Biochemistry (Moscow), Vol. 69, No. 5, 2004, pp. 536-541. Translated from Biokhimiya, Vol. 69, No. 5, 2004, pp. 658-665.

Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM03-207, January 20, 2004.

Comparative Analysis of Different Typing Methods for *Helicobacter pylori* Clinical Isolates

V. M. Govorun[1,2]*, P. G. Lokhov[1], S. A. Moshkovskii[1], K. T. Momynaliev[2], O. V. Selesnyova[2], L. V. Kudryavtseva[2], M. V. Serebryakova[1], O. V. Tikhonova[1], E. I. Goufman[1], and A. I. Archakov[1]

[1]Orekhovich Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Pogodinskaya ul. 10, Moscow 119121, Russia; fax: (7-095) 245-0857; E-mail: proteomics@ibmh.msk.su
[2]Institute of Physico-Chemical Medicine, Ministry of Public Health of Russian Federation, Malaya Pirogovskaya ul. 1a, Moscow 119828, Russia; fax: (7-095) 246-4401

Received August 6, 2003
Revision received October 6, 2003

Abstract—The goal of the present work was to compare different techniques of molecular typing using as an example clinical isolates of *Helicobacter pylori* obtained from patients in different regions of Russia. DNA-macroarray genome scanning using individual genes was employed to set up our basic classification of isolates that did or did not contain pathogenicity islands. In parallel, DNA of the same isolates was used in the conventional random amplified polymorphic DNA (RAPD) PCR analysis, and the isolates were also genotyped (*cagA*, *vacA*, *iceA*, and *babA* status) and their proteomic maps were obtained by means of unidimensional SDS polyacrylamide gel electrophoresis (1D-SDS-PAGE) coupled with identification using peptide mass fingerprinting by MALDI-TOF mass spectrometry. A statistically significant correlation (coefficient of correlation $r = 0.25$, $p = 0.005$) was observed between the results of genome scanning and 1D-SDS-PAGE. No correlation was found between RAPD-PCR typing and genome scanning.

Key words: *Helicobacter pylori*, molecular typing, genome scanning, DNA array, RAPD analysis, proteomic map

Classification of microbial strains within a species has become an important task in modern applied microbiology. This task is due to the variability and plasticity inherent in bacteria and is particularly urgent for human pathogens. The typing of reference strains together with clinical isolates reveals the genetic and metabolic traits responsible for pathogenicity and virulence and also allows elaborating specific approaches for prognosis and treatment of infectious diseases. Recent advances in structural and functional genomics of microorganisms has led to the establishment of a large depository of typing methods which include the RAPD-PCR analysis, subtractive hybridization, and multilocus sequencing as well as genome scanning on hybridization DNA arrays.

Helicobacter pylori is a Gram-negative bacterium that colonizes the human stomach, mainly persisting within the gastric mucosal layer. The infection by *H. pylori* leads to active gastritis, which may progress to chronic gastritis, gastric ulcer, or duodenal ulcer. Besides, *H. pylori* infection is associated with mucosa-associated lymphoid tissue (MALT) lymphoma and gastric adenocarcinoma [1, 2]. Clinical studies have allowed classifying *H. pylori* as an overt carcinogen [2]. Epidemiological studies indicate that infection by *H. pylori* is one of the most common in the world, with 20 to 100% of the adult population infected in different regions of the world [3].

Efforts to classify *H. pylori* isolates by means of the above-mentioned techniques have revealed their heterogeneity [4-6]. In particular, the pathogenicity island (PAI) was absent from many isolates. There is evidence suggesting that PAI may be associated with a more severe course of the disease, for example, with a higher risk of gastric adenocarcinoma [7]. Analysis of complete genome sequences of two PAI-containing *H. pylori* strains, 26695 and J99 [8, 9], has shown that in the overall genome organization and the order of genes the two strains are fairly similar. Only from 6 to 7% of the genes are strain-specific. These strain-specific genes are believed to determine the resistance to antibiotics, cell surface structure, and virulence and pathogenicity [4].

The present work is aimed at comparing the different methods of molecular typing using as an example some classical *H. pylori* isolates obtained in different Russian regions.

* To whom correspondence should be addressed.

0006-2979/04/6905-0536 ©2004 MAIK "Nauka/Interperiodica"

Our main classification of isolates that did and did not contain the pathogenicity island is based on DNA macroarray genome scanning using individual genes. In parallel, we examined the same DNA samples using conventional RAPD-PCR, genotyping *cagA*, *vacA*, *iceA*, and *babA* genes and constructed their proteomic maps by means of 1D-SDS-PAGE combined with trypsin fragment identification using MALDI-TOF mass spectrometry.

MATERIALS AND METHODS

Strains and cultivation. A total 17 *H. pylori* isolates were obtained from different locations of Russia, including Moscow and Moscow Region ($n = 9$), St. Petersburg ($n = 4$), Kazan ($n = 2$), and Novosibirsk ($n = 2$). Each isolate was obtained from patients with duodenal ulcer and/or chronic gastritis. Upper endoscopy was performed and multiple gastric biopsy specimens were taken from the antrum.

Freshly retrieved biopsy specimens were transported to the laboratory for culture on the same day. *H. pylori* was cultured by rubbing gastric biopsy specimens onto Pylori agar (Biomerieux, France). The plates were incubated under microaerobic conditions (90% N_2, 5% O_2, 5% CO_2) at 37°C for 4 to 7 days. The plates were observed at days 3 and 7. Suspected colonies were tested by gram staining, as well as for urease, oxidase, and catalase activities [10]. If all the tests gave positive results, the culture was identified as *H. pylori*.

PCR analysis and genotyping of virulence genes. DNA isolation from biopsy material and from clinical isolates of *H. pylori* as well as *H. pylori* assay was done using the Helicopol II kit (Lytech, Russia).

The primers used for genotyping virulence genes *cagA*, *vacA*, *iceA*, and *babA2* have been described earlier [11-14]. The PCR reaction mixture (30 μl) contained 1 μg *H. pylori* DNA, 50 mM KCl, 10 mM Tris-HCl (pH 8.8 at 25°C), 1.5 mM $MgCl_2$, 0.08% Nonidet P-40, 200 μM of each deoxynucleotide triphosphate (Fermentas, Lithuania), 20 pM of each primer, and 2.5 U of Taq DNA polymerase (Fermentas). The following PCR conditions were used for *vacA*, *cagA*, and *iceA*: denaturation; 94°C for 5 min; then 35 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 2 min; and final extension at 72°C for 5 min in a Perkin-Elmer (USA) GeneAmp PCR System 2400. The conditions for *babA2* were the same except for annealing at 55 instead of 52°C and extension at 72°C for 1 min instead of 2 min.

The PCR products were separated in 2% agarose gels containing 0.5 μg/ml ethidium bromide in 1× TAE buffer (Trevigen, USA). The results were recorded using a GelDoc 1000 (BioRad, USA) video system and then processed by means of the Molecular Analyst Software (BioRad).

RAPD analysis. This was performed as described in the literature [15, 16] with certain modifications. PCR fragments were analyzed in 1% agarose gels, and RAPD-PCR fragments were studied with the aid of the Gene Profiler 4.03 software (Scanalytics, Inc., USA). Cluster analysis was done using the TreeCon program [17].

Genome scanning. For DNA/DNA hybridization, high density nylon membranes (Eurogentec, Belgium) carrying immobilized 1564 PCR products, in duplicate, out of the 1590 ORFs of *H. pylori* strain 26625. Chromosomal DNA (25 ng) was purified from clinical isolates using the genomic Wizard Genomic DNA Purification Kit (Promega, USA), labeled with [α-^{33}P]dATP, and used as a probe in hybridization. The DNA arrays were exposed for two days with a storage phosphor screen (Molecular Dynamics, USA) and then scanned by the Storm 820 (Molecular Dynamics) at a resolution of 50 μm and color depth of 16 bits. The images were studied by means of the ImageQuant software 5.1 (Molecular Dynamics).

The background value was found as a median of signals around the spot site. An algorithm described earlier [18] was used to normalize individual spot intensities. Genomic DNA for each isolate was obtained from two independent cultivations of bacteria and then used for independent DNA labeling and hybridization. The values of normalized signal intensity were computed based on two independent replicas. Genomic DNA of *H. pylori* strain 26695 kindly provided by Prof. D. Tallibi (Eurogentec) was used as a control. The gene was considered absent if the signal intensity was less than 10% in the control hybridization. When the intensity was between 10 and 80% of the control, the result was checked using PCR and primers for each ORF studied. For cluster analysis, the genome scanning results were converted into a binary code (gene present, 1; gene absent, 0). The results were processed with the J-express 2.1 software (http://www.molmine.com/).

1D-SDS-PAGE. The cells were collected from the solid medium and washed twice with ice-cold phosphate-buffered saline (50 mM sodium phosphate, 0.15 M sodium chloride, pH 7.2) containing 1 mM protease inhibitor phenylmethylsulfonyl fluoride (PMSF; Sigma, USA), and resuspended in distilled water at a protein concentration in the sample equal to 50 μg/μl. After mixing the samples with a sample buffer (100 mM DTT, 4% (w/v) SDS, 20% (v/v) glycerol) [19] at a 1 : 2 ratio and keeping for 5 min at 95°C, they were applied to the gel at 100 μg protein per lane.

SDS-PAGE was run on a Protean II xi Cell (BioRad) using 9-16% vertical gradient gels (30.8% T, 2.6% C), 20 × 20 cm and 1.5 mm thick, prepared according to the standard protocol [20]. Coomassie Brilliant Blue R-250 staining was performed as described earlier [21].

Identification of proteins. Proteins were identified by mass-spectrometry of fingerprints of their trypsin-gener-

Table 1. Clinical isolates of *H. pylori*: the origin and DNA-typing by four virulence genes

H. pylori strain	Host's condition	District	Genotype			
			vacA	*cagA*	*iceA*	*babA*
8	DU	Moscow	*s1/m1/m2*	+	*iceA1/A2*	+
9	CG	Kazan	*s1/m1*	+	*iceA2*	+
19	DU	Moscow	*s2/m2*	—	*iceA1*	—
33	CG	Kazan	*s1/m1*	+	*iceA1*	+
36	CG	Novosibirsk	*s2/m2*	—	*iceA1*	—
37	DU	Moscow	*s2/m2*	—	*iceA1*	—
46	DU	Moscow	*s1/m2*	+	—	+
50	unknown	Moscow	*s1/m2*	+	*iceA1*	+
52	CG	St. Petersburg	*s1/m2*	+	*iceA1*	+
53	CG	St. Petersburg	*s1/m1*	+	*iceA1*	+
57	CG	St. Petersburg	*s1/m1/m2*	+	*iceA1*	+
58	DU	Moscow	*s1/m2*	+	—	+
59	CG	St. Petersburg	*s1/m1/m2*	+	*iceA1/A2*	+

Note: DU, duodenal ulcer; CG, chronic gastritis.

ated fragments. In-gel digestion by modified porcine trypsin (Promega) was done according to a published protocol [22].

MALDI-TOF mass spectrometry was performed in a Bruker (Germany) Reflex III instrument, and the proteins were identified as described earlier [23].

Statistical analysis. The RAPD-PCR fingerprints and the protein SDS-PAGE electrophoregrams were analyzed using the SPSS 11 Demo software (SPSS Inc., USA). The bands obtained after protein separation were coded in the binary mode (presence of band, 1; absence of band, 0). For cluster analysis, the matrix of squares of Euclidean distances was used. Final dendrograms were calculated by the Ward method [24].

Correlation between the matrices of distances was cross-analyzed by the non-parametric Mantel test [25, 26]. For each test, 10,000 iterations and randomization of one matrix were used.

RESULTS

Genotyping. The genotyping of 13 clinical isolates according to the four virulence genes (*cagA*, a marker for genomic pathogenicity island; *babA*, mediator of adhesion of *H. pylori* on gastric epithelium; the vacuolating cytotoxin *vacA*, and the recently discovered *iceA*, *i*nduced by *c*ontact with *e*pithelium) has revealed two groups of

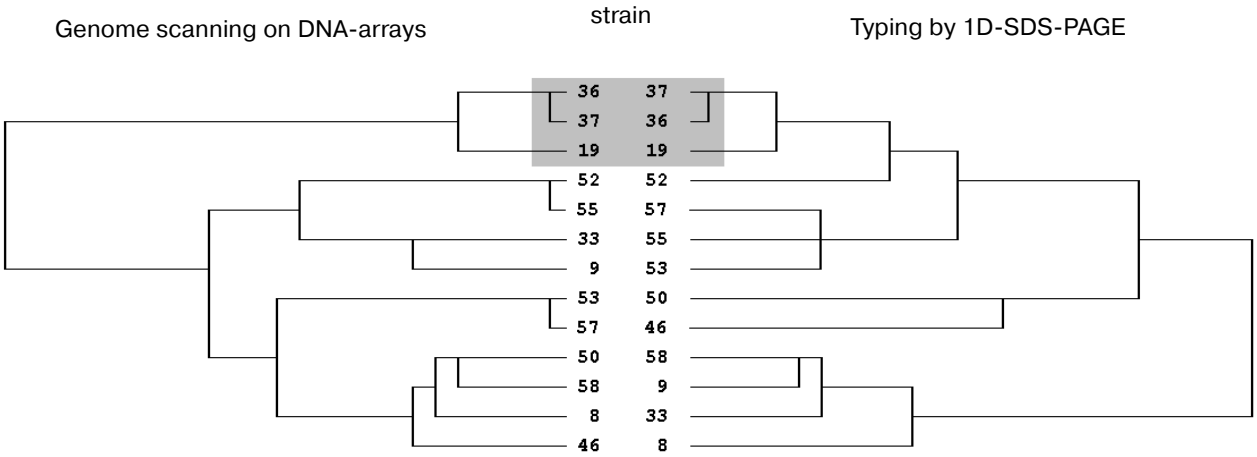

Fig. 1. Clustering dendrograms for typing *H. pylori* isolates by means of genome scanning and 1D-SDS-PAGE.

strains that were obvious without using special mathematical techniques (see Table 1). One group includes strains 19, 36, and 37 whose genome lacks *babA* as well as *cagA*, and hence the entire PAI. These strains also contain the *vacA* s2/m2 and *iceA* 1 alleles. The 10 remaining strains are consequently *cagA*- and *babA*-positive.

Genome scanning. Genome scanning of the strains using DNA-arrays revealed heterogeneity in 309 loci out of the total 1565 (19.7%). In this work, we did not intend to discuss the composition of these genes, but it is worth noting that our DNA scanning data in general agree with those obtained earlier [4]. The table of DNA scanning results is available at http://www.imbh.msk.su/proteomics/helicobacter. When classifying, strains 19, 36, and 37 were found in a separate cluster like in the case of genotyping (Fig. 1). One of the reasons for that is the absence in them of a large PAI covering over 20 genes.

RAPD-PCR analysis. Clustering of RAPD-PCR results obtained from 13 isolates with four pairs of random primers is presented in Fig. 2 as dendrograms, each of them corresponding to a pair of primers. It is noteworthy that *cagA*-negative isolates fail to produce a single cluster in all the four cases. In three cases (Fig. 2, RAPD 1-3), strains 36 and 37 but not strain 19 fall into a single cluster.

Analysis of proteomic maps. The clinical isolates of *H. pylori* have been characterized by 1D-SDS-PAGE (Fig. 3). For one of the strains (No. 46), proteomic inventory has been made and 71 individual proteins have been identified (the list is available at http://www.ibmh.msk.su/proteomics/helicobacter). This number, as expected, is lower than that identified by means of 2D-electrophoresis (126 gene products) described earlier [27]. The 1D-proteomic maps have been classified without the use of identification but only based on 1D-SDS-PAGE (Fig. 1).

Correlation analysis. The results of the non-parametric Mantel test used to determine correlation between the dendrograms obtained in classification of 13 isolates using DNA scanning, four versions of RAPD-PCR, and 1D-SDS-PAGE are given in Table 2. A statistically significant correlation (coefficient of correlation $r = 0.25$, $p = 0.05$) has been noted between genome scanning and 1D-SDS-PAGE. No correlation was found between the dendrograms in the rest of the cases.

DISCUSSION

Our results are indicative of the degree to which the existing methods of strain typing reflect the real relation between strains of the bacterium. A comparison of the clustering results obtained with different primer sets in RAPD-PCR as well as in genome scanning and protein 1D-SDS-PAGE provides information on relative applicability of these methods for typing *H. pylori* strains.

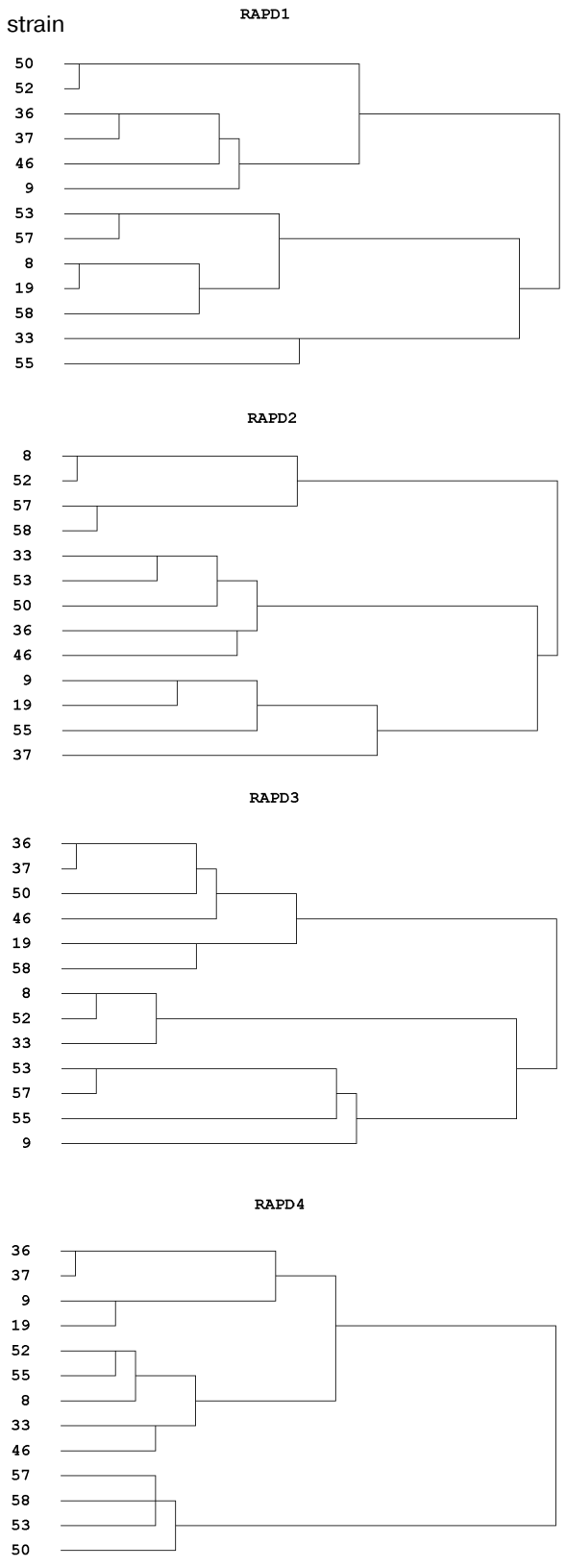

Fig. 2. Clustering dendrograms for typing *H. pylori* isolates by means of RAPD-PCR with four pairs of random primers (RAPD 1-4).

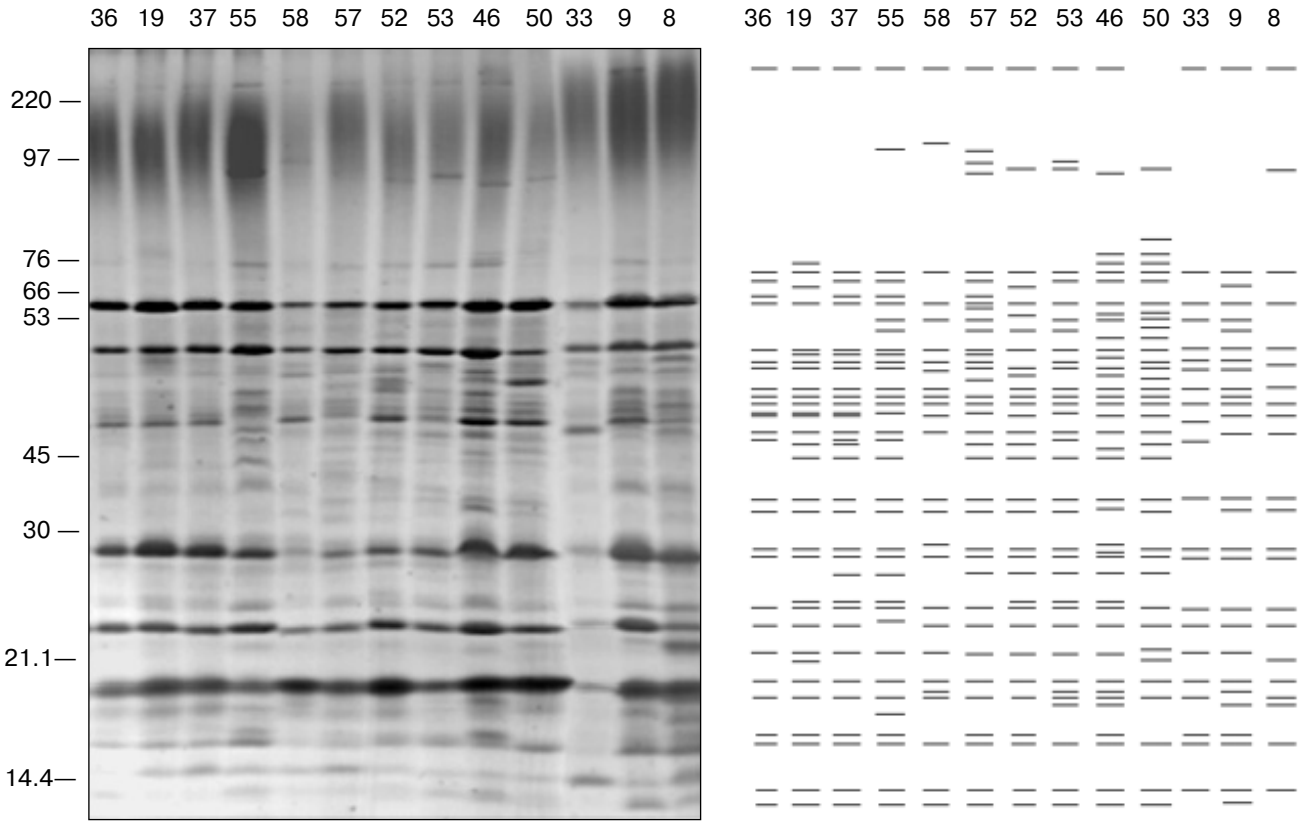

Fig. 3. 1D-SDS-PAGE of proteins from *H. pylori* isolates (left) and a schematic representation of the bands obtained (right). Ordinate, molecular masses (kD) of standard proteins; electrophoretic lanes correspond to strain numbers.

The choice of distance matrix correlation as an objective indicator of correct typing seems to be justified. Obviously, if clustering is made adequately, the same strains in different dendrograms should fall into identical clusters. The use of the non-parametric Mantel test has offered evidence on statistical significance of the correlation coefficients.

We conclude that strains having similar characteristics of virulence or pathogenicity may fail to become clustered together when typing based on RAPD-PCR is employed, and thus the applicability of this method is limited. The method of typing that is based on protein separation offers statistically significant correlation with genotyping. The fact that PAI-negative strains fell into one cluster suggests that the proteomic typing provides adequate results against the background of various traits that a strain may carry. The genomic and proteomic typing gave similar results even in spite of the frequent cases when the protein separation pattern did not show the products of the genes that were instrumental in determining the joint clustering of the PAI-negative strains in genome scanning. This particularly concerned the proteins of the virulence island PAI. There were stained bands corresponding to the protein molecular mass range of markers of this island (100-150 kD) in all protein separation patterns, including those of PAI-minus strains (Fig. 3), suggesting that this proteomic map region does not significantly affect clustering results.

In spite of this, one of the proteins identified in that mass range from PAI-plus strains was CagA. On the other hand, chief components corresponding to PAI-minus strains 19, 36, and 37 were DNA polymerase III (DnaE) α-subunit and type I restriction enzyme (HsdR) protein. At the same time, not a single peptide could be found that had relevance to cleavage products of the CagA protein (data not shown).

Therefore, the genomic relation between *H. pylori* strains has an indirect effect on relation between their proteomic 1D-SDS-PAGE maps. Such genome–proteome co-variability is a result of structural and functional integrity of the microorganism, i.e., a situation when similar alteration in the genome becomes reflected in a transformed but similar manner in the cell protein composition.

In summary, it may be said that the typing method based on a conventional and affordable SDS-PAGE procedure is indeed informative with respect to the genetic relation among different strains of bacteria (*H. pylori* in

Table 2. Correlation between distance matrices of binary-coded densitograms (1D-SDS-PAGE, RAPD-PCR analysis with four pairs of random primers) and the results of genome scanning of *H. pylori* clinical isolates*

Pair of distance matrices	*r*	*p*
Genome scanning – proteome (1D-SDS-PAGE)	**0.25**	**0.025**
Proteome – RAPD1	0.09	0.31
Proteome – RAPD2	0.17	0.17
Proteome – RAPD3	0.14	0.22
Proteome – RAPD4	–0.17	0.85
Genome scanning – RAPD1	0.15	0.11
Genome scanning – RAPD2	–0.17	0.9
Genome scanning – RAPD3	0.01	0.52
Genome scanning – RAPD4	0.15	0.11
RAPD1 – RAPD2	–0.08	0.69
RAPD1 – RAPD3	0.02	0.48
RAPD1 – RAPD4	0.13	0.17
RAPD2 – RAPD3	0.17	0.16
RAPD2 – RAPD4	0.12	0.17
RAPD3 – RAPD4	0.06	0.32

* Statistically significant correlation is highlighted; *r* is a correlation coefficient; *p* is a significance level (1 is probability).

our case) and can be used for that purpose even without protein identification by means of mass spectrometry.

This work was supported by the program "Proteomics for Medicine" of Ministry of Industry, Science, and Technology of the Russian Federation. The work of S. A. Moshkovskii is supported in part by the National Medicine Assistance Fund (Russia).

REFERENCES

1. El-Omar, E. M., Chow, W. H., and Rabkin, C. S. (2001) *Gastroenterology*, **121**, 1002-1004.
2. Meining, A., Riedl, B., and Stolte, M. (2002) *J. Clin. Pathol.*, **55**, 770-773.
3. Poundeyr, R. E., and Ng, D. (1995) *Aliment. Pharmacol. Ther.*, **9** (Suppl. 2), 33-39.
4. Salama, N., Guillemin, K., McDaniel, T. K., Sherlock, G., Tompkins, L., and Falkow, S. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 14668-14673.
5. Israel, D. A., Salama, N., Krishna, U., Rieger, U. M., Atherton, J. C., Falkow, S., and Peek, R. M., Jr. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 14625-14630.
6. Israel, D. A., Salama, N., Arnold, C. N., Moss, S. F., Ando, T., Wirth, H. P., Tham, K. T., Camorlinga, M., Blaser, M. J., Falkow, S., and Peek, R. M., Jr. (2001) *J. Clin. Invest.*, **107**, 611-620.
7. Nomura, A. M., Lee, J., Stemmermann, G. N., Nomura, R. Y., Perez-Perez, G. I., and Blaser, M. J. (2002) *J. Infect. Dis.*, **186**, 1138-1144.
8. Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., Venter, J. C., et al. (1997) *Nature*, **388**, 539-547.
9. Alm, R. A., Ling, L. S., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., Smith, D. R., Noonan, B., Guild, B. C., deJonge, B. L., Carmel, G., Tummino, P. J., Caruso, A., Uria-Nickelsen, M., Mills, D. M., Ives, C., Gibson, R., Merberg, D., Mills, S. D., Jiang, Q., Taylor, D. E., Vovis, G. F., and Trust, T. J. (1999) *Nature*, **397**, 176-180.
10. On, S. L., and Holmes, B. (1992) *J. Clin. Microbiol.*, **30**, 746-749.
11. Cover, T. L., and Blaser, M. J. (1992) *J. Biol. Chem.*, **267**, 10570-10575.
12. Van Doorn, L. J., Figueiredo, C., Sanna, R., and Pena, S. (1998) *J. Clin. Microbiol.*, **36**, 2597-2603.
13. VanDoorn, L. J., Figueiredo, C., Sanna, R., Plaisier, A., Schneeberger, P., deBoer, W., and Qoit, W. (1998) *Gastroenterology*, **115**, 58-66.
14. VanDoorn, L. J., Henskens, Y., Nouhan, N., and Verschuuren, A. (2000) *J. Clin. Microbiol.*, **38**, 13-17.
15. Akopyants, N., Bukanov, N. O., Westblom, T. U., Kresovich, S., and Berg, D. E. (1992) *Nucleic Acids Res.*, **20**, 5137-5142.
16. Van der Ende, A., Pan, Z. J., Bart, A., van der Hulst, R. W., Feller, M., Xiao, S. D., Tytgat, G. N., and Dankert, J. (1998) *Infection Immunity*, **66**, 1822-1826.
17. Van de Peer, Y., and de Wachter, R. (1997) *Comput. Appl. Biosci.*, **13**, 227-230.
18. Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J. D., Volker, U., and Hecker, M. (2001) *J. Bacteriol.*, **183**, 5613-5617.
19. Laemmli, U. K. (1970) *Nature*, **227**, 680-685.
20. Goerg, A., and Weiss, W. (1999) in *2-D Proteome Analysis Protocols (Methods in Molecular Biology)* (Link, A. J., ed.) Vol. 112, Humana Press, Totowa, New Jersey, pp. 235-244.
21. Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988) *Electrophoresis*, **9**, 255-262.
22. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Analyt. Chem.*, **68**, 850-858.
23. Govorun, V. M., Moshkovskii, S. A., Tikhonova, O. V., Goufman, E. I., Serebryakova, M. V., Momynaliev, K. T., Lokhov, P. G., Khryapova, E. V., Kudryavtseva, L. V., Smirnova, O. V., Toropyguine, I. Yu., Maksimov, B. I., and Archakov, A. I. (2003) *Biochemistry (Moscow)*, **68**, 42-49.
24. Ward, J. (1963) *J. Am. Stat. Assoc.*, **58**, 236-244.
25. Sokal, R. R., and Rohlf, F. J. (1995) *Biometry*, 3rd Edn., W. H. Freeman and Company, New York.
26. Mantel, N. (1967) *Cancer Res.*, **27**, 209-220.
27. Jungblut, P. R., Bumann, D., Haas, G., Zimny-Arndt, U., Holland, P., Lamer, S., Siejak, F., Aebischer, A., and Meyer, T. F. (2000) *Mol. Microbiol.*, **36**, 710-725.