JOURNAL OF CLINICAL MICROBIOLOGY, June 2002, p. 2244–2246 Vol. 40, No. 6
0095-1137/02/$04.00+0 DOI: 10.1128/JCM.40.6.2244–2246.2002

Discrimination between Cases of Duodenal Ulcer and Gastritis on the Basis of Putative Virulence Factors of *Helicobacter pylori*

Yoshio Yamaoka,[1]* Julianne Souchek,[2] Stefan Odenbreit,[3] Rainer Haas,[3] Anna Arnqvist,[4,5] Thomas Borén,[5] Tadashi Kodama,[6] Michael S. Osato,[1] Oscar Gutierrez,[7] Jong G. Kim,[8]† and David Y. Graham[1]

Department of Medicine[1] and Department of Health Services Research,[2] Baylor College of Medicine and Veterans Affairs Medical Center, Houston, Texas; Max Von Pettenkofer-Institut, München, Germany[3]; Division of Oral Microbiology, Department of Odontology,[4] and Department of Molecular Biology,[5] Umeå University, Umeå, Sweden; Third Department of Internal Medicine, Kyoto Prefectural University of Medicine, Kyoto, Japan[6]; Universidad Nacional de Colombia, Bogota, Colombia[7]; and Guro Hospital, Korea University College of Medicine, Seoul, Korea[8]

Received 22 January 2002/Returned for modification 24 March 2002/Accepted 2 April 2002

The BabA, *cagA*, and *vacA* statuses of 827 *Helicobacter pylori* isolates were used in logistic regression models to discriminate duodenal ulcer from gastritis. Only BabA was a candidate for a universal virulence factor, but the low *c* statistic value (0.581) indicates that none of these factors were helpful in predicting the clinical presentation.

As a general rule, important disease-associated bacterial toxins are tightly associated with their respective diseases irrespective of geographic region (e.g., cholera toxin and cholera) (3). Recently it was suggested that the *babA2* gene of *Helicobacter pylori* was associated with duodenal ulcer in a German population (2). Because of a strong correlation between the presence of *babA2*, *vacA* s1, and *cagA* (triple-positive strains) and the prevalence of duodenal ulcers, it was suggested that this pattern could be used to identify patients at higher risk for specific *H. pylori*-related diseases (2). We examined a large number of strains from both Western and East Asian countries and constructed models to discriminate duodenal ulcer from gastritis on the basis of the presence of *H. pylori* putative virulence factors *babA*, *cag* pathogenicity island (PAI), and *vacA*.

(Preliminary results of this study were presented in abstract form at the 102nd Annual Meeting of the American Gastroenterological Association, May 2001, Atlanta, Georgia [Gastroenterology **120**(Suppl. 1)**:**A652], and at the 14th International Workshop on Gastroduodenal Pathology and *Helicobacter pylori*, September 2001, Strasbourg, France [Gut **48** (Suppl. 11)**:**A15].)

H. pylori organisms were cultured from antral biopsy specimens by using standard methods (8, 9), and multiple colonies on the plates were used to extract the genomic DNA. The *vacA* genotypes (s1 or s2 and m1 or m2) were determined on the basis of PCR results as previously described (1, 8). Samples with mixed infections of different *vacA* genotypes were excluded. The *cag* PAI status was evaluated by PCR for seven loci in the island: *cagA*, *cagE*, *cagG*, *cagM*, *cagT*, HP0527, and HP0524 (strains with HP numbers are derived from strain 26695; GenBank accession number AE000511) (7). Specific primers for the *cag* empty site were used to confirm the absence of the *cag* PAI, as previously described (5). We regarded samples as *cag* PAI positive if all seven genes yielded positive PCR results. Samples positive for all *cag* PAI primers and *cag* empty site primers were regarded as harboring mixed infections with *cag* PAI-positive and -negative strains and were excluded from further analysis. When samples were positive by PCR for only some of the seven genes, Southern blot hybridization was performed. When a given gene was detected by hybridization and not by PCR, the isolate was considered positive for that gene. We regarded isolates with results positive for only some of the genes as *cag* PAI deletion strains and excluded them from analysis.

We evaluated the BabA status of the isolates by a combination of PCR and Western blot analysis. The functional BabA adhesin is encoded by the *babA2* gene, which differs from the *babA1* gene by the presence of a 10-bp insert with a repeat motif in the signal peptide sequence. It is therefore technically difficult to evaluate *babA2*/BabA status by PCR alone. We assessed BabA protein based on the premise that a combination of PCR and Western blotting techniques would provide a more accurate evaluation of the BabA status than PCR alone. BabA-specific antiserum was generated using a fusion protein with an N-terminal MS2 polymerase, and the His tag was produced using the *Escherichia coli* expression vector pEV40 (6). Parts of the *babA2* gene lacking its signal sequence were amplified by PCR and the fragment was cloned into pEV40, resulting in plasmid pSO129a. The fusion protein was expressed by temperature induction of the λ_{PL} promoter and purified by Ni^{2+}-nitrilotriacetic acid affinity chromatography. The purified fusion protein was used to immunize a rabbit to obtain the polyclonal antiserum AK253. BabA status was defined as positive if the Western blotting yielded positive results. The samples were BabA negative if both the PCR and Western blotting yielded negative results. We excluded samples that were PCR positive but Western blotting negative from further analysis.

* Corresponding author. Mailing address: Veterans Affairs Medical Center (111D), 2002 Holcombe Blvd., Houston, TX 77030. Phone: (713) 794-7234. Fax: (713) 790-1040. E-mail: yyamaoka@bcm.tmc.edu.
† Deceased.

TABLE 1. Univariate analysis of the relationship between *H. pylori* typing and duodenal ulcer prevalence in four countries[a]

Location(s)	Disease (no. of cases)	No. of cases (% of total cases)			
		BabA	*cag* PAI	*vacA* s1	*vacA* m1
East Asian countries					
Japan and Korea	Duodenal ulcer (295)	283 (96)	288 (98)	295 (100)	282 (96)
	Gastritis (179)	159 (89)	176 (98)	179 (100)	174 (97)
P value		0.003*	NS	NS	NS
Japan	Duodenal ulcer (180)	172 (96)	177 (98)	180 (100)	178 (99)
	Gastritis (126)	112 (89)	124 (98)	126 (100)	124 (98)
P value		0.024*	NS	NS	NS
Korea	Duodenal ulcer (115)	111 (97)	111 (97)	115 (100)	104 (90)
	Gastritis (53)	47 (89)	52 (98)	53 (100)	50 (94)
P value		0.054	NS	NS	NS
Western countries					
United States and Columbia	Duodenal ulcer (209)	176 (84)	196 (94)	187 (89)	137 (66)
	Gastritis (144)	103 (72)	124 (86)	114 (79)	100 (69)
P value		0.003*	0.013*	0.006*	NS
United States	Duodenal ulcer (145)	123 (85)	138 (95)	136 (94)	98 (68)
	Gastritis (92)	66 (72)	77 (84)	73 (79)	64 (70)
P value		0.012*	0.003*	0.001*	NS
Colombia	Duodenal ulcer (64)	53 (83)	58 (91)	51 (80)	39 (61)
	Gastritis (52)	37 (71)	47 (90)	41 (79)	36 (69)
P value		0.102	NS	NS	NS

[a] *P* values were determined by Fisher's exact test. *, significantly different ($P < 0.05$); NS, not significant ($P > 0.1$).

Overall, the analysis was based on data from 827 patients, 504 (61%) with duodenal ulcer and 323 with gastritis, including 474 from Japan and Korea (East Asia), of whom 295 (62%) had duodenal ulcer. There were 353 from Colombia and the United States (Western), of whom 209 (59%) had duodenal ulcer. There were no significant differences in gender or age distribution among the populations or in disease presentation.

Independent univariate analysis using Fisher's exact test showed that the presence of BabA was significantly related to that of duodenal ulcer in both East Asian and Western populations (Table 1). In the United States, the presence of *cag* PAI and *vacA* s1 genotypes was also significantly related to that of duodenal ulcer. However, as the BabA, *cag* PAI, and *vacA* genotypes were closely linked (data not shown), univariate analysis had limited power. We therefore constructed logistic regression models to test whether it was possible to discriminate duodenal ulcer from gastritis. For logistic regression analysis, *cag* PAI and BabA were coded as 1 for positive and 0 for negative and *vacA* s and *vacA* m were coded as 1 (for s2 or m2) or 0 (for s1 or m1). Regions (East Asian and Western) and diseases were included as independent variables. The reference group in the logistic regression consisted of *cag* PAI positive, BabA positive, *vacA* s1, and *vacA* m1, because it was by far the largest group. From the best subsets of predictors, the largest set which had significantly nonzero coefficients and maximized the area under the receiver operating characteristic (ROC) curve was chosen as the discriminant function. The *c* statistic, or area under the ROC curve, expressed the probability of correctly identifying from a random pair (normal and abnormal) the abnormal individual. The selection rule for logistic regression is to pick the individual with the higher probability of duodenal ulcer. A significance level of 0.05 was used for tests of hypotheses. The *c* statistic ranged from 0 to 1, with 1 indicating a perfect prediction and 0.5 indicating a chance prediction. SAS 6.12 procedures were used for the statistical analyses (SAS Institute Inc., Cary, N.C.).

The best model to discriminate duodenal ulcer from gastritis contained BabA, irrespective of the regional category (Table 2). For example, the absence of BabA decreased the odds of the presence of duodenal ulcer by a factor of 0.49 in all four countries ($P < 0.01$). Although many investigators have sought putative virulence markers in East Asian countries, no reliable disease specificity has emerged, in part because the *cag* PAI-positive and *vacA* s1 genotypes are almost uniformly present in these countries (8, 9). The fact that the best model contained BabA even in East Asia implies that BabA is important. BabA is one of the outer membrane proteins in *H. pylori*, and the biological importance of BabA as an adhesin to Lewis b antigen has been confirmed in vitro (4). However, it is important to note that although the best model contained BabA, the *c* statistic value was relatively low (0.583 for all four countries). Although the absence of BabA decreased the risk of duodenal ulcer by an odds ratio of 0.34 in East Asian countries, which was statistically significant, the overall *c* statistic value of the model in East Asian countries (0.543) did not reach statistical significance. Because the prevalence of BabA was very high even in gastritis cases in all four countries (Table 1), the presence of BabA alone was not a useful marker for duodenal ulcer.

TABLE 2. The best models to discriminate duodenal ulcer from gastritis on the basis of *H. pylori* types

Population (countries)	Predicted strain or characteristic	Coefficient	Odds ratio	*c* statistic
Combined (Japan, Korea, United States, and Colombia)	BabA negative	−0.721[a]	0.49[b]	0.583[c]
	vacA s1-m2	0.867[a]	2.38[b]	
	vacA s2-m2	−1.209[a]	0.30[b]	
Intercept		0.497		
East Asian (Japan and Korea)	BabA negative	−1.079[a]	0.34[b]	0.543
	vacA s1-m2	0.387	1.47	
Intercept		0.569		
Western (United States and Colombia)	*vacA* s1-m2	1.070[a]	2.92[b]	0.622[c]
	vacA s2-m2	−1.428[a]	0.24[b]	
	BabA negative	−0.486	0.62	
Intercept		0.374		

[a] Coefficient significantly different from 0 at the 0.05 significance level.
[b] Odds ratio is significantly greater or less than 1 at the 0.05 significance level.
[c] Area under the ROC curve is significantly greater than 0.50 at the 0.05 significance level.

The best logistic regression model also contained *vacA* s1/*vacA* m2 irrespective of the country of origin (Table 2). It is of interest that although the presence of *vacA* s1 and *vacA* m2 increased the odds of the presence of duodenal ulcer in our model, if both *vacA* s2 and *vacA* m2 were present, the odds of a duodenal ulcer were significantly decreased, a finding which is consistent with prior data showing that *vacA* m2 genotypes alone are not markers for duodenal ulcer disease. Taken together, our results therefore show that *vacA* m genotypes alone do not have value in predicting disease presentation.

Overall, although BabA appears to be a good candidate for universal *H. pylori* virulence factor status as well as being biologically plausible, neither BabA nor combinations of the BabA, *cag* PAI, and *vacA* genotypes were helpful in predicting the clinical presentation of an *H. pylori* infection.

The material presented here is based upon work supported in part by the Office of Research and Development, Medical Research Service, Department of Veterans Affairs, and by Public Health Service grants DK-53659 and DK56338, funding the Texas Gulf Coast Digestive Diseases Center. We also acknowledge the generous support of Hilda Schwartz.

We also thank Shogo Kikuchi (Department of Public Health, Aichi Medical University, Aichi, Japan) for help with statistical analyses.

REFERENCES

1. **Atherton, J. C., P. Cao, R. M. Peek, M. K. R. Tummuru, M. J. Blaser, and T. L. Cover.** 1995. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. J. Biol. Chem. **270:**17771–17777.
2. **Gerhard, M., N. Lehn, N. Neumayer, T. Borén, 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.
3. **Graham, D. Y., and Y. Yamaoka.** 2000. Disease-specific *Helicobacter pylori* virulence factors: the unfulfilled promise. Helicobacter **5**(Suppl. 1)**:**S3–S9.
4. **Ilver, D., A. Arnquist, J. Ögren, I. Frick, D. Kersulyte, E. T. Incecik, D. E. Berg, A. Covacci, L. Engstrand, and T. Borén.** 1998. *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. Science **279:**373–377.
5. **Mukhopadhyay, A. K., D. Kersulyte, J. Y. Jeong, S. Datta, Y. Ito, A. Chowdhury, S. Chowdhury, A. Santra, S. K. Bhattacharya, T. Azuma, B. Nair, and D. E. Berg.** 2000. Distinctiveness of genotypes of *Helicobacter pylori* in Calcutta, India. J. Bacteriol. **182:**3219–3227.
6. **Pohlner, J., J. Kramer, and T. F. Meyer.** 1993. A plasmid system for high-level expression and in vitro processing of recombinant proteins. Gene **130:**121–126.
7. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann, K. A. Ketchum, 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.
8. **Yamaoka, Y., T. Kodama, M. Kita, J. Imanishi, K. Kashima, and D. Y. Graham.** 1998. Relationship of *vacA* genotypes of *Helicobacter pylori* to *cagA* status, cytotoxin production, and clinical outcome. Helicobacter **4:**241–253.
9. **Yamaoka, Y., T. Kodama, O. Gutierrez, J. G. Kim, K. Kashima, and D. Y. Graham.** 1999. Relationship between *Helicobacter pylori iceA*, *cagA*, and *vacA* status and clinical outcome: studies in four different countries. J. Clin. Microbiol. **37:**2274–2279.