© 1992 Oxford University Press

Nucleic Acids Research, 1992, Vol. 20, No. 23 6221–6225

PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*

Natalia Akopyanz, Nickolai O.Bukanov, T.Ulf Westblom[1] and Douglas E.Berg*
Department of Molecular Microbiology, Campus Box 8230, Washington University Medical School, St Louis, MO 63110 and [1]Division of Infectious Diseases and Immunology, St Louis University School of Medicine, St Louis, MO 63104, USA

Received September 28, 1992; Accepted November 2, 1992

ABSTRACT

DNA sequence diversity among 60 independent isolates of the gastric pathogen *Helicobacter pylori* was assessed by testing for restriction fragment length polymorphisms (RFLPs) in several PCR-amplified gene segments. 18 MboI and 27 HaeIII RFLPs were found in the 2.4 kb *ureA-ureB* (urease) segment from the 60 strains; this identified 44 separate groups, with each group containing one to four isolates. With one exception, each isolate not distinguished from the others by RFLPs in *ureA-ureB* was distinguished by MboI digestion of the neighboring 1.7 kb *ureC-ureD* segment. The 1.5 kb *flaA* (flagellin) gene, which is not close to *ure* gene cluster, was also highly polymorphic. In contrast, isolates from initial and followup biopsies yielded identical restriction patterns in each of the three cases tested. The potential of this method for detecting population heterogeneity was tested by mixing DNAs from different strains before amplification: the arrays of restriction fragments obtained indicated co-amplification from both genomes in each of the five pairwise combinations tested. These results show that *H. pylori* is a very diverse species, that indicate PCR-based RFLP tests are almost as sensitive as arbitrary primer PCR (RAPD) tests, and suggest that such RFLP tests will be useful for direct analysis of *H. pylori* in biopsy and gastric juice specimens.

INTRODUCTION

Helicobacter pylori is a fastidious, slow growing, microaerophilic bacterium whose long-term carriage in the upper gastrointestinal tract is implicated in the pathogenesis of chronic active gastritis, peptic ulcers and gastric carcinoma (1,2). Efficient, reliable and potentially automatable methods of strain identification are needed to facilitate epidemiological studies and analyses of mechanisms of *H. pylori* infection and long term carriage. Inspection of plasmid DNAs and profiles of total genomic restriction digests showed that clinical isolates could differ significantly, and more sensitive Southern blotting with ribosomal DNA probes ('ribotyping') has distinguished many individual strains (3–6). These methods require considerable amounts of DNA (about 2 μg per test), which is an impediment in large scale studies, since *H. pylori* is difficult and costly to grow. In addition, the genomic DNAs of some isolates are partially or completely resistant to several commonly used restriction enzymes (6, 7), and this might lead to occasional misclassification based on the seeming absence of a restriction site.

Polymerase chain reaction (PCR)-based methods offer an attractive alternative to chromosomal fingerprinting methods, because they are sensitive and fast, require much less DNA, and are not affected by in vivo DNA modification. One method uses arbitrary primers and results in strain-specific arrays of anonymous DNA fragments that sample the genome in its entirety (termed 'RAPD' or arbitrary primer PCR) (8, 9). We found that each of 60 independent *H. pylori* isolates yielded a unique RAPD fingerprint (10), which illustrates the high sensitivity of this method. One needs pure cultures for the RAPD method, however, because all DNAs in a reaction mix are sampled randomly. This is an impediment for some studies of *H. pylori*, because culturing it from clinical samples is laborious and not always successful. RAPD PCR on such clinical samples would yield anonymous bands from host cells and other microbes, as well as from the target genome (see ref. 8).

H. pylori specific DNAs can be amplified from biopsies and from gastric juice (11–13), however, and given sufficient sequence diversity, such PCR-amplified DNAs should allow strains to be distinguished by restriction fragment length polymorphisms (RFLPs). Preliminary tests using the 2.4 kb *ureA-ureB* gene segment from 22 isolates had revealed ten different HaeIII RFLPs, although the MboI fragment pattern was said to be constant (14). Our results indicate that PCR-based RFLP fingerprinting can distinguish individual strains of *H. pylori* efficiently, and with a sensitivity close to that of the RAPD method, and that PCR-based RFLP analysis should be useful in detecting mixed *H. pylori* populations.

MATERIAL AND METHODS

The *H. pylori* strains used were NCTC11637 and NCTC11638 from Australia, P466 from Peru (gift of Dr. R.H.Gilman), MO19 from Missouri, and 63 'WV' strains (Table 1, below) collected by one of us (TUW) at the Veterans Administration Medical

* To whom correspondence should be addressed

Center in Huntington, West Virginia (10). Bacteria were grown to stationary phase under microaerophilic conditions as described (15), and PCR amplification of specific fragments was carried out either on about 1 ng of phenol extracted DNA (10), or on 0.1 μl of stationary phase cultures that had been boiled for 4 min. The oligonucleotides used for PCR amplification of specific segments are: 5'AGGAGAATGAGATGA and 5'ACTTTATTGGCTGGT, for *ureA-ureB* (14); 5'TGGGACTGATGGCGTGAGGG and 5'ATCATGACATCAGCGAAGTTAAAAATGG for *ureC-ureD* (16); and 5'ATGGCTTTTCAGGTCAATAC and 5'GCTTAAGATATTTTGTTGAACG for *flaA* (17). PCR amplification was carried out in 100 μl, using boiled cells or DNA, 0.2 U Amplitaq polymerase, standard PCR buffer (Cetus), and 20 pmoles of each primer. The reaction was carried out in a Perkin Elmer 480 thermal cycler programmed for 30 cycles as follows: 94°C, 1 min (denaturation); 50°C (*ureA-ureB*) or 60°C (*flaA* or *ureC-ureD*) (annealing); and then 72°C, 2 min (extension). After amplification, DNA products were ethanol precipitated, the pellets were washed with 70% ethanol, and resuspended in 25 μl of sterile distilled water, and 3 μl of product was electrophoresed on a 1% agarose gel to ensure homogeneity and yield. PCR amplification resulted in a homogeneous DNA fragment of the size expected (2.4 kb for *ureA-ureB*; 1.7 kb for *ureC-ureD*; and 1.5 kb for *flaA*) with each strain tested. Usually 10 μl of product was digested with 5 to 10 U of restriction enzyme for one hour in the buffer recommended by the supplier (NE Biolabs), and electrophoresed in a gel consisting of 3.5% NuSieve GTG agarose and 1.5% Seakem LE agarose (FMC Corporation).

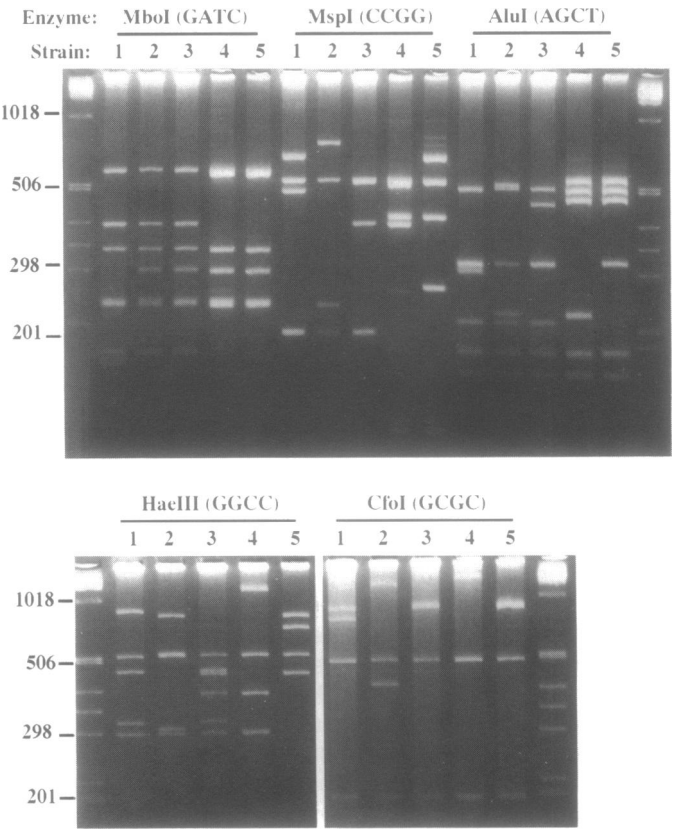

Figure 1. RFLP analysis of the PCR-amplified 2.4 kb *ureA-ureB* segment of *H. pylori*. Strains 1 though 5 are NCTC11638, WV229, NCTC11637, P466 and MO19, respectively.

RESULTS

Diversity detected by RFLP analysis

The diversity of sequences in known genes of *H. pylori* was assessed by restriction analysis of the 2.4 kb *ureA-ureB* segment, amplified from five strains from three continents. We found five distinct patterns with HaeIII, MspI and AluI, four patterns with CfoI, and three patterns with MboI (Fig. 1). Retests of the *ureA-ureB* segment from isolates that were initially said to yield just one MboI pattern (14) also revealed multiple MboI RFLPs (18).

The diversity of *H. pylori* gene sequences was tested further using a set of 60 independent isolates from patients at one hospital: 18 MboI patterns, and also 27 HaeIII patterns were observed in the *ureA-ureB* segment (Fig. 2); most pairs of strains that were matched in MboI patterns were distinguishable by HaeIII patterns, and vice versa, and 44 distinct MboI-HaeIII fingerprint groups containing one to four members were identified (Table 1). The adjoining *ureC-ureD* segment from the 25 isolates in the groups containing at least two members was amplified and digested with MboI: all but two isolates (WV230 and WV231) yielded a different MboI pattern (Fig. 3, for representative data). The 1.5 kb *flaA* gene, which is not near the urease gene cluster (19), was also tested after PCR amplification from 14 randomly chosen isolates: eight MboI and 11 HhaI RFLPs were found (data not shown).

Sequential isolates from the same patient

Isolates from initial and followup biopsies were available for three patients. In each case, no difference was found between initial and followup biopsy isolates in the RFLPs in any of the three target genes tested (see Fig. 4). Appropriate negative controls ruled out possible PCR artifacts such as DNA contamination. This outcome is in accord with RAPD tests of the same isolates (10), and indicates persistent carriage of the original strain despite treatment, not re-colonization by different strains after eradication of the organism.

Detection of mixed populations

Given the diversity of restriction sites in *H. pylori* genes, RFLP analysis should reveal bacterial heterogeneity, as would result from mixed infections with different strains. Heterogeneity would be signalled by either of two results: (i) by restriction fragment sizes that, when added together, exceed that of the starting PCR fragment; or (ii) unequal yields of different restriction fragments, since only a subset of fragments are common to most strains (see Fig. 2). This test for heterogeneity assumes that target DNA segments will be amplified from divergent strains with similar efficiency with a given set of primers. To test the validity of this assumption, we mixed genomic DNAs from five strains in all pairwise combinations, amplified the *ureA-ureB* segment, and digested the segment with HaeIII and MboI. Comparison of RFLP patterns from mixtures and pure samples demonstrated co-amplification in each case (Fig. 5). Interestingly, extra bands were obtained from one of the mixtures (strains 1 and 2; white arrows). These can be attributed to 'crossover PCR', the jumping of templates during amplification (see 20); such products do not detract from the promise of RFLP analysis for detecting population heterogeneity.

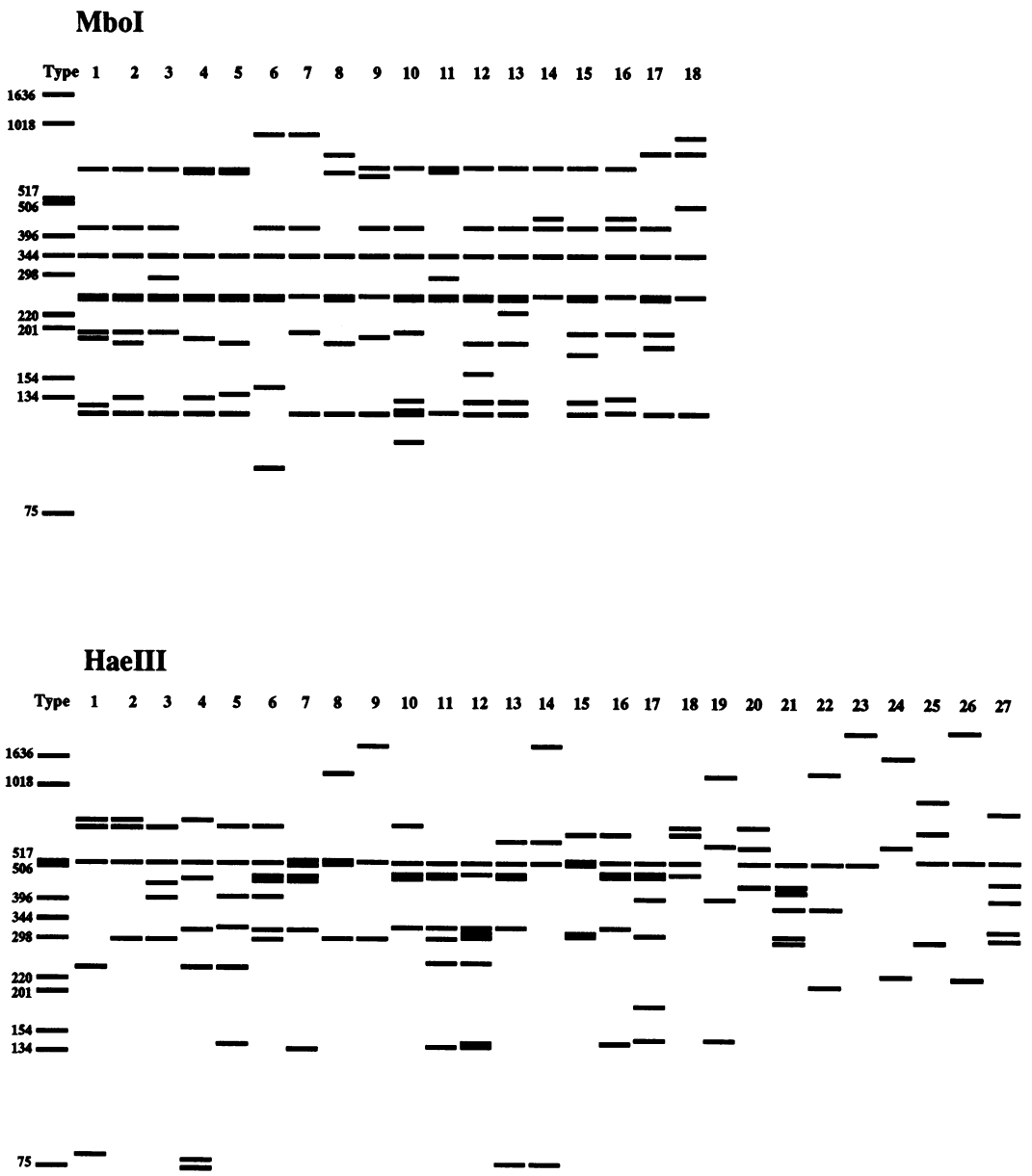

Figure 2. Summary tracings of profiles of *Mbo*I and *Hae*III digests of the *ureA-ureB* DNA segment amplified from WV strains and analyzed as in Fig. 1.

DISCUSSION

The many RFLPs found in PCR-amplified *H. pylori* gene segments from 60 independent strains indicates a high level of diversity in this gastric pathogen. In particular, digestion of the *ureA-ureB* gene segment with MboI and with HaeIII divided these 60 isolates into 44 distinct groups. With only one exception, each isolate not distinguished from all others by MboI and HaeIII RFLPs in *ureA-ureB* was distinguished by MboI RFLPs in the neighboring 1.7 kb *ureC-ureD* segment. The 1.5 kb *flaA* gene, which is not close to the *ure* gene cluster, was also highly polymorphic, but was not needed to discriminate among the strains tested. In contrast to the diversity of strains from different patients, isolates obtained from initial and followup biopsies of the same patient were indistinguishable in the three cases tested.

WV230 and WV231, the only two strains not distinguished by RFLP tests, were isolated on the same day, but were nominally

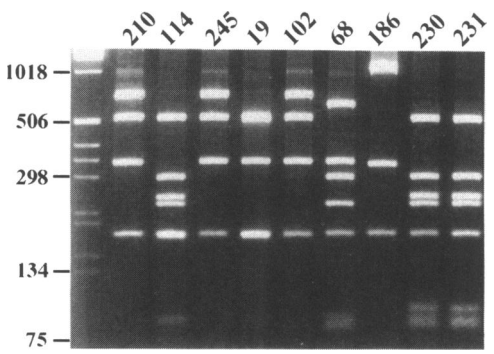

Figure 3. Representative MboI digests of the 1.7 kb *ureC-ureD* segment from WV strains that were not distinguished by RFLP analysis of the *ureA-ureB* segment (see Table 1).

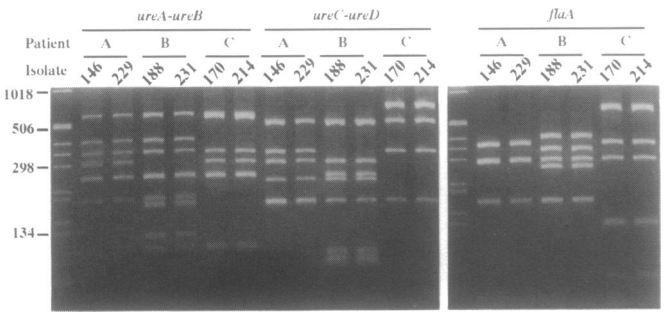

Figure 4. MboI restriction analysis of pairs of strains from initial and second (followup) biopsies from three patients. Initial and followup biopsies also resulted in identical RFLP patterns after HaeIII digestion of the *ureA-ureB* and HhaI digestion of the *ureC-ureD* and *flaA* gene segments (data not shown).

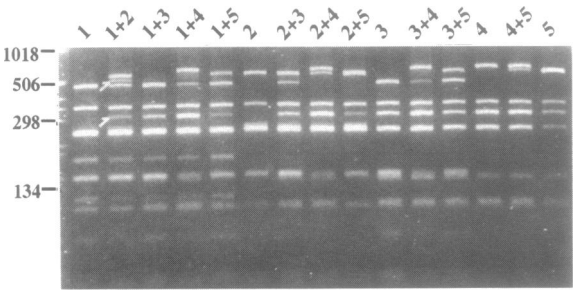

Figure 5. Detection of mixed populations by PCR-based RFLP analysis. Equal amounts of genomic DNAs from the indicated strains (1 through 5, as in Fig. 1) were mixed, and used for amplification of the 2.4 kb *ureA-ureB* segment. The resultant fragments (and control fragments amplified from pure instead of mixed DNA samples) were digested with MboI and HaeIII, and electrophoresed. The arrows in the mixture 1+5 indicate bands that are not present in the profiles of parental strains 1 or 5, and thus may represent crossover PCR products (see text; ref 20).

independent (from unrelated patients). It is possible, however, that they actually came from the same patient, perhaps by carryover on a contaminated endoscope (see ref. 21). Small, reproducible differences were seen in the profiles generated from these two isolates by arbitrary primer (RAPD) PCR with three of six primers tested (7,10). These differences suggest divergence during long-term carriage, and thus suggest that the RAPD method can be somewhat more sensitive than RFLP analysis, which focuses on only one or a few specific segments.

It is worth emphasizing the power of PCR-based RFLP analysis, relative to ribotyping or pulse field gel fingerprinting methods. Only minute amounts of DNA (< 0.1 ng) are sufficient for PCR-based RFLP tests. In contrast a typical genomic fingerprinting assay uses several micrograms of DNA. This 10,000-fold difference is of particular significance for organisms whose growth is slow or requires costly media or facilities, where fast strain-identification is needed to help guide therapy, or where high virulence mandates use of very small cultures to safeguard laboratory workers. The ability to generate informative fingerprints from less than a microliter of culture should also contribute to the automation of DNA fingerprinting protocols. Given recent demonstrations of PCR amplification of *H. pylori* sequences directly from stomach biopsies and gastric juice specimens (11–13), PCR-based RFLP tests should allow direct use of such clinical samples for strain identification, for example: (i) to trace patterns of infection in families and larger populations; (ii) to examine persistence of individual strains during long term carriage and antimicrobial therapy; (iii) to identify *H. pylori* mixtures that could reflect co-existence, or secondary infection; and (iv) to detect horizontal gene transfer and recombination, which could contribute to the development of genotypes with increased virulence for particular hosts.

Several factors might contribute to the great diversity among *H. pylori* isolates observed here and in our earlier RAPD analyses (10) *H. pylori* infections are infrequent ($\sim 1\%$ per year) in societies with adequate sanitation, but the organism is carried for decades once acquired. Since most patients in the present study were over 50 years old, and former members of the US Armed Forces, they may have been infected many years before biopsy and in many different geographical locations. The present results might thus reflect the diversity of *H. pylori* world-wide, the relative absence of competition between clones in a single host, due to the rarity of person-to-person spread, and evolutionary changes including periodic selection (22) during long term carriage.

ACKNOWLEDGMENTS

We are grateful to Drs. C.M.Berg, S.Drazek, P.Foxall and H.Mobley for stimulating discussions, to P.Foxall and H.Mobley for pemission to cite unpublished results, and to J.Dunford-Shore and B.R.Midkiff for skillful technical assistance. This work was supported by grants DEFGO2-89ER60862 from the US Department of Energy, and by grants HG00563 to D.E.B. and RR05870 to T.U.W. from the US Public Health Service.

REFERENCES

1. Blaser, M. J. (1992) Clin. Infect. Dis. 15:386–393.
2. Marshall, B. J., McCallum, R. W. and Guerrant. R. L. (1991) *Helicobacter pylori* in peptic ulceration and gastritis. Blackwell Scientific Publications, Boston, MA.
3. Oudbier, J. H., Langenberg, W., Rauws, E. A. J. and Bruin-Mosch, C. (1990) J. Clin. Microb. 28:559–565.
4. Owen, R. J., Bickley, J. , Costas, M., and Morgan, D. R. (1991). Scand. J. Gastroenterol. 26 (suppl), 43–50.
5. Simor, A. E., Shames, B., Drumm, B., Sherman, P., Low , D. E. and Penner, J. L. (1990) J. Clin. Microbiol. 28:83–86.
6. Tee, W., Lambert, J., Smallwood, R., Schembri, M., Ross, B. C. and Dwyer, B. (1992). J. Clin. Microbiol. 30:1562–1567.
7. Akopyanz, N., Bukanov, N. Westblum, T.U. and Berg, D. E. unpublished data.
8. Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. (1990). Nuc. Acids Res. 18:6531–6535.
9. Welsh, J. and McClelland, M. (1990) Nuc. Acids Res. 18:7213–7218.
10. Akopyanz, N., Bukanov, N. Westblum, T.U., Kresovich, S. and Berg, D. E. 1992. Nuc. Acids. Res. 20:5137–5142.
11. Hammar, M., T. Tyskiewicz, T. Wadström, and P. O'Toole. J. Clin. Microbiol. 30:54–58.
12. Clayton, C. L., H. Leanthous, P. J. Coates, D. D. Morgan and S. Tabaqchali. 1992. J. Clin. Microb. 30:192–200.
13. Westblom, T. U., Phadnis, S., Yang, P. and Czinn, S. J. 1992. Clin. Infect. Dis. in press.
14. Foxall, P. A., Hu , L.-T. and Mobley, H. L. T. (1992). J. Clin. Microb. 30:739–741.
15. Westblom, T. U. (1991) In Marshall, B. J., R. W. McCallum, and R. L. Guerrant (eds.) *Helicobacter pylori* in peptic ulceration and gastritis. Blackwell Scientific Publications, Boston, MA. pp. 81–91.
16. Labigne, A., Cussac, V., and Courcoux, P. 1991. J. Bacteriol. 173:1920–1931.

17. Leying, H. J., Suerbaum, C., Geis, G. and Haas, R. 1992. Molec. Microbiol. 6:2863–2874.
18. Foxall, P. A. and Mobley, H. L. T. personal communication.
19. Bukanov, N. Westblum, T.U. and Berg, D. E. unpublished data.
20. Krishnan, B. R., Kersulyte, D., Brikun, I., Berg, C. M., and Berg, D. E. 1991. Nuc. Acids Res. 19: 6177–6182.
21. Langenberg, W., Rauws, E. A., Oudbier, J. H. and Tytgat, G. N. 1990. J. Inf. Dis. 161:507–511.
22. Atwood, K. C., Schneider, L. K. and Ryan, F. J. (1951). Cold Spring Harb. Symp. Quant. Biol. 16:345–355.