q, Taylor DE, Vovis GF, Trust TJ. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature 1999;397:176–180.
24. Fawcett P, Eichenberger P, Losick R, Youngman P. The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. Proc Natl Acad Sci USA 2000;97:8063–8068.
25. Peek RM Jr, Miller GG, Tham KT, Perez-Perez GI, Cover TL, Atherton JC, Dunn GD, Blaser MJ. Detection of *Helicobacter pylori* gene expression in human gastric mucosa. J Clin Microbiol 1995; 33:28–32.
26. Jungblut PR, Bumann D, Haas G, Zimny-Arndt U, Holland P, Lamer S, Siejak F, Aebischer A, Meyer TF. Comparative proteome analysis of *Helicobacter pylori*. Mol Microbiol 2000;36:710–725.
27. Alm RA, Bina J, Andrews BM, Doig P, Hancock RE, Trust TJ. Comparative genomics of *Helicobacter pylori*: analysis of the outer membrane protein families. Infect Immun 2000;68:4155–4168.
28. Iliopoulos I, Tsoka S, Andrade MA, Janssen P, Audit B, Tramontano A, Valencia A, Leroy C, Sander C, Ouzounis CA. Genome sequences and great expectations. Genome Biol 2001;2:1.1–1.3.
29. Wooff E, Michell SL, Gordon SV, Chambers MA, Bardarov S, Jacobs WR Jr, Hewinson RG, Wheeler PR. Functional genomics reveals the sole sulphate transporter of the *Mycobacterium tuberculosis* complex and its relevance to the acquisition of sulphur in vivo. Mol Microbiol 2002;43:653–663.
30. Donahue JP, Peek RM, Van Doorn LJ, Thompson SA, Xu Q, Blaser MJ, Miller GG. Analysis of *iceA1* transcription in *Helicobacter pylori*. Helicobacter 2000;5:1–12.
31. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzegerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature 1997;388:539–547.
32. Huynen M, Dandekar T, Bork P. Differential genome analysis applied to the species-specific features of *Helicobacter pylori*. FEBS Lett 1998;426:1–5.
33. Daigle F, Graham JE, Curtiss R III. Identification of *Salmonella typhi* genes expressed within macrophages by selective capture of transcribed sequences (SCOTS). Mol Microbiol 2001;41: 1211–1222.
34. Liu S, Graham JE, Bigelow L, Morse PD II, Wilkinson BJ. Identification of *Listeria monocytogenes* genes expressed in response to growth at low temperature. Appl Environ Microbiol 2002;68: 1697–1705.
35. Hou JY, Graham JE, Clark-Curtiss JE. *Mycobacterium avium* genes expressed during growth in human macrophages detected by selective capture of transcribed sequences (SCOTS). Infect Immun 2002;70:3714–3726.

Received August 2, 2001. Accepted August 8, 2002.

Address requests for reprints to: James E. Graham, Ph.D., Department of Microbiology and Immunology, 319 Abraham Flexner Wy., University of Louisville School of Medicine, Louisville, Kentucky 40202; e-mail: j.graham@louisville.edu; fax: (502) 852-2781.

This work was supported by a Vanderbilt University Discovery grant (to J.E.G.); grants RO1 DK53623, RO1 DK58587, and R29 CA77955; an F.D.H.N. Fiterman Foundation Award; and the Medical Research Service of the Department of Veterans Affairs.

We thank Rama Subramanian of Sigma–Genosys for preliminary data analyses, Mark McClain for assistance with computations, and John Donahue, Dawn Israel, and Wayne Shraw for helpful discussions.