Absence in *Helicobacter pylori* of an uptake sequence for enhancing uptake of homospecific DNA during transformation

Nigel J. Saunders,¹ John F. Peden² and E. Richard Moxon¹

Author for correspondence: Nigel J. Saunders. Tel: +44 1865 222347. Fax: +44 1865 222626.
e-mail: saunders@molbiol.ox.ac.uk

¹ Molecular Infectious Diseases Group, Institute of Molecular Medicine, University of Oxford, Headington, Oxford OX3 9DS, UK

² Oxford University Bioinformatics Centre, Sir William Dunn School of Pathology, South Parks, Oxford OX1 3RE, UK

Uptake sequences are abundant sequence motifs, often located downstream of ORFs, that are used to facilitate the within-species horizontal transfer of DNA. A frequent word analysis of the complete genome sequence of *Helicobacter pylori* strain 26685 was performed to search for and determine the identity of an uptake sequence in this species. The results demonstrated that *Hel. pylori* does not possess an uptake sequence. This is the first naturally transformable Gram-negative species shown to lack such a transformation-targeting system.

Keywords: uptake sequence, *Helicobacter pylori*, transformation

INTRODUCTION

Helicobacter pylori is a Gram-negative bacterial species which demonstrates a wide spectrum of genetic diversity (Marshall *et al.*, 1998): its genome size ranges from 1·0 to 1·8 Mb (Takami *et al.*, 1993), extensive intra-strain variability is revealed by RFLP (Akopyanz *et al.*, 1992) and there is little conservation of gene order in unrelated strains (Jiang *et al.*, 1996). The index of association (I_A) is a measure of linkage equilibrium between alleles and this provides a means to measure the genetic relatedness of bacterial strains. The extent to which bacterial populations can then be considered to be 'clonal' (high linkage disequilibrium/low frequency of horizontal exchange) or 'panmictic' (linkage equilibrium/high frequency of horizontal exchange) populations can then be determined (Smith *et al.*, 1993). The I_A , determined following analysis by multilocus enzyme electrophoresis, does not differ significantly from zero. This indicates that the clonal population structure of *Hel. pylori* has been completely disrupted due to frequent recombination between strains (Go *et al.*, 1996) and that the population can therefore be considered 'panmictic'. Furthermore, several genes investigated in *Hel. pylori* demonstrate a large degree of sequence divergence (Marshall *et al.*, 1998), as well as mosaicism (Atherton *et al.*, 1995). Finally, there is a clearly documented example of strains that have undergone horizontal exchange during co-infection of a human host (Kersulyte *et al.*, 1999). Taken together, these findings are consistent with

Hel. pylori being a population of bacteria that exchange DNA horizontally within the species at a high frequency.

Most strains of *Hel. pylori* are naturally transformable (Nedenskov-Sorensen *et al.*, 1990; Wang *et al.*, 1993; Tsuda *et al.*, 1993). Naturally transformable bacteria are able to take up DNA readily from their environment and have the capacity to use it as a means of horizontal exchange of genetic information. Other naturally transformable bacteria have developed strategies to increase the likelihood that the substrate for recombination is derived from related bacteria (Solomon & Grossman 1996; Saunders *et al.*, 1999). In Gram-positive bacteria this is typically achieved through signals of growth phase and cell density that foster transformation only when the majority of the potential donor bacteria are the same species. In Gram-negative bacteria, as exemplified by *Haemophilus* and *Neisseria* spp., it is achieved by using uptake sequences (USs) which are involved in enhancing the binding and uptake of homospecific DNA.

The US of *Haemophilus* spp. is a sequence containing a 9 bp core motif, AAGTGCGGT, of which there are 1465 copies in the *Haemophilus influenzae* strain Rd genome (Smith *et al.*, 1995). This US influences the process of DNA uptake into the recipient cell (Sisco & Smith 1979; Danner *et al.*, 1982). Similarly, *Neisseria* spp. possess a 10 bp US (GCCGTCTGAA) that performs a similar role (Goodman & Scocca, 1988; Elkins *et al.*, 1991). In both cases, USs often occur in inverted pairs and are typically located at the 3' end of ORFs where they are believed to act as transcriptional terminators – a function which may account for their frequency and

Abbreviation: US, uptake sequence.

distribution in the chromosome. By using USs, related species provide a pool of genes that are available to the other members of the species.

Searching the complete *Hae. influenzae* genome sequence for frequent words readily identifies the US (Karlin *et al.*, 1996) and this approach also identifies the US in those *Neisseria meningitidis* and *Neisseria gonorrhoeae* genome sequences that are available (N. J. Saunders, unpublished). We have used the approach of frequent word searching to demonstrate the absence of an US in the *Hel. pylori* strain 26685 genome sequence.

METHODS

The complete genome sequences of *Hel. pylori* (Tomb *et al.*, 1997) and *Hae. influenzae* (Fleischmann *et al.*, 1995) were downloaded by anonymous FTP from The Institute for Genomic Research FTP site (<ftp://ftp.tigr.org/pub/data/>). The genomes were initially stored and interrogated in the ACNUC database retrieval system (<http://pbil.univ-alyon1.fr/databases/acnuc.html>) (Gouy *et al.*, 1985). Each entire genome was extracted and processed by SCAN (J. F. Peden, unpublished), a C program which counts the frequency of all oligonucleotide words up to n -tuple size 12; SCAN also calculates the expected frequency of each word using a high-order Markov chain. A Markov chain is a stochastic process that can use the probability of the shorter component words (e.g. TCCT n -tuple=4 has two-component words with n -tuple=3, TCC and CCT) to predict the probability of a word with a greater n -tuple (Cox & Miller, 1965).

Where investigated, the sequence context of frequently occurring words was examined using ACEDB, a graphical user interface, to a whole genome analysis system (Saunders *et al.*, 1998).

RESULTS AND DISCUSSION

The observed and expected frequencies of all words up to n -tuple=12 were calculated from the genome sequences. The expected frequencies of each oligonucleotide were estimated by a high-order Markov chain. The 20 most frequent n -tuple=12 words in *Hel. pylori* and *Hae. influenzae* are shown in Table 1. The US of *Hae. influenzae* is underlined. There are no words in the *Hel. pylori* genome that occur at a frequency that would represent an US as seen in either *Haemophilus* or *Neisseria*. On the basis of the word list presented in the table a US of comparable length to those seen in *Haemophilus* or *Neisseria* occurs at no more than 10% of the comparable frequency in *Hel. pylori*.

The 20 most frequent words from *Hae. influenzae* include 2688 instances of words that include the *Haemophilus* US. *Haemophilus* word numbers 16, 19 and 20 include the tetrameric repeat TTGG which occurs in five locations in the genome, predominantly in association with iron-binding proteins (Hood *et al.*, 1996).

The most frequent words in *Hel. pylori* are composed of poly(A) and poly(T) repeats. Many of these have the potential to form hairpin loops that could form a

structure with similarity to that formed by inverted pairs of USs in *Hae. influenzae* and *N. meningitidis*. In order to determine the distribution of these sequences and whether they are distributed similarly to other USs, the context of each 12-mer word (TTTTTTAAAAAA) was determined. They occur both intra and inter-genically (15 intergenically, 24 within probable coding sequence within ORFs, and in one case the central TAA is the stop codon of an ORF) and are therefore not located in a way similar to that of known USs.

Finally, a measure of functional selection for word 12 on the basis of Markov chain predictions (Table 2) was made. The analysis revealed that this word is present less frequently than the occurrence of its components would predict and that the largest discrepancy in the occurrence of its components is at the transition between 3- and 4-mers. This is indicated by the abundance of the heptamer sequences over that predicted on the basis of the occurrence of di- and trinucleotides (as indicated by the Markov chain n -3 prediction in Table 2). In contrast, predictions based on longer words are much closer to observed results. Looking at the table as a whole, there would appear to be a bias towards the presence of two adenines after the run of thymidines. It is therefore clear that the abundance of the longer words is a consequence of the abundance of its component tetramers, rather than selection of the larger words. The abundance of the similar words is a product of the same phenomenon.

The *Hel. pylori* 'word 2' (AAAAAAAAAAAA) was also assessed in the same fashion (results shown in Table 3). This set of results is unusual and suggests the action of two quite different processes that influence the presence of homopolymeric tracts of A or T. Despite their abundance (they are the most frequent 6-mer word present in the genome) there are many fewer homopolymeric tracts of A or T of up to 8 bp in length present than would be expected on the basis of the abundance of their component parts. The abundance of these repeats is a product of the prevalence of the di- and trinucleotides AA/TT and AAA/TTT, whilst the longer repeats up to 8 bp are selected against. However, in marked contrast, repeats of 9–11 bp in length are much more abundant than predicted by Markov chain analysis. This suggests that there is a specific mutational tendency to generate and maintain repeats of this length – perhaps due to polymerase slippage, in repeats of 8 or more bp in length. This observation is consistent with, and extends, our previous observation of a comparative excess of homopurine:homopyrimidine repeats in the *Hel. pylori* genome (Saunders *et al.*, 1998).

The absence of an identifiable US in *Hel. pylori* is novel for a Gram-negative naturally transformable bacterium. The *in vivo* function of natural transformation is unknown but it is likely to be involved in horizontal transfer of partial, and perhaps also complete, genes between organisms where they become the substrate for homologous recombination. Through this mechanism they can be the substrates for the generation of allelic diversity and perhaps also the repair of degenerate sequences. In this context, *Hel. pylori* has a panmictic

Table 1. Most abundant words in *Hel. pylori* (strain 26685) and *Hae. influenzae* (strain Rd) on both strands of the complete genome sequences

Underlined sections indicate the *Hae. influenzae* US. Bold indicates the component of an oligonucleotide repeat.

Word no.	<i>Hel. pylori</i>		<i>Hae. influenzae</i>	
	Word	Frequency	Word	Frequency
1	TTTTTAAAAAT	60	AAAGTGC GGT CA	251
2	AAAAAAAAAAAA	59	<u>ACCGCACTTTT</u>	242
3	CTTTTTTAAAA	49	TA <u>ACCGCACTT</u>	210
4	AAAAGCGTTTT	48	<u>AAGTGC</u> GGTCAA	202
5	TTTTTTTAAAA	47	TT <u>ACCGCACTT</u>	182
6	TTTTTAGAAAA	47	A <u>ACCGCACTT</u>	180
7	CTTTTTCTAAAA	44	<u>AAGTGC</u> GGTTAA	170
8	TTCTTTTCTAA	43	<u>ACCGCACTTT</u> TA	162
9	GATTTTAAAAA	43	AAAAGTGC GGT C	149
10	TTTTAGAAAAA	42	TTT <u>ACCGCACT</u>	141
11	TCTTTTTCTAAA	41	AAAAGTGC GGT A	141
12	TTTTTTTAAAA	40	TC <u>ACCGCACTT</u>	117
13	AAAAAGCGTTTT	40	<u>AGTGC</u> GGTTAAA	116
14	AACGCTTTAAAA	40	<u>CACCGCACTT</u>	115
15	TTTTAAAAAGA	39	CAAAAGTGC GGT	107
16	GCTTTTTTAAAA	39	TTGG TTGGTTGG	103
17	TTTAAAAGCGTT	39	<u>AGTGC</u> GGTCAAT	102
18	TTTAGCGTTTTT	38	<u>AGTGC</u> GGTCAAA	101
19	TTTTAAAAAATT	38	TGG TTGG TTGGT	101
20	AGCGTTTTTTTAA	38	GG TTGG TTGGTT	101

population structure indicating a rate of recombination that exceeds the rate of mutation-driven diversification (Go *et al.*, 1996). This population structure is similar to that seen in other naturally transformable Gram-negative species such as *N. gonorrhoeae* and *Hae. influenzae* (Smith *et al.*, 1993). The presence of mosaic genes (Atherton *et al.*, 1995) is also a feature common to naturally transformable species (e.g. Zhou & Spratt, 1992). Finally, there is a clearly documented example of horizontal transfer involving six exchanges between two strains during a mixed infection in the human stomach (Kersulyte *et al.*, 1999). From these observations it can be concluded that horizontal transfer between *Hel. pylori* strains is important in their diversification.

The function of the targeting systems such as USs in Gram-negative bacteria and indicators of cell density in Gram-positive bacteria is to enhance the probability that the transforming DNA in donor and recipient is similar. This relative exclusivity increases the likelihood that products of recombination are going to be functional. There are several possible explanations for why *Hel. pylori* lacks an US. As described previously, a typical US of *Neisseria* and *Haemophilus* is located 3' of expressed ORFs where it is thought to act as a transcriptional terminator. At some point in the evolution of these species this sequence has been co-opted to a second function in transformation. This second function may have affected the sequence, structure and

location of the US but it is hard to see how it could have developed without the pre-existing wide distribution of the sequence from which it arose. One possibility is that *Hel. pylori* has never possessed a suitable transcriptional terminator that could evolve this second function.

Another interesting possibility is that *Hel. pylori* has never been under selective pressure to develop a targeting system akin to those present in other species. Other naturally transformable bacterial species live in niches colonized by a mixed bacterial flora and would be exposed to DNA derived from unrelated bacteria much of the time. *Hel. pylori* resides for a large proportion of its life cycle and generations in the gastric mucosa. Colonization by multiple strains of *Hel. pylori* does occur (Taylor *et al.*, 1995; van der Ende *et al.*, 1996; Jorgensen *et al.*, 1996; Berg *et al.*, 1997), so opportunity for horizontal exchange between strains exists in nature. However, the stomach is not colonized by many other bacterial species. It may be that this ecological separation makes an additional mechanism unnecessary. The potential for such separation to contribute to the maintenance of distinct populations is seen in the pathogenic *Neisseria* spp., which share a common US and are able to exchange DNA, but exist as clearly separate species (Vazquez *et al.*, 1993).

Another factor acting to maintain the integrity of the species is the availability of substrates for homologous recombination. However, there are examples of the

Table 2. Markov chain results for *Hel. pylori* 'word 12' from Table 1

Markov chain $n-1$ is based upon the component word two bases shorter than the word predicted; Markov chain $n-2$ is based upon the component word three bases shorter than the word predicted; Markov chain $n-3$ is based upon the component word four bases shorter than the word predicted.

Word	Actual frequency	Predicted frequency		
		Markov chain $n-3$	Markov chain $n-2$	Markov chain $n-1$
TTTTTTAAAAAA	40	58.0	65.7	62.7
TTTTTTAAAAA	99	83.4	83.7	92.2
TTTTTAAAAAA	85	78.7	81.4	89.8
TTTTTTAAAAA	220	208.8	205.9	211.5
TTTTTAAAAAA	218	198.6	202.4	209.7
TTTTTTAAAAA	206	214.1	198.8	193.9
TTTTTAAAAA	1262	1312.7	1325.4	1314.9
TTTAAAAA	1057	1044.6	1040.1	1069.6
TTTTTAAAA	985	1036.8	1011.0	1006.1
TTTTTTTAA	860	768.4	687.9	818.7
AATTTTTTT	844	754.6	688.0	885.4
TTTAAAAA	2607	2030.4	2679.0	2616.6
TTTTTAAA	2535	2039.6	2642.4	2545.8
TTAAAAA	2134	1719.2	1935.6	2067.6
TTTTTTAA	2047	1770.5	1957.9	2048.1
TTTTTTTA	1883	1931.1	1825.9	1581.3
TAAAAAA	1740	1886.4	1808.4	1572.8

Table 3. Markov chain results for *Hel. pylori* 'word 2'

For each homopolymeric tract of length L (n -tuple = L) the highest order Markov chain that can be calculated is n ($n = L - 1$); we present here the $n-1$, $n-2$ and $n-3$ results.

Length of A/T homopolymeric tract	Actual frequency	Predicted frequency		
		Markov chain $n-3$	Markov chain $n-2$	Markov chain $n-1$
12	59	7.1	56	102
11	78	5.9	14.5	58
10	99	45	22	50
9	171	401	176	127
8	589	1923	1201	693
7	2729	6513	4917	3592
6	10742	17186	15159	12568

acquisition of whole genes for which the presence of a similar sequence in the recipient cannot be a pre-requisite (Vazquez *et al.*, 1995). In the absence of a targeting system, *Hel. pylori* might be expected to have an increased likelihood of incorporating DNA from unrelated species when it is encountered. In this context, it is interesting to note that the sigma-factor-encoding *rpoD* gene of *Hel. pylori* has strongest similarity to those from Gram-positive rather than Gram-negative bacteria (Solnick *et al.*, 1997). Further, the cytochrome-*c* bio-

genesis system in *Hel. pylori* is a Type II system, which is typically found in Gram-positive bacteria, cyanobacteria and chloroplasts rather than Gram-negative bacteria, which normally have a Type I system (Goldman & Kranz, 1998). Finally, it is striking that a large number of genes share an unusually high amount of sequence similarity with genes from other taxonomic groups – although it is necessary to be circumspect about this interpretation given the bias in current databases used for the comparisons.

Horizontal transfer is undoubtedly important in the population biology and evolution of *Hel. pylori*. Although additional factors may exist, including a deoxyribonuclease resistance mechanism (Kuipers *et al.*, 1998), natural transformation is probably an important means of genetic exchange in this species. Natural transformation can be observed *in vitro*, the nature of the genetic diversity observed in *Hel. pylori* is similar to that present in other naturally transformable species and co-colonization by multiple strains has been demonstrated. The absence of an US in such an organism is novel and we suggest reflects the environment in which *Hel. pylori* resides. Furthermore, it suggests that this species may have a greater propensity to be the recipient for interspecific horizontal transfer of genes than is the case for other naturally transformable bacteria.

ACKNOWLEDGEMENTS

N. J. S. is supported by a Wellcome Trust research fellowship in medical microbiology.

REFERENCES

- Akopyanz, N., Bukanov, N., Westblom, T. U. & Berg, D. E. (1992). PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*. *Nucleic Acids Res* **20**, 6221–6225.
- Atherton, J. C., Cao, P., Peek, R. M., Tummuru, M. K. R., Blaser, M. J. & Cover, T. L. (1995). Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. *J Biol Chem* **270**, 17771–17777.
- Berg, D. E., Gilman, R. H., Lelwala-Guruge, J. & 9 other authors (1997). *Helicobacter pylori* populations in Peruvian patients. *Clin Infect Dis* **25**, 996–1002.
- Cox, D. R. & Miller, H. D. (1965). *The Theory of Stochastic Processes*. New York: Chapman & Hall.
- Danner, D. B., Smith, H. O. & Narang, S. A. (1982). Construction of DNA recognition sites active in *Haemophilus* transformation. *Proc Natl Acad Sci USA* **79**, 2393–2397.
- Elkins, C., Thomas, C. E., Seifert, H. S. & Sparling, P. F. (1991). Species-specific uptake of DNA by gonococci is mediated by a 10-base-pair sequence. *J Bacteriol* **173**, 3911–3913.
- van der Ende, A., Rauws, E. A., Feller, M., Mulder, C. J., Tytgat, G. N. & Dankert, J. (1996). Heterologous *Helicobacter pylori* isolates from members of a family with a history of peptic ulcer disease. *Gastroenterology* **111**, 638–647.
- Fleischmann, R. D., Adams, M. D., White, O. & 37 other authors (1995). Whole-genome random sequencing and assembly of *Haemophilus* Rd. *Science* **269**, 496–512.
- Go, M. F., Kapur, V., Graham, D. Y. & Musser, J. M. (1996). Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinatorial population structure. *J Bacteriol* **178**, 3934–3938.
- Goldman, B. & Kranz, R. (1998). Evolution and horizontal transfer of an entire biosynthetic pathway for cytochrome *c* biogenesis: *Helicobacter*, *Deinococcus*, *Archaea* and more. *Mol Microbiol* **27**, 871–873.
- Goodman, S. D. & Scoocca, J. J. (1988). Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *Proc Natl Acad Sci USA* **85**, 6982–6986.
- Gouy, M., Gautier, C., Attimonelli, M., Lanave, C. & Dipaola, G. (1985). ACNUC – a portable retrieval-system for nucleic-acid sequence databases – logical and physical designs and usage. *Comput Appl Biosci* **1**, 167–172.
- Hood, D. W., Deadman, M. E., Jennings, M. P., Bisceric, M., Fleischmann, R. D., Venter, J. C. & Moxon, E. R. (1996). DNA repeats identify novel virulence genes in *Haemophilus influenzae*. *Proc Natl Acad Sci USA* **93**, 11121–11125.
- Jiang, Q., Hiratsuka, K. & Taylor, D. E. (1996). Variability of gene order in different *Helicobacter pylori* strains contributes to genome diversity. *Mol Microbiol* **20**, 833–842.
- Jorgensen, M., Daskalopoulos, G., Warburton, V., Mitchell, H. M. & Hazell, S. L. (1996). Multiple strain colonisation and metronidazole resistance in *Helicobacter pylori*-infected patients: identification from sequential and multiple biopsy specimens. *J Infect Dis* **174**, 631–635.
- Karlin, S., Mrazek, J. & Campbell, A. M. (1996). Frequent oligonucleotides and peptides of the *Haemophilus influenzae* genome. *Nucleic Acids Res* **24**, 4263–4272.
- Kersulyte, D., Chalkauskas, H. & Berg, D. E. (1999). Emergence of recombinant strains of *Helicobacter pylori* during human infection. *Mol Microbiol* **31**, 31–43.
- Kuipers, E. J., Israel, D. A., Kusters, J. G. & Blaser, B. J. (1998). Evidence for a conjugation-like mechanism of DNA transfer in *Helicobacter pylori*. *J Bacteriol* **180**, 2901–2905.
- Marshall, D. G., Dundon, W. G., Beesley, S. M. & Smyth, C. J. (1998). *Helicobacter pylori* – a conundrum of genetic diversity. *Microbiology* **144**, 2925–2939.
- Nedenskov-Sorensen, P., Buckholm, G. & Bovre, K. (1990). Natural competence for genetic transformation in *Campylobacter pylori*. *J Infect Dis* **161**, 365–366.
- Saunders, N. J., Peden, J. F., Hood, D. W. & Moxon, E. R. (1998). Simple sequence repeats in the *Helicobacter pylori* genome. *Mol Microbiol* **27**, 1091–1098.
- Saunders, N. J., Hood, D. W. & Moxon, E. R. (1999). Bacterial evolution: bacteria play pass the gene. *Curr Biol* **9**, 180–183.
- Sisco, K. L. & Smith, H. O. (1979). Sequence-specific DNA uptake in *Haemophilus* transformation. *Proc Natl Acad Sci USA* **76**, 972–976.
- Smith, J. M., Smith, N. H., O'Rourke, M. & Spratt, B. G. (1993). How clonal are bacteria? *Proc Natl Acad Sci USA* **90**, 4384–4388.
- Smith, H. O., Tomb, J.-F., Dougherty, B. A., Fleischmann, R. D. & Venter, J. C. (1995). Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. *Science* **269**, 538–540.
- Solnick, J. V., Hansen, L. M. & Syvanen, M. (1997). The major sigma factor (RpoD) from *Helicobacter pylori* and other Gram-negative bacteria shows an enhanced rate of divergence. *J Bacteriol* **179**, 6196–6200.
- Solomon, J. M. & Grossman, A. D. (1996). Who's competent and when: regulation of natural genetic competence in bacteria. *Trends Genet* **12**, 150–155.
- Takami, S., Hayashi, T., Tonokatsu, Y., Shimoyama, T. & Tamura, T. (1993). Chromosomal heterogeneity of *Helicobacter pylori* isolates by pulsed-field gel electrophoresis. *Int J Med Microbiol Virol Parasitol Infect Dis* **280**, 120–127.
- Taylor, N. S., Fox, J. G., Akopyants, N. S. & 8 other authors (1995). Long-term colonisation with single and multiple strains of *Helicobacter pylori* assessed by DNA fingerprinting. *J Clin Microbiol* **33**, 918–923.
- Tomb, J.-F., White, O., Kerlavage, A. R. & 39 other authors (1997).

The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539–547.

Tsuda, M., Karita, M. & Nakazawa, T. (1993). Genetic transformation in *Helicobacter pylori*. *Microbiol Immunol* **37**, 85–89.

Vazquez, J. A., Fuente, L., Berron, S., O'Rourke, M., Smith, N. H., Zhou, J. & Spratt, B. G. (1993). Ecological separation and genetic isolation of *Neisseria gonorrhoeae* and *Neisseria meningitidis*. *Curr Biol* **3**, 567–572.

Vazquez, J. A., Berron, S., O'Rourke, M., Carpenter, G., Feil, E., Smith, N. H. & Spratt, B. G. (1995). Interspecies recombination in nature: a meningococcus that has acquired a gonococcal PIB porin. *Mol Microbiol* **15**, 1001–1007.

Wang, Y., Roos, K. P. & Taylor, D. E. (1993). Transformation of *Helicobacter pylori* by chromosomal metronidazole resistance and by a plasmid with a selectable chloramphenicol resistance marker. *J Gen Microbiol* **139**, 2485–2493.

Zhou, J. & Spratt, B. G. (1992). Sequence diversity within the *argF*, *fbp* and *recA* genes of natural isolates of *Neisseria meningitidis*: interspecies recombination within the *argF* gene. *Mol Microbiol* **6**, 2135–2146.

Received 4 May 1999; revised 17 August 1999; accepted 10 September 1999.