Vol. 74, No. 5

Multiple Chromosomal Loci for the *babA* Gene in *Helicobacter pylori*

Ewa E. Hennig,[1] Johnna M. Allen,[2] and Timothy L. Cover[2,3,4]*

*Department of Gastroenterology, Medical Center for Postgraduate Education, Cancer Center, 02-781 Warsaw, Poland[1];
Departments of Medicine[2] and Microbiology and Immunology,[3] Vanderbilt University School of Medicine,
Nashville, Tennessee 37232; and VA Medical Center, Nashville, Tennessee 37212[4]*

Helicobacter pylori babA **encodes an outer membrane protein that binds to fucosylated Lewis b blood group antigen. We analyzed a panel of 35** *H. pylori* **strains and identified three possible chromosomal loci for** *babA*. **There was a significant association between the presence of** *babA* **and the presence of** *cagA* **($P = 0.0001$). Phylogenetic analysis of** *babA* **alleles revealed two divergent families of signal sequences. Among 17 strains in which an intact in-frame** *babA* **allele was identified, 10 expressed a detectable BabA protein. Expression of a BabA protein and the Lewis b-binding phenotype were not dependent on the chromosomal locus of** *babA*. **These data indicate that there is marked heterogeneity among** *H. pylori* **strains in** *babA* **genetic content and BabA expression.**

Helicobacter pylori is a gram-negative bacterium that colonizes the human stomach. *H. pylori* genomes contain about 30 related *hop* genes, which are predicted to encode outer membrane proteins (1). Several Hop proteins, including BabA, SabA, AlpA, AlpB, and HopZ, can mediate adherence of *H. pylori* to gastric epithelial cells. The BabA adhesin binds to the fucosylated Lewis b blood group antigen on the surfaces of these cells (3, 8).

Some *H. pylori* strains bind Lewis b antigen, whereas other strains do not (3, 5, 8). Differences among *H. pylori* strains in the Lewis b binding phenotype can potentially be explained by the observation that only some of them express a BabA protein (7, 15). Peptic ulcer disease and gastric adenocarcinoma occur more commonly in persons infected with BabA-expressing *H. pylori* strains than in persons infected with strains that do not express BabA (5, 15). The molecular basis for variation among strains in expression of the BabA protein has not yet been investigated in any detail. In a previous study, we analyzed BabA expression by immunoblotting with two anti-BabA ScFv antibodies and detected expression of BabA in only about half of the *H. pylori* strains tested (7). In the current study, we studied the same panel of strains in order to analyze *babA* genetic content and investigate the molecular basis for variation among strains in BabA expression.

Genome sequence analyses of *H. pylori* strains J99 and 26695 indicated that *babA* can be present in one of two possible chromosomal loci (2, 14), which we designate locus A and locus B, respectively. Strain J99 expresses BabA and binds to Lewis b antigen, whereas strain 26695 does not express a detectable BabA protein and does not bind to Lewis b antigen (7, 8, 11). The mechanistic basis for the absence of BabA expression in strain 26695 is not known. As a first step in analyzing *babA*, we designed PCR primers that could amplify locus A and locus B (Table 1), and we analyzed a panel of 35 previously described *H. pylori* strains (7). PCR products of the expected sizes were successfully amplified from locus A (>3.5 kb) and from locus B (>3 kb) in 32 and 33 of the *H. pylori* strains tested, respectively (Table 2). These amplicons were analyzed initially by nucleotide sequencing using primers A1-S, A3-AS, or A3b-AS (Table 1). The sequence data (about 600 to 800 nucleotides from each strain) were then compared with nucleotide sequences deposited in GenBank in order to identify the most closely related sequences.

As shown in Tables 2 and 3, 19 of the 32 sequences from locus A and 4 of the 33 sequences from locus B were most closely related to *babA*. Two strains (92-25 and J195) contained *babA* in both locus A and locus B. The other sequences were most closely related either to *babB* (exemplified by the genes designated HP0896 and jhp1164 in strains 26695 and J99, respectively) or to *omp9* (*hopU*) (exemplified by HP0317 in strain 26695) (Table 2). These experiments demonstrated that *babA* is found more commonly in locus A than in locus B and that *babA* chromosomal loci frequently contain one of two alternate *hop* genes (*babB* or *omp9*).

Based on the detection of *omp9* in *H. pylori* chromosomal loci known to commonly contain *babA* or *babB*, we reasoned that intragenomic recombination events might involve *babA* and *omp9* sequences, similar to what has been described for *babA* and *babB* sequences (10, 13). To analyze the locus corresponding to the site of *omp9* in *H. pylori* strain 26695 (designated locus C), appropriate primers (derived from genes flanking the *omp9* locus in *H. pylori* strain 26695) were designed (Table 1) and used for PCR assays. Amplicons of >4 kb were obtained from seven strains, and the remaining strains yielded amplicons of ≤2.8 kb. Sequence analysis revealed that three of the larger amplicons contained *babA*, two contained *omp9*, and two contained *babB* (Table 2). These data indicate that *babA* may be found in one of at least three possible chromosomal loci (Table 3).

In the analyses described above, 13 of the 35 strains did not contain detectable *babA* sequences in locus A, B, or C (Table 3). To investigate by another approach whether *babA* sequences were present, we attempted to PCR amplify fragments

* Corresponding author. Mailing address: Division of Infectious Diseases, A2200 Medical Center North, Vanderbilt University School of Medicine, Nashville, TN 37232. Phone: (615) 322-2035. Fax: (615) 343-6160. E-mail: timothy.L.cover@vanderbilt.edu.

TABLE 1. Oligonucleotide primers used for analysis of *H. pylori babA*

Primer name	Primer sequence	Description[a]
A99-S	5′GTATTTTGTGTAGTCTTTGTTGGTGG	jhp0832 (HP0893); locus A
A99-AS	5′GCAGTTGCATGGTCAGCTCTGAGG	jhp0835 (HP0898); locus A
AHp-S	5′CCACAGCATGATCATAGAGTATAGAG	jhp1163 (HP1242); locus B
AHp-AS	5′GGCTTTAATCCCCTACATTGTGGA	jhp1165 (HP1244); locus B
Hp1-S	5′GAAGCTTTGGTGTGGGGTTA	jhp0298 (HP0313); locus C
Hp1-AS	5′ACCCTAATGGGCATGTGGTA	jhp0301 (HP0318); locus C
OPE1243	5′ACCATCTTCAACAACGAGCCAGGG	jhp0833 (*babA*), bp 418–441 (forward)
OP9647	5′GGTATCGATCCACTTCCATCAC	jhp0833 (*babA*), bp 440–461 (forward)
OPE1244	5′AGCTCTTGCGCGTCCGTGATC	jhp0833 (*babA*), bp 1002–1022 (reverse)
OPE1245	5′GTGAAAGGGTTGAAAGGCTTGC	jhp0833 (*babA*), bp 1085–1106 (reverse)
A1-S	5′CCCGGGAGACGACGGCTTTTACACA[b]	jhp0833 (*babA*), bp 63–81 (forward)
A3-AS	5′CTCGAGGATGCTGTTTTTAAGATTT[b]	jhp0833 (*babA*), bp 1380–1362 (reverse)
A3b-AS	5′GCTGTATCTGCTGCTCTTGAGTGCC	jhp0833 (*babA*), bp 1414–1390 (reverse)

[a] Primers were derived from the indicated *H. pylori* chromosomal loci in *H. pylori* strain 26695 or J99. Locus A corresponds to the site of *babA* (jhp0833) in strain J99 and the site of *babB* (HP0896) in strain 26695. Locus B corresponds to the site of *babA* (HP1243) in strain 26695 and the site of *babB* (jhp1164) in strain J99. Locus C corresponds to the site of *omp9* (HP0317) in strain 26695.
[b] Underlined sequences represent restriction sites for SmaI and XhoI.

of *babA* from these strains, using primers OPE1243, OP9647, OPE1244, and OPE1245 in all four possible combinations (Table 1). These primers were designed, based on aligned sequences from many different *H. pylori* strains, to amplify *babA* but not related *H. pylori hop* gene sequences. PCR products of the expected size were amplified from 2 strains (87-91 and 87-203) but not from the remaining 11 strains. Sequence analysis confirmed that these two amplicons were fragments of *babA*. We hypothesize that there is an unidentified chromosomal locus for *babA* in these two strains. Thus, among the 35 strains analyzed in the current study, 24 contained *babA* sequences and 11 did not. Our conclusion that some strains lack *babA* is consistent with results obtained in other studies using microarray analysis (12), Southern hybridization (13), or PCR using different primers (5, 15). Twenty of the 24 strains that contained *babA* contained *cagA* (a marker for the *cag* pathogenicity island); *cagA* was present in only 1 of the 11 strains that lacked *babA* ($P = 0.0001$) (Table 3).

To gain further insight into *babA* genetic diversity, we analyzed the complete *babA* nucleotide sequences from 18 different *H. pylori* strains, including 3 strains (87-199, 60190, and J195) that contained two copies of *babA* (Table 3). This was accomplished by sequencing the PCR products described above on both strands with appropriate primers. In strain 87-199 (one of the strains that contained multiple copies of *babA*), the *babA* sequences found in loci B and C were identical and encoded full-length BabA proteins. In strain 60190, the *babA* sequences from loci A and C were almost identical except for

TABLE 2. Analysis of three potential chromosomal loci for *babA* in a panel of 35 *H. pylori* strains

Locus[a]	No. of strains[b]	babA	babB	omp9
		No. of strains in which the following gene was present in locus A, B, or C		
A	32	19	8	5
B	33	4	24	5
C	7	3	2	2

[a] Chromosomal loci A, B, and C are defined as described in the text.
[b] Total number of strains containing a detectable *hop* gene in locus A, B, or C.

TABLE 3. Analysis of BabA expression and *babA* genetic content in 35 *H. pylori* strains

Strain	cagA[a]	BabA protein[b]	Lewis b binding[c]	babA localization[d]	BabA length[e]
J99	+	+	2.3	A	744
87-29	+	+	1.7	A	741
87-33	+	+	2.5	A	740
J258	+	+	1.4	A	744
J133	+	+	1.7	A	ND
87-199	+	+	1.8	B	744
				C	744
92-25	+	+	2.3	A	ND
				B	ND
87-81	+	+	1.8	C	735
J262	−	+	1.1	A	ND
J190	−	+	1.0	—	NA
87-91	+	+	0.9	ID	ND
87-203	−	+	1.0	ID	ND
92-18	+	+	0.9	A	742
92-26	+	+	1.0	A	737
J116	+	+	1.0	A	742
J123	+	+	0.9	A	744
92-19	+	−	1.1	A	738
92-21	+	−	1.1	A	736
J166	+	−	0.9	A	ND
J195	−	−	1.0	A	55
				B	55
J223	+	−	1.0	A	745
26695	+	−	1.0	B	733
86-338	−	−	1.1	A	736
J178	+	−	0.9	A	737
60190	+	−	1.2	A	721
				C	721
92-28	−	−	1.2	—	NA
J154	−	−	1.1	—	NA
86-313	−	−	1.0	—	NA
J128	+	−	0.9	—	NA
92-20	−	−	1.1	—	NA
92-23	−	−	1.0	—	NA
87-230	−	−	0.9	—	NA
87-75	−	−	0.9	—	NA
J63	−	−	1.1	—	NA
Tx30a	−	−	0.9	—	NA

[a] Presence or absence of *cagA* nucleotide sequences (7).
[b] Immunoreactivity with anti-BabA ScFv antibodies (7).
[c] Values represent a ratio of optical density values, comparing binding of *H. pylori* strains to Lewis b versus binding to an albumin control. In this assay, values of >1.5 have been considered indicative of binding to Lewis b (5, 9, 13).
[d] Chromosomal loci A, B, and C are defined as described in the text. ID (indeterminate), *babA* sequences were detected in these strains but the *babA* chromosomal locus was not identified. —, *babA* sequences were not detected.
[e] Deduced lengths (in amino acids) of the encoded BabA proteins. Five previously determined *babA* sequences correspond to GenBank accession numbers AY549174 to AY549178 (7), and new *babA* sequences have been assigned GenBank accession numbers DQ225153 to DQ225165. ND, not determined; NA, not applicable.
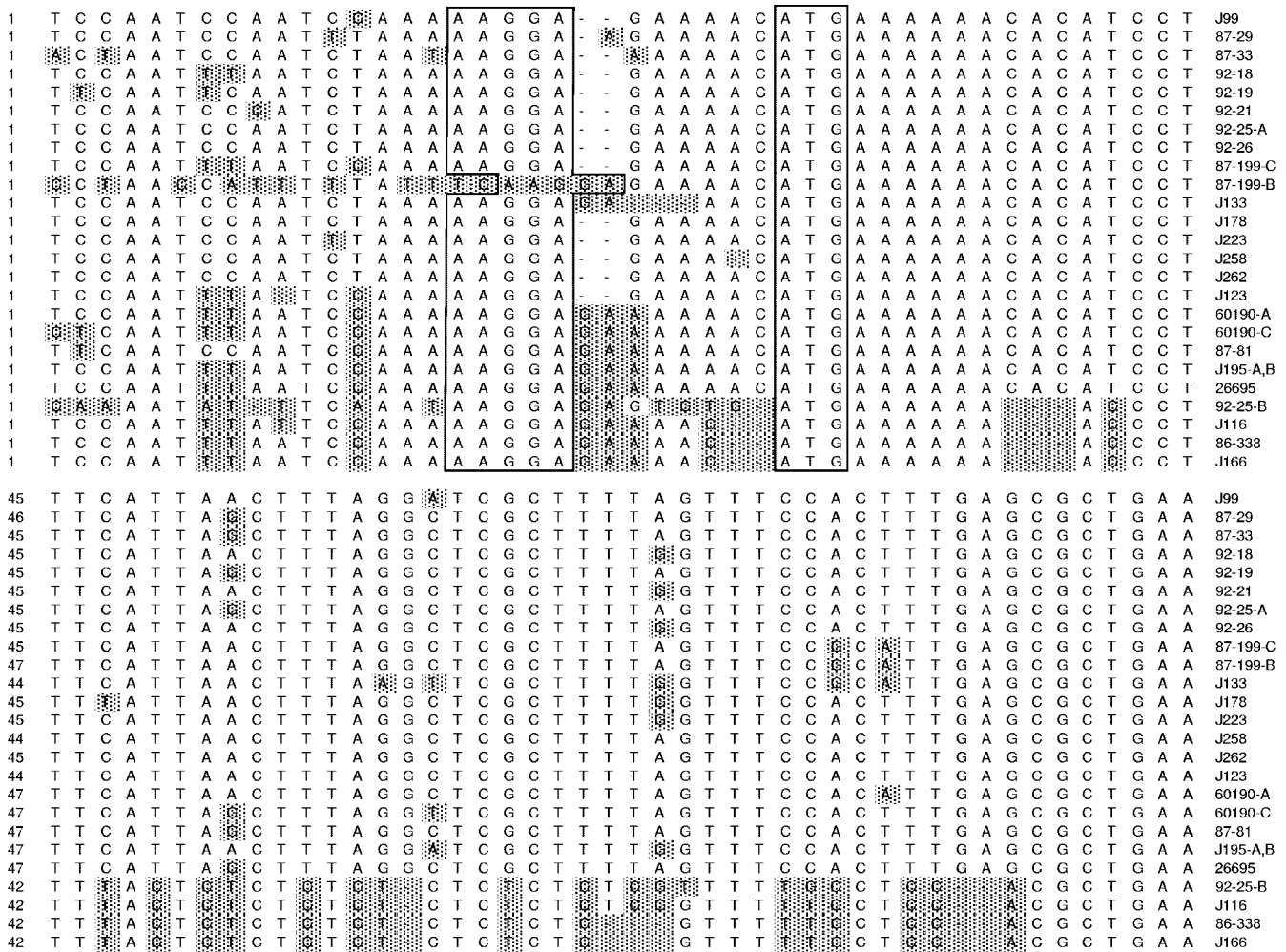
```
1   T C C A A T C C A A T C G A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   J99
1   T C C A A T C C A A T T T A A A  A A G G A  - A G A A A A C  A T G  A A A A A A C A C A T C C T   87-29
1   A C A A T C C A A T C T A A A A  A A G G A  - - A A A A A C  A T G  A A A A A A C A C A T C C T   87-33
1   T C C A A T A A A T C T A A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   92-18
1   T C C A A T C A A T C T A A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   92-19
1   T C C A A T C C A A T C T A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   92-21
1   T C C A A T C C A A T C T A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   92-25-A
1   T C C A A T C C A A T C T A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   92-26
1   T C C A A T A A A T C G A A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   87-199-C
1   C C A A C C A T T T A A T G A A G C A  G A A A A C  A T G  A A A A A A C A C A T C C T   87-199-B
1   T C C A A T C C A A T C T A A A  A A G G A  G A A C  A T G  A A A A A A C A C A T C C T   J133
1   T C C A A T C C A A T C T A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   J178
1   T C C A A T C C A A T T T A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   J223
1   T C C A A T C C A A T C T A A A  A A G G A  - - G A A A C  A T G  A A A A A A C A C A T C C T   J258
1   T C C A A T C C A A T C T A A A  A A G G A  - - G A A A A C  A T G  A A A A A A C A C A T C C T   J262
1   T C C A A T A A T C G A A A  A A G G A  - - G A A A C  A T G  A A A A A A C A C A T C C T   J123
1   T C C A A T A A T C G A A A  A A G G A  G A A A A A A C  A T G  A A A A A A C A C A T C C T   60190-A
1   C C A A T A A T C G A A A  A A G G A  G A A A A A C  A T G  A A A A A A C A C A T C C T   60190-C
1   T T C A A T C C A A T C G A A A  A A G G A  G A A A A A C  A T G  A A A A A A C A C A T C C T   87-81
1   T C C A A T A A T C G A A A  A A G G A  G A A A A A C  A T G  A A A A A A C A C A T C C T   J195-A,B
1   T C C A A T A A T C G A A A  A A G G A  G A A A A C  A T G  A A A A A A C A C A T C C T   26695
1   C A A A A T A T C A A A  A A G G A  C A G G T G  A T G  A A A A A A A C C C T   92-25-B
1   T C C A A T A A T C G A A A  A A G G A  G A A A A G  A T G  A A A A A A A C C C T   J116
1   T C C A A T A A T C G A A A  A A G G A  G A A A A G  A T G  A A A A A A A C C C T   86-338
1   T C C A A T A A T C G A A A  A A G G A  G A A A A G  A T G  A A A A A A A C C C T   J166
```

```
45  T T C A T T A A C T T T A G G T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   J99
46  T T C A T T A G C T T T A G G C T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   87-29
45  T T C A T T A G C T T T A G G C T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   87-33
45  T T C A T T A A C T T T A G G C T C G C T T T T T A G T T T C C A C T T T G A G C G C T G A A   92-18
45  T T C A T T A G C T T T A G G C T C G C T T T T T A G T T T C C A C T T T G A G C G C T G A A   92-19
45  T T C A T T A A C T T T A G G C T C G C T T T T T G G T T T C C A C T T T G A G C G C T G A A   92-21
45  T T C A T T A G C T T T A G G C T C G C T T T T T G G T T T C C A C T T T G A G C G C T G A A   92-25-A
45  T T C A T T A A C T T T A G G C T C G C T T T T T G G T T T C C A C T T T G A G C G C T G A A   92-26
45  T T C A T T A A C T T T A G G C T C G C T T T T A G T T T C C G C A T T G A G C G C T G A A   87-199-C
47  T T C A T T A A C T T T A G G C T C G C T T T T A G T T T C C G C A T T G A G C G C T G A A   87-199-B
44  T T C A T T A A C T T T A A G T C G C T T T T G G T T T C C A C T T T G A G C G C T G A A   J133
45  T T T A T T A A C T T T A G G C T C G C T T T T T G G T T T C C A C T T T G A G C G C T G A A   J178
45  T T C A T T A A C T T T A G G C T C G C T T T T T G G T T T C C A C T T T G A G C G C T G A A   J223
44  T T C A T T A A C T T T A G G C T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   J258
45  T T C A T T A A C T T T A G G C T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   J262
44  T T C A T T A A C T T T A G G C T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   J123
47  T T C A T T A A C T T T A G G C T C G C T T T T A G T T T C C A C A T T G A G C G C T G A A   60190-A
47  T T C A T T A G C T T T A G G C T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   60190-C
47  T T C A T T A A C T T T A G G A T C G C T T T T T G G T T T C C A C T T T G A G C G C T G A A   87-81
47  T T C A T T A G C T T T A G G A T C G C T T T T T G G T T T C C A C T T T G A G C G C T G A A   J195-A,B
47  T T C A T T A G C T T T A G G C T C G C T T T T A G T T T C C A C T T T G A G C G C T G A A   26695
42  T T A G T G C T C T C T C T C T G G T T T G G C T G C A C G C T G A A   92-25-B
42  T T A G T G C T C T C T C T C T G G T T T G G C T G C A C G C T G A A   J116
42  T T A G T G C T C T C T C T G G T T T G G C T G C A C G C T G A A   86-338
42  T T A G T G C T C T C T C T G G T T T G G C T G C A C G C T G A A   J166
```

FIG. 1. Nucleotide diversity in the 5′ region of *babA*. Nucleotide sequences in the 5′ region of *babA* from 22 *H. pylori* strains (including 4 strains with multiple copies of *babA*) were aligned. For strains with multiple copies of *babA*, the locus of the sequence (A, B, or C) is given next to the strain designation. ATG start codons and putative Shine-Dalgarno sequences are boxed. The *babA* sequences in 4 strains (J166, 86-338, J116, and one copy of *babA* from strain 92-25) were markedly different from those in the other 18 strains. Shading indicates nucleotides different from the consensus.

three substitutions near the 5′ ends of the genes; both copies of *babA* in this strain encoded full-length BabA proteins. In strain J195, the *babA* sequences found in loci A and B were identical, and each contained a frameshift mutation that prevented expression of a full-length BabA protein. The multiple copies of *babA* in these strains presumably resulted from gene conversion (intragenomic nonreciprocal recombination) events (4, 10, 13). With the exception of the two copies of *babA* in strain J195, all of the *babA* alleles analyzed were predicted to encode full-length BabA proteins, ranging from 721 to 745 amino acids (Table 3).

Analysis of a consensus alignment of the 5′ portion of *babA* alleles revealed two divergent families of sequences (Fig. 1). Four sequences (from strains J116, 86-338, J166, and 92-25 locus B) were markedly different in this region from *babA* sequences found in the other strains. However, further downstream, all of the sequences were closely related (data not shown). Multiple sequential CT repeats were present in one family of sequences (Fig. 1). Such regions are predicted to be subject to frameshift mutations, which could result in phase-variable expression of the encoded proteins.

To analyze the conservation and diversity of BabA amino acid sequences, we examined a consensus alignment of deduced BabA sequences. Consistent with the nucleotide sequence data shown in Fig. 1, two divergent families of amino acid sequences were detected near the amino terminus of BabA (Fig. 2). The region of divergence (amino acids 4 to 19) is predicted to constitute part of an amino-terminal signal peptide. The signal peptides encoded by four strains (J116, 86-338, J166, and 92-25-B) were closely related to signal peptides of paralogous BabB proteins (exemplified by HP0896 in strain 26695 and jhp1164 in strain J99) (Fig. 2). Notably, other parts of these four proteins were most closely related to BabA sequences (Fig. 2). It is striking that all four proteins had BabB-type signal peptides but sequences further downstream were typical BabA sequences.

We next sought to investigate potential links between *babA* genetic features, BabA protein expression, and binding to
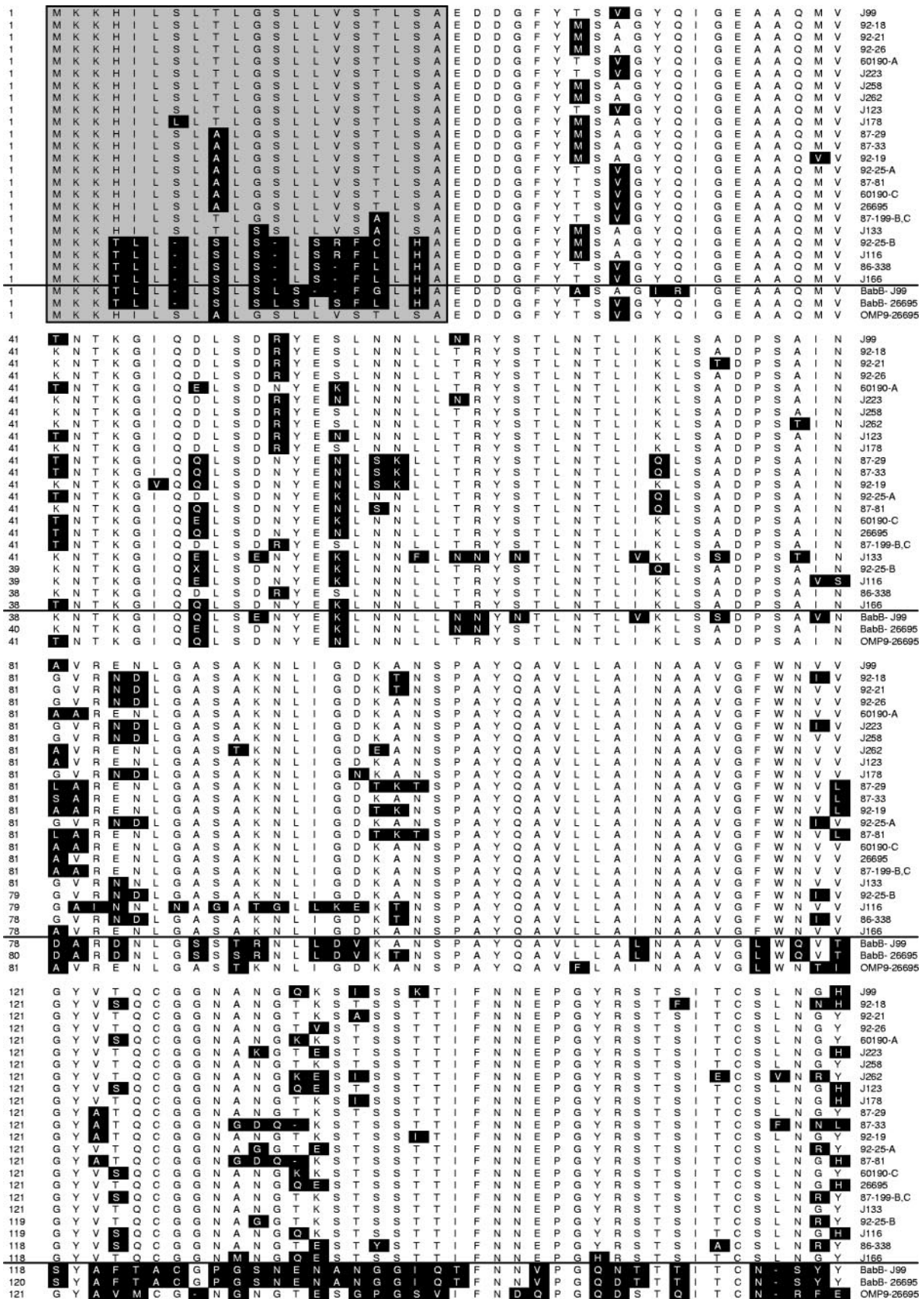
FIG. 2. Divergence among BabA signal peptides. Deduced N-terminal amino acid sequences of BabA from 21 *H. pylori* strains (including 3 strains with multiple copies of *babA*) were aligned. Deduced amino acid sequences of BabB (from strains J99 and 26695) and Omp9 (from strain 26695) are included in the alignment. Putative signal peptides are shaded and boxed. Two distinct families of BabA signal peptides can be distinguished. The signal peptides encoded by four strains are identical to BabB signal peptides. Black highlighting indicates amino acids different from the consensus.

Lewis b antigen. BabA protein expression in this panel of strains was characterized previously by immunoblot analysis (7). Sixteen of the 35 strains expressed a detectable BabA protein (Table 3). Nearly all of these 16 strains contained *babA* in locus A, but 1 contained *babA* in locus C and 1 contained two copies of *babA* in loci B and C. Among 17 strains in which an intact in-frame *babA* allele (predicted to encode a BabA protein of >700 amino acids) was identified, only 10 expressed a detectable BabA protein (Table 3). Potentially, BabA expression was not detected in some strains due to BabA amino sequence diversity and failure of anti-BabA antibodies to recognize certain epitopes. However, in a phylogenetic analysis, the *babA* alleles from BabA-expressing strains did not cluster separately from strains that failed to express BabA (data not shown).

We then tested all 35 strains for the ability to bind Lewis b antigen, using an in vitro adhesion assay (9). Among strains that expressed BabA, there was considerable variation in the level of binding to Lewis b antigen (Table 3). In previous studies, Lewis b/control binding ratios of >1.5 have been considered indicative of binding to Lewis b antigen (5, 9, 13). Based on this criterion, 7 strains exhibited detectable binding to Lewis b antigen and the other 28 strains did not (Table 3). As expected, all seven strains that bound to Lewis b antigen expressed a detectable BabA protein. Five of these seven strains contained *babA* in locus A. Among the four strains with divergent sequences near the 5′ end of *babA* (J166, 86-338, J116, and 92-25) (Fig. 1 and 2), two strains produced a detectable BabA protein and two did not. Strain J116 expressed a detectable BabA protein but did not exhibit detectable binding to Lewis b antigen. Strain 92-25 expressed a BabA protein and bound to Lewis b antigen, but this strain contained two nonidentical copies of *babA*. Therefore, it is not possible to reach any conclusions about the functional role of the divergent copy of *babA* in this strain.

It has been reported previously that the *babA1* gene in strain CCUG17875 lacks a translation initiation codon (8). To determine whether the lack of BabA expression in some strains was due to this phenomenon, we analyzed the 5′ regions of *babA* alleles. Each of the *babA* alleles contained an ATG start codon (Fig. 1); therefore, the absence of a translational initiation codon, as described for *babA1* from CCUG17875, is probably uncommon. Each of the genes analyzed contained an AAGGA sequence, corresponding to a putative ribosomal binding site, upstream of the ATG start codon (Fig. 1). There was variation among strains in the spacing between the ribosomal binding site and the BabA translational start codon (Fig. 1), but there was no significant correlation between the spacing of the Shine-Dalgarno sequence and expression of a BabA protein (compare Fig. 1 and Table 3). A TGNTATAAT consensus sequence, corresponding to an extended −10 motif, was present upstream of the *babA* translation initiation codon in all of the strains analyzed, with the exception of the locus B copy of *babA* in strain 87-199 (data not shown). Further upstream of the putative −10 sequence, there was considerable nucleotide sequence variation among strains, but we did not identify any motifs that correlated with the presence or absence of BabA protein expression.

In summary, these data indicate that there is extensive variation among *H. pylori* strains in *babA* genetic content. There are multiple possible chromosomal loci for *babA*. Individual strains can contain anywhere from 0 to 2 copies of *babA*. Frameshift mutations can result in the absence of BabA expression. Chimeric *babB/babA* alleles are commonly detected, and polymorphisms are present throughout *babA* alleles. In addition to this extensive variation among strains in *babA* content, a recent study showed that there can be metastability in *babA* sequences within the same strain (4). Based on the observation that many strains do not contain *babA* and do not bind to Lewis b antigen, it may be presumed that BabA expression is not required for persistent *H. pylori* colonization of the human stomach. However, adherence of *H. pylori* to gastric epithelial cells via the BabA adhesin may be a factor that contributes to the development of gastroduodenal disease (5, 6).

Nucleotide sequence accession numbers. The *babA* sequences newly determined in this study have been assigned GenBank accession numbers DQ225153 to DQ225165.

REFERENCES

1. Alm, R. A., J. Bina, B. M. Andrews, P. Doig, R. E. Hancock, and T. J. Trust. 2000. Comparative genomics of *Helicobacter pylori*: analysis of the outer membrane protein families. Infect. Immun. **68:**4155–4168.
2. Alm, R. A., L. S. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:**176–180.
3. Aspholm-Hurtig, M., G. Dailide, M. Lahmann, A. Kalia, D. Ilver, N. Roche, S. Vikstrom, R. Sjostrom, S. Linden, A. Backstrom, C. Lundberg, A. Arnqvist, J. Mahdavi, U. J. Nilsson, B. Velapatino, R. H. Gilman, M. Gerhard, T. Alarcon, M. Lopez-Brea, T. Nakazawa, J. G. Fox, P. Correa, M. G. Dominguez-Bello, G. I. Perez-Perez, M. J. Blaser, S. Normark, I. Carlstedt, S. Oscarson, S. Teneberg, D. E. Berg, and T. Boren. 2004. Functional adaptation of BabA, the *Helicobacter pylori* ABO blood group antigen binding adhesin. Science **305:**519–522.
4. Backstrom, A., C. Lundberg, D. Kersulyte, D. E. Berg, T. Boren, and A. Arnqvist. 2004. Metastability of *Helicobacter pylori* bab adhesin genes and dynamics in Lewis b antigen binding. Proc. Natl. Acad. Sci. USA **101:**16923–16928.
5. Gerhard, M., N. Lehn, N. Neumayer, T. Boren, R. Rad, W. Schepp, S. Miehlke, M. Classen, and C. Prinz. 1999. Clinical relevance of the *Helicobacter pylori* gene for blood-group antigen-binding adhesin. Proc. Natl. Acad. Sci. USA **96:**12778–12783.
6. Guruge, J. L., P. G. Falk, R. G. Lorenz, M. Dans, H. P. Wirth, M. J. Blaser, D. E. Berg, and J. I. Gordon. 1998. Epithelial attachment alters the outcome of *Helicobacter pylori* infection. Proc. Natl. Acad. Sci. USA **95:**3925–3930.
7. Hennig, E. E., R. Mernaugh, J. Edl, P. Cao, and T. L. Cover. 2004. Heterogeneity among *Helicobacter pylori* strains in expression of the outer membrane protein BabA. Infect. Immun. **72:**3429–3435.
8. Ilver, D., A. Arnqvist, J. Ogren, I. M. Frick, D. Kersulyte, E. T. Incecik, D. E. Berg, A. Covacci, L. Engstrand, and T. Boren. 1998. *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. Science **279:**373–377.
9. Olfat, F. O., Q. Zheng, M. Oleastro, P. Voland, T. Boren, R. Karttunen, L. Engstrand, R. Rad, C. Prinz, and M. Gerhard. 2005. Correlation of the *Helicobacter pylori* adherence factor BabA with duodenal ulcer disease in four European countries. FEMS Immunol. Med. Microbiol. **44:**151–156.
10. Pride, D. T., and M. J. Blaser. 2002. Concerted evolution between duplicated genetic elements in *Helicobacter pylori*. J. Mol. Biol. **316:**629–642.
11. Pride, D. T., R. J. Meinersmann, and M. J. Blaser. 2001. Allelic variation within *Helicobacter pylori* babA and babB. Infect. Immun. **69:**1160–1171.
12. Salama, N., K. Guillemin, T. K. McDaniel, G. Sherlock, L. Tompkins, and S. Falkow. 2000. A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. Proc. Natl. Acad. Sci. USA **97:**14668–14673.
13. Solnick, J. V., L. M. Hansen, N. R. Salama, J. K. Boonjakuakul, and M. Syvanen. 2004. Modification of *Helicobacter pylori* outer membrane protein expression during experimental infection of rhesus macaques. Proc. Natl. Acad. Sci. USA **101:**2106–2111.
14. Tomb, J.-F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketcham, H. P. Klenk, S. Gill, B. A. Dougherty, K.

Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzegerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature **388:** 539–547.

15. **Yamaoka, Y., J. Souchek, S. Odenbreit, R. Haas, A. Arnqvist, T. Boren, T. Kodama, M. S. Osato, O. Gutierrez, J. G. Kim, and D. Y. Graham.** 2002. Discrimination between cases of duodenal ulcer and gastritis on the basis of putative virulence factors of *Helicobacter pylori*. J. Clin. Microbiol. **40:**2244–2246.