Effect of FliK mutation on the transcriptional activity of the σ^{54} sigma factor RpoN in *Helicobacter pylori*

Francois P. Douillard,¹ Kieran A. Ryan,¹ Jason Hinds²
and Paul W. O'Toole¹

Correspondence
Paul W. O'Toole
pwotoole@ucc.ie

¹Department of Microbiology and Alimentary Pharmabiotic Centre, University College Cork, Western Road, Cork, Ireland

²Bacterial Microarray Group, Division of Cellular and Molecular Medicine, St George's University of London, Cranmer Terrace, London SW17 0RE, UK

Received 21 November 2008
Revised 5 March 2009
Accepted 9 March 2009

Helicobacter pylori is a motile Gram-negative bacterium that colonizes and persists in the human gastric mucosa. The flagellum gene regulatory circuitry of *H. pylori* is unique in many aspects compared with the *Salmonella*/*Escherichia coli* paradigms, and some regulatory checkpoints remain unclear. FliK controls the hook length during flagellar assembly. Microarray analysis of a *fliK*-null mutant revealed increased transcription of genes under the control of the σ^{54} sigma factor RpoN. This sigma factor has been shown to be responsible for transcription of the class II flagellar genes, including *flgE* and *flaB*. No genes higher in the flagellar hierarchy had altered expression, suggesting specific and localized FliK-dependent feedback on the RpoN regulon. FliK thus appears to be involved in three processes: hook-length control, export substrate specificity and control of RpoN transcriptional activity.

INTRODUCTION

Helicobacter pylori infection is responsible for gastrointestinal disorders such as peptic and duodenal ulcers (Veldhuyzen van Zanten & Sherman, 1994), and is a predisposing factor for gastric adenocarcinoma (EUROGAST, 1993) and B-cell MALT lymphoma (Parsonnet *et al.*, 1994). Epidemiological studies show large disparities in *H. pylori* infection rates between geographical regions (Kikuchi & Dore, 2005). Developing countries are the most severely infected by *H. pylori*, due to poor living conditions and limited access to therapies (Frenck & Clemens, 2003). The continuing emergence of new antibiotic-resistant strains underlines the urgent need to expand our understanding of the biology of *H. pylori*, in order to facilitate the development of new treatments.

Motility is a key feature of *H. pylori* and is required for colonization and persistence (Eaton *et al.*, 1991, 1992, 1996). In motile bacteria, flagella contribute to motility,

adhesion and inflammatory response by the host cells (Eaton *et al.*, 1992; Galkin *et al.*, 2008). Three RNA polymerase sigma factors, σ^{80} , σ^{54} and σ^{28} , control the transcription of *H. pylori* flagellar genes (Niehus *et al.*, 2004; Scarlato *et al.*, 2001). σ^{80} modulates the transcription of the early flagellar genes (class I). Middle flagellar structural genes (class II) are under the control of RpoN (σ^{54}). RpoN transcriptional activity is tightly regulated by the FlgR/FlgS activation system and HP0958, an RpoN chaperone (Brahmachary *et al.*, 2004; Pereira & Hoover, 2005; Ryan *et al.*, 2005a). The class II regulon includes HP0906/*fliK*, which encodes the hook-length-control protein. Transcription of class III genes (late flagellar genes) is controlled by FliA (σ^{28}) and its anti-sigma factor FlgM (Colland *et al.*, 2001; Josenhans *et al.*, 2002).

The initial annotations of *H. pylori* genome sequences did not identify all flagellar genes expected by comparison with the *Salmonella*/*Escherichia coli* paradigm (Alm *et al.*, 1999; Tomb *et al.*, 1997). Thus, the anti-sigma 28 factor FlgM was only identified subsequently, by rigorously searching for low-level conservation (Colland *et al.*, 2001; Josenhans *et al.*, 2002). Similarly, the hook-length protein FliK has only been recently identified based on bioinformatic analyses (Ryan *et al.*, 2005b). The HP0906 gene product has been shown to be FliK in *H. pylori* (Ryan *et al.*, 2005b). FliK in *Salmonella* controls hook length (Ferris & Minamino, 2006). Ablation of the *H. pylori* *fliK* gene impairs motility, and similar to a *fliK*-null mutant in *Salmonella*, the cells harbour polyhook structures (Ryan

Abbreviations: CGH, comparative genome hybridization; qRT-PCR, quantitative real-time PCR.

The microarray design is available in BjuG@Sbase (accession number A-BUGS-18; <http://bugs.sgul.ac.uk/A-BUGS-18>) and also ArrayExpress (accession number A-BUGS-18). Fully annotated microarray data have been deposited in BjuG@Sbase (accession number E-BUGS-78; <http://bugs.sgul.ac.uk/E-BUGS-78>) and also ArrayExpress (accession number E-BUGS-78).

A supplementary table, listing oligonucleotide primers used in this study, is available with the online version of this paper.

et al., 2005b). A dramatic reduction in flagellin production and overproduction of the FlgE hook protein are also observed, and *flgE* and *flaB* transcription levels are significantly increased in the *fliK* mutant (Ryan *et al.*, 2005b). These genes have been shown to be under the control of RpoN, the σ^{54} sigma factor of *H. pylori* (Beier & Frank, 2000; Brahmachary *et al.*, 2004; Niehus *et al.*, 2004; Scarlato *et al.*, 2001). Three genes belonging to the intermediate gene class do not show altered transcription levels according to targeted quantitative real-time PCR (qRT-PCR) analysis (Ryan *et al.*, 2005b). We thus suggested that the FliK protein is required to turn off the RpoN regulon during flagellar assembly. However, the transcriptional profile of the whole flagellar regulon was not investigated, leaving the extent of FliK feedback on flagellar gene expression unclear.

In the present study, we used a pan-*H. pylori* array, based on the genomes of strains NCTC26695 and J99. Array comparative genomic hybridization (CGH) was performed to identify specific chromosomal regions that were missing in CCUG17874, the motile strain used in our laboratory for flagellum genetics. Global transcript analysis of a CCUG17874 mutant lacking the *fliK* gene was performed to further investigate the role of FliK in flagellar biogenesis in *H. pylori*. We integrated these microarray data into the currently known flagellar regulatory model.

METHODS

Bacterial strains, media and growth conditions. *H. pylori* type strain CCUG 17874 (Culture Collection, University of Gothenburg, Sweden) was cultured as previously described (Ryan *et al.*, 2005b). *H. pylori* mutants defective in the *fliK* gene (Ryan *et al.*, 2005a, b) and *flgE* gene (O'Toole *et al.*, 1994) have been previously described. The mutants used in this study were cultured on Columbia agar base (CBA) plates containing the appropriate antibiotic: 10 µg chloramphenicol ml⁻¹ (Sigma) or 15 µg kanamycin ml⁻¹ (Sigma).

Extraction of genomic DNA from *H. pylori*. Genomic DNA from 2 day-old *H. pylori* plate cultures was isolated using the DNeasy tissue kit (Qiagen). The genomic DNA was then quantified using a Nanodrop ND-1000 spectrophotometer.

RNA isolation from *H. pylori*. Total RNA was isolated from 20 h *H. pylori* liquid cultures using the Qiagen RNeasy Mini kit. *H. pylori* cells were harvested and centrifuged for 15 s at $\geq 10\,000$ g. Cell pellets were then resuspended in 750 µl Bacteria RNA Protect reagent (Qiagen). The remainder of the protocol was performed as per the manufacturer's instructions (RNeasy Mini kit; Qiagen). RNA quality was assessed using a Bioanalyser 2100 Instrument (Agilent Technologies) and quantified by NanoDrop ND-1000 spectrophotometer. Prior to further array experiments, total RNA samples were DNase-treated using a Turbo DNA-free kit (Ambion).

***H. pylori* microarray design and construction.** Design and construction of the *H. pylori* microarray (BµG@S HPv1.0.0; Bacterial Microarray Group at St George's, University of London) was completed using the approaches described by Hinds *et al.* (2002a, b). In brief, PCR products were designed to represent all 1576 ORFs in *H. pylori* NCTC26695 and all 1495 ORFs in *H. pylori* J99 (Alm *et al.*, 1999; Tomb *et al.*, 1997). The microarrays were constructed by

robotic spotting of PCR products in duplicate on UltraGaps amino-silane-coated glass slides (Corning) using a MicroGrid II automated microarrayer (BioRobotics) and post-print-processed according to the slide manufacturer's instructions.

CGH. Because of its history of genetic analysis and its motility, *H. pylori* strain CCUG17874 (equivalent to the type strain NCTC11637) was used to study global transcription patterns of different flagellar mutants. The genome of this strain has not been sequenced. The macrodiversity of CCUG17874 was therefore investigated using CGH. Nucleic acid labelling was undertaken using a modified protocol described by Hinds *et al.* (2002a). Briefly, 5 µg wild-type CCUG17874 genomic DNA was labelled with dCTP Cy3. A 41.5 µl mixture containing 5 µg genomic DNA and 3 µg random primers (Promega) was heated at 95 °C for 5 min, snap-cooled and quickly centrifuged at 10 000 g. A 5 µl volume of 10 × REact2 buffer (Invitrogen), 1 µl dNTP (5 mM dDTP, 2 mM dCTP), 1 µl DNA polymerase I large Klenow fragment (Invitrogen) and 1.5 µl Cy3 dCTP Fluorolink (GE Healthcare) was then added. Similarly, wild-type NCTC26695 genomic DNA was labelled with dCTP Cy5. The mixture was incubated at 37 °C for 90 min in the dark. Subsequent steps, including purification of labelled cDNA, hybridization, washing and scanning, were performed as described below. Array hybridizations for the wild-type and for the mutant were performed in duplicate. Following quantile normalization (Bolstad *et al.*, 2003), a dynamic cut-off-based analysis was performed as described in Kim *et al.* (2002). The output files give three lists: divergent genes, uncertain genes and present genes.

Confirmatory analysis of selected CGH data by PCR. PCRs were performed under standard conditions to confirm selected results obtained by CGH. Seven genes were selected from the CGH data. Primer pair sequences were designed using the Primer3 online software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and are listed in Supplementary Table S1.

Type II microarray analysis. To compare the transcriptional profiles of the wild-type and HP0906 mutant strains, the *H. pylori* whole genome microarray was used in a common reference or type II experimental design whereby Cy5-labelled cDNA from each strain was co-hybridized to an array with a Cy3-labelled genomic DNA reference. Nucleic acid labelling and microarray hybridizations were undertaken according to BµG@S standard protocols (Hinds *et al.*, 2002a). In brief, for the common reference, 5 µg wild-type CCUG17874 genomic DNA was labelled with dCTP Cy3 using random primers (Promega) and DNA polymerase I large Klenow fragment (Invitrogen). Cy5-labelled cDNA was generated from 4 µg total RNA during first-stand synthesis using random primers (Promega) and Superscript II reverse transcriptase (Invitrogen). The Cy3- and Cy5-labelled nucleic acid mixtures were then co-purified using the MinElute PCR Purification kit (Qiagen), mixed in a hybridization solution of 4 × SSC and 0.3 % SDS, and hybridized under a LifterSlip (Erie Scientific) for 18 h at 65 °C. Microarray slides were washed once in 1 × SSC, 0.06 % SDS at 65 °C for 2 min and twice in 0.06 × SSC for 2 min, dried by centrifugation, and scanned using a dual-laser Affymetrix 428 scanner. The images and data were analysed using the GeneDirector software package, which includes ImaGene and GeneSight v2.0 (Biodiscovery). Genes designated as missing or uncertain in CCUG17874 were filtered out. Array hybridizations for the wild-type and for the mutant were performed in triplicate. Ratios for wild-type DNA versus wild-type RNA and wild-type DNA versus mutant RNA were exported to Excel (Microsoft). Following the within-slide normalization in GeneSight, the genes containing empty values were discarded. Quantile normalization was then performed to make the entire distribution of the values identical between each array slide (Bolstad *et al.*, 2003). The triplicates for the wild-type and for the mutant were averaged

and final ratios ($\log_2$) were calculated. One-way ANOVA was used to calculate statistical confidences. Genes with a P value less than 0.05 and a fold-change greater than 2.00 were designated differentially expressed in the mutant.

Quantitative analysis of transcription by real-time PCR. qRT-PCR was performed as a confirmatory test on four flagellar genes following global transcript analysis by microarray. Real-time PCR primers were designed using the Primer3 software package (Rozen & Skaletsky, 2000) and are listed in Supplementary Table S1. RNA (500 ng) was reverse-transcribed using Improm-II reverse transcriptase (Promega) and 500 ng random hexamers, as described in the manufacturer's manual. qRT-PCR was performed on flagellar genes using primers listed in Supplementary Table S1. The reaction mixture was prepared as described in the manufacturer's protocol. Briefly, the amplification by qRT-PCR was performed in a final volume of 12.5 μ l including 1 μ l cDNA, 50 nM of each primer, 6.25 μ l 2 $\times$ Master Mix (Biogen) and 1:60 000 Sybr Green I (Bio/Gene). qRT-PCRs were run and monitored using an ABI 7000 Thermo cycler and ABI Prism 7000 SDS software (both from Applied Biosystems). Reactions were performed in triplicate (technical replicates) from at least two independent RNA preparations (biological replicates). Relative fold-changes of expression were calculated as described by Pfaffl (2001). The *era* gene was used as a housekeeping gene (Sebert *et al.*, 2002). Each transcript abundance was therefore calculated relative to the *era* gene transcript abundance.

Reverse transcription and amplification of intergenic regions by PCR. RNA (500 ng) was reverse-transcribed using Improm-II reverse transcriptase (Promega) and 500 ng random hexamers, as described in the manufacturer's manual. Primer pairs were designed to amplify the intergenic region between *hp0114* and *hp0115* and an internal region of the *hp0115* gene (Supplementary Table S1). Next, the two regions were amplified by PCR using cDNA preparations.

Bioinformatics analysis. Predictions of stress-induced DNA duplex destabilization were performed using the SIDD online server developed by the C. Benham laboratory (Bi & Benham, 2004). Predictions of transcriptional terminators were performed using TransTermHP (Kingsford *et al.*, 2007).

RESULTS

CGH of *H. pylori* strain CCUG17874

We and others have performed genetic and microbiological analysis of *H. pylori* motility in strain CCUG17874, whose genome has not been sequenced. To investigate its genome content, we used an *H. pylori* DNA array representing all the ORFs from strains NCTC26695 and J99. To analyse the level of genetic conservation between CCUG17874 and the sequenced strains, and to improve the subsequent type II array experiment analyses, CGH was performed. Based on the trinary output analysis method described by Kim *et al.* (2002), genes were classified into three groups: absent (highly divergent), uncertain and absolutely present. The genes designated uncertain have values that are in the transition region, where they cannot be confidently assigned to one or the other group. The uncertain group consisted of 130 genes. Seven genes annotated as highly divergent, uncertain or absolutely present were analysed by PCR using genomic DNA from CCUG17874 (test strain) and NCTC26695 (reference strain) (Fig. 1).

Fig. 1. Confirmatory PCR for CGH analysis. Seven genes were amplified using the genomic DNA of CCUG17874 and NCTC26695. *omp27* and *flgI* were assigned as divergent. *groES*, *flaG* and *omp13* were classified as uncertain. *cheA* and *omp22* were classified as present.

The PCR data confirmed the CGH results, with one exception. The HP1192 gene was shown to be present in CCUG17874 using a PCR approach, whereas CGH data indicated its absence (data not shown). Sequence divergence is correlated with hybridization values (Kim *et al.*, 2002). It has been reported that the Cy5: Cy3 ratio decreases twofold if the corresponding gene has 85 % identity with the gene used for the array design (van Hijum *et al.*, 2008). Thus, some of the divergent genes in CCUG17874 may be present but their sequences may be poorly conserved compared with the sequenced reference strain. The confirmatory PCRs provided empirical data to allow rational adjustment of the cut-off value for spot intensity in the array data, and thus to narrow the break-point between genes classified as present and divergent (Fig. 1). The uncertain gene list was thus reduced from 130 to seven genes following the PCR screening. Out of the seven, five genes encode hypothetical proteins. The two remaining genes encode transketolase and geranyltransferase. We classified these genes as divergent for the data analysis of further type II array experiments. *flgI*, encoding the flagellar basal-body P-ring protein, is an essential structural component of the flagellar superstructure. CGH data indicated that *flgI* was absent from the genome of strain CCUG17874. However, the standard deviation of hybridization values for this gene was very high, suggesting that the CGH output value for this gene was not reliable. PCR investigation confirmed the presence of *flgI*, which is consistent with the motile phenotype of the CCUG17874 strain used in this study (Fig. 1).

Table 1 lists genes present in NCTC26695 and J99 but missing in CCUG17874. Significantly, 14.25 % of *H. pylori* NCTC26695 genes were absent in the CCUG17874 genome. Most of the missing genes encode hypothetical proteins and appear to be strain- and *H. pylori*-specific (Table 1). Some genes that encode restriction enzymes, outer-membrane proteins, insertion elements (possibly in

Table 1. Genes absent from the genome of *H. pylori* strain CCUG17874, relative to two sequenced genomes, as determined by CGH

Strain	Number of divergent genes	Cellular role category
NCTC26695	147	Unknown function
	14	DNA metabolism (restriction enzyme)
	7	Cell envelope
	5	Pseudogene
	4	Energy metabolism
	5	Mobile and extrachromosomal element function
	2	Chemotaxis and motility
	3	Cellular processes: pathogenesis
	16	Others
CCUG17874	24	Unknown function
	2	DNA metabolism (topoisomerase)
	1	Other (DNA transfer protein)

reduced copy number in the CCUG17874 genome), virulence-associated proteins and transposases (possibly in reduced copy number in the CCUG17874 genome) were also absent. Transposable elements are often present in multiple copies in the *H. pylori* genome (Logan & Berg, 1996; Tomb *et al.*, 1997) and may be associated with disruption of operons and large regions in the CCUG17874 genome.

Consideration of the location of genes identified as missing by the CGH analysis revealed that five large regions have been disrupted in the genome of CCUG 17874 relative to

the NCTC26695 genome (Fig. 2). These regions consist of genes encoding hypothetical proteins and/or mobile and extrachromosomal elements, such as transposases and recombinases. These genes designated as divergent in CCUG17874 were filtered out in further type II array experiments using that *H. pylori* strain.

Transcription analysis of the *fliK* mutant

A previous transcriptional analysis showed that *FliK* is under the control of the σ^{54} sigma factor (Niehus *et al.*,

Fig. 2. Overview of the divergence of the genome of *H. pylori* CCUG17874 relative to NCTC26695. Outer rings show strand location of genes, with genes missing in CCUG17874 (relative to NCTC26695) in red, and genes present in green. The middle ring shows average mol% G+C content. The inner ring is GC skew.

2004). To provide the first genome-wide analysis of FliK-related regulatory circuitry in *H. pylori*, we performed global transcript analysis of the HP0906 insertional mutant. Fifty-four genes were differentially transcribed in the HP0906 mutant, including 34 hypothetical or putative proteins. Differentially expressed genes with functional annotations are listed in Table 2. Seven genes encoding ribosomal proteins were significantly upregulated in the FliK mutant. The microarray data also showed upregulation of stress-related proteins, such as ferredoxin and thioredoxin, in the FliK mutant. HP0923, encoding an outer-membrane protein (Omp22) that is highly immunoreactive (Kim *et al.*, 2000), was downregulated. Three RpoN-dependent genes, *flaB*, *flgE* and HP1076, were upregulated (Table 2).

Because of its likely role in the flagellar regulon, we carefully examined expression fold-changes of the flagellar genes in the HP0906 mutant (Table 3). Only one class I gene (*fliP*/HP0684) was significantly downregulated. The genes for RpoN, the regulator FlhA, FlgR (HP0703) and FlgS (HP0244) were transcribed at wild-type levels in the *fliK* mutant. All class III genes were also transcribed at wild-type levels. The FliA sigma factor and its anti-sigma factor FlgM were not significantly differentially expressed. Only one gene in the intermediate class was significantly upregulated, *hp0367*. The response regulator *ompR* was upregulated 1.9-fold and this upregulation was confirmed by qRT-PCR, which indicated a fold upregulation of

2.25 ± 0.54 -fold. *fliK* mutation affected the transcription of almost all class II genes in the flagellar regulon (Table 3), with only three exceptions.

HP0114, part of the RpoN regulon and located downstream of *flaB*/HP0115, was not upregulated compared with other RpoN-dependent genes in the *fliK* mutant (Fig. 3, Table 3). HP0114 also has a different transcript profile compared with the other RpoN-dependent genes in an HP0958 mutant (Douillard *et al.*, 2008). We therefore performed bioinformatic analysis to identify DNA regulatory regions, such as promoters, terminations and intrinsic termination motifs (rho-independent transcription terminators) (Bi & Benham, 2004; Kingsford *et al.*, 2007). The algorithm developed by Bi & Benham (2004) identifies regions where the DNA duplex is destabilized, i.e. promoters and terminators. The results suggested that HP0114 and *flaB* are not co-transcribed (Fig. 3a), which is consistent with our data. In addition, a transcription terminator prediction algorithm (Kingsford *et al.*, 2007) identified a potential terminator GGGC GTTA GCCC located downstream of *flaB* (confidence score 88 out of 100), followed immediately by a U-rich region. Although reverse-transcription investigations showed a weak band corresponding to the intergenic region between HP0114 and *flaB*, it appears that these two genes are not completely co-transcribed. The weak PCR band obtained could actually correspond to low-level read-through from the *flaB* transcript.

Table 2. Selected genes differentially expressed in the HP0906 mutant

Only genes with fold-change ≥ 2.00 , $P \leq 0.05$ and known function are listed in this Table. Fold-changes and P values were calculated based on three independent biological replicates, as described in Methods. Genes discussed in this paper are shown in bold type.

TIGR ORF number	Annotation	Fold-change	P value
Downregulated genes			
Hp26695-0683	UDP- <i>N</i> -acetylglucosamine pyrophosphorylase (<i>glmU</i>)	0.237	0.009
Hp26695-1406	Biotin synthetase (<i>bioB</i>)	0.386	0.004
Hp26695-1206	Multidrug resistance protein (<i>hetA</i>)	0.402	0.009
Hp26695-0923	Outer-membrane protein (<i>omp22</i>)	0.410	0.039
Hp26695-1480	Seryl-tRNA synthetase (<i>serS</i>)	0.425	0.001
Hp26695-0684	Flagellar biosynthesis protein (<i>fliP</i>)	0.432	0.000
Hp26695-0662	RNase III (<i>rnc</i>)	0.440	0.013
Hp26695-1208	Ulcer-associated adenine-specific DNA methyltransferase	0.451	0.003
Upregulated genes			
Hp26695-1315	Ribosomal protein S19 (<i>rps19</i>)	2.053	0.047
Hp26695-1246	Ribosomal protein S6 (<i>rps6</i>)	2.073	0.002
Hp26695-1306	Ribosomal protein S14 (<i>rps14</i>)	2.082	0.030
Hp26695-0084	Ribosomal protein L13 (<i>rpl13</i>)	2.167	0.035
Hp26695-1310	Ribosomal protein S17 (<i>rps17</i>)	2.273	0.047
Hp26695-0870	Flagellar hook (<i>flgE</i>)	2.407	0.008
Hp26695-0115	Flagellin B (<i>flaB</i>)	2.432	0.007
Hp26695-1314	Ribosomal protein L22 (<i>rpl22</i>)	2.464	0.009
Hp26695-0824	Thioredoxin (<i>trxA</i>)	2.466	0.002
Hp26695-0125	Ribosomal protein L35 (<i>rpl35</i>)	2.514	0.000
Hp26695-0277	Ferredoxin	2.550	0.008
Hp26695-1076	Hypothetical protein	3.374	0.007

Table 3. Differential expression ratios of all known flagellar genes in the HP0906 mutant relative to the wild-type

Fold-changes and *P* values were calculated based upon three independent biological replicates, as described in Methods. ORFs and gene annotations are based on the TIGR database (Tomb *et al.*, 1997). The genes were assigned to previously proposed flagellar classes (Niehus *et al.*, 2004). Genes with fold-changes in expression with $P \leq 0.05$ are shown in bold type. Dashes indicate values excluded during array data analysis due to variation or technical problems with array features.

Proposed class	TIGR ORF no.	Putative gene product (gene)	$\Delta hp0906$	<i>P</i> value
Class I	HP0019	Chemotaxis protein (<i>cheV</i>)	1.175	0.16
	HP0082	Methyl-accepting chemotaxis transducer (<i>tlpC</i>)	0.952	0.50
	HP0099	Methyl-accepting chemotaxis protein (<i>tlpA</i>)	1.051	0.77
	HP0103	Methyl-accepting chemotaxis protein (<i>tlpB</i>)	1.592	0.01
	HP0173	Flagellar biosynthetic protein (<i>fliR</i>)	0.978	0.79
	HP0244	Signal-transducing protein, histidine kinase (<i>atoS</i>)	0.756	0.12
	HP0246	Flagellar basal-body P-ring protein (<i>flgI</i>)	–	–
	HP0325	Flagellar basal-body L-ring protein (<i>flgH</i>)	1.340	0.01
	HP0326	CMP- <i>N</i> -acetylneuraminic acid synthetase (<i>neuA</i>)	0.889	0.20
	HP0327	Flagellar protein G (<i>flaG</i>)	0.684	0.16
	HP0351	Basal-body M-ring protein (<i>fliF</i>)	1.266	0.05
	HP0352	Flagellar motor switch protein (<i>fliG</i>)	1.293	0.02
	HP0391	Purine-binding chemotaxis protein (<i>cheW</i>)	1.780	0.02
	HP0392	Histidine kinase (<i>cheA</i>)	1.513	0.02
	HP0393	Chemotaxis protein (<i>cheV</i>)	1.249	0.20
	HP0584	Flagellar motor switch protein (<i>fliN</i>)	1.451	0.05
	HP0599	Haemolysin secretion protein precursor (<i>hylB</i>)	1.025	0.84
	HP0616	Chemotaxis protein (<i>cheV</i>)	1.147	0.44
	HP0684	Flagellar biosynthesis protein (<i>fliP</i>)	0.432	0.00
	HP0685	Flagellar biosynthesis protein (<i>fliP</i>)	1.176†	0.44
	HP0703	Response regulator	0.855	0.46
	HP0714	RNA polymerase sigma 54 factor (<i>rpoN</i>)	0.706*	0.12
	HP0770	Flagellar biosynthesis protein (<i>flhB</i>)	0.649	0.19
	HP0815	Flagellar motor rotation protein (<i>motA</i>)	1.405	0.06
	HP0816	Flagellar motor rotation protein (<i>motB</i>)	1.009†	0.48
	HP0840	<i>flaA1</i> protein	1.211	0.26
	HP1041	Flagellar biosynthesis protein (<i>flhA</i>)	0.983	0.79
	HP1067	Chemotaxis protein (<i>cheY</i>)	1.347	0.31
	HP1092	Flagellar basal-body rod protein (<i>flgG</i>)	1.593	0.03
	HP1286	Conserved hypothetical secreted protein (<i>fliZ</i>)	1.309	0.66
	HP1419	Flagellar biosynthesis protein (<i>fliQ</i>)	0.791	0.15
	HP1420	Flagellar export protein ATP synthase (<i>fliI</i>)	0.774	0.12
	HP1462	Secreted protein involved in flagellar motility	1.324	0.00
	HP1575	Homologue of FlhB protein (<i>flhB2</i>)	1.122	0.15
	HP1585	Flagellar basal-body rod protein (<i>flgG</i>)	0.748	0.09
Class II	HP0114	Hypothetical protein	1.199	0.43
	HP0115	Flagellin B (<i>flaB</i>)	2.432*	0.01
	HP0295	Flagellin B homologue (<i>fla</i>)	1.459	0.02
	HP0869	Hydrogenase expression/formation protein (<i>hypA</i>)	1.808	0.00
	HP0870	Flagellar hook (<i>flgE</i>)	2.407*	0.01
	HP0906	Hook-length-control regulator (<i>fliK</i>)	Not valid	Not valid
	HP1076	Hypothetical protein	3.374	0.01
	HP1119	Flagellar hook-associated protein 1 (HAP1) (<i>flgK</i>)	1.442	0.04
	HP1120	Hypothetical protein	1.772	0.03
	HP1154	Hypothetical protein (operon with <i>murG</i>)	1.614	0.09
	HP1155	Transferase, peptidoglycan synthesis (<i>murG</i>)	1.668	0.06
	HP1233	Putative flagellar muraminidase (<i>flgJ</i>)	1.396	0.00
Class III	HP0472	Outer-membrane protein (<i>omp11</i>)	1.142	0.47
	HP0601	Flagellin A (<i>flaA</i>)	1.204	0.47
	HP1051	Hypothetical protein	1.083	0.59
	HP1052	UDP-3- <i>O</i> -acetyl <i>N</i> -acetylglucosamine deacetylase (<i>envA</i>)	1.013	0.94
Intermediate	HP0165	Hypothetical protein	1.422	0.44

Table 3. cont.

Proposed class	TIGR ORF no.	Putative gene product (gene)	$\Delta h p 0 9 0 6$	P value
	HP0166	Response regulator (<i>ompR</i>)	1.923*	0.06
	HP0366	Spore coat polysaccharide biosynthesis protein C	0.941	0.77
	HP0367	Hypothetical protein	1.790	0.02
	HP0488	Hypothetical protein	0.789	0.06
	HP0907	Hook-assembly protein, flagella (<i>flgD</i>)	0.909	0.40
	HP0908	Flagellar hook (<i>flgE</i>)	0.802	0.06
	HP1028	Hypothetical protein	1.271	0.17
	HP1029	Hypothetical protein	1.035	0.57
	HP1030	FliY protein (<i>fliY</i>)	1.094	0.54
	HP1031	Flagellar motor switch protein (<i>fliM</i>)	1.018	0.82
	HP1032	Alternative transcription initiation factor, sigma 28 (<i>fliA</i>)	1.120	0.18
	HP1033	Hypothetical protein	0.921	0.44
	HP1034	ATP-binding protein (<i>ylxH</i>)	0.921	0.55
	HP1035	Flagellar biosynthesis protein (<i>flhF</i>)	0.999	0.96
	HP1122	Anti-sigma 28 factor (<i>flgM</i>)	1.392	0.03
	HP1440	Hypothetical protein	0.323	0.00
	HP1557	Flagellar basal-body protein (<i>fliE</i>)	1.254	0.08
	HP1558	Flagellar basal-body rod protein (<i>flgC</i>) (proximal rod protein)	1.055	0.71
	HP1559	Flagellar basal-body rod protein (<i>flgB</i>) (proximal rod protein)	1.716	0.01
	HP0751	Polar flagellin (<i>flaG2</i>)	1.130	0.27
	HP0752	Flagellar cap protein (<i>fliD</i>)	0.831	0.19
	HP0753	Flagellar chaperone (<i>fliS</i>)	1.118	0.39
	HP0754	Flagellar chaperone (<i>fliT</i>)	–	–
Not assigned	HP0410	Flagellar sheath-associated protein (<i>hpaA2</i>)	0.944	0.52
	HP0492	Flagellar sheath-associated protein (<i>hpaA3</i>)	1.161	0.18
	HP0797	Flagellar sheath-associated protein (<i>hpaA</i>)	1.420	0.11

*Confirmatory analysis by qRT-PCR was performed for these genes.

†The values for these genes were missing in the microarray analysis and were then investigated by qRT-PCR.

Four genes, *flaB*/HP0115, *flgE*/HP0870, *motB*/HP0816 and *rpoN*/HP0714, were subsequently analysed by qRT-PCR to confirm the differential expression indicated by array data (Fig. 4). RpoN controls transcription of *flaB* and *flgE*. *motB*/HP0816 is a class I gene that is expressed at wild-type levels in the HP0906 mutant. The fold-changes obtained were in good agreement with the microarray data (Fig. 4).

RpoN-dependent gene transcription in a *flgE* mutant

The *flgE* gene of *H. pylori* flagellar hook protein is controlled by the RpoN sigma factor. Completion of the hook is an important morphogenetic check-point in *H. pylori* flagellum biogenesis (Niehus *et al.*, 2004; O'Toole *et al.*, 1994). Targeted analysis of RpoN-dependent transcription was carried out by performing qRT-PCR for the *rpoN* and *flaB* genes (Fig. 4). Mutation of *flgE* caused a significant increase in *rpoN* and *flaB* transcription.

DISCUSSION

In the present study, we investigated the effect of *fliK* mutation on RpoN activity and flagellin synthesis in *H.*

pylori. The FliK protein HP0906 controls hook length. In the model bacterium *Salmonella enterica* serovar Typhimurium, FliK is known to be involved in the switch of substrate export specificity (Minamino *et al.*, 1999). Previous analyses of flagellum regulatory mechanisms have catered for strain-specific effects by performing array analyses on multiple strains harbouring the same mutation (Niehus *et al.*, 2004). We approached this problem by first establishing the flagellar gene complement of the model strain CCUG17874, relative to two *H. pylori* genome sequences, by CGH. This analysis indicated that all known *H. pylori* flagellar genes were present in CCUG17874, consistent with the motile phenotype of that strain. This had the additional value of definitively identifying, for the first time in this commonly employed strain, the presence of several variable areas of the genome, including the plasticity zone, the *cag* pathogenicity island and the DNA modification/restriction genes. These loci have DNA with a low mol% G + C content and are known to be unstable (Alm & Trust, 1999). Most of the divergent genes in CCUG17874 were located in regions with low mol% G + C content and had no attributed function. Some low mol% G + C-content regions in the NCTC26695 and J99 genomes have also been found in self-replicating plasmids, high-

Fig. 3. Transcription analysis of the *flaB* gene region. (a) Probability profile of stress-induced DNA duplex destabilization of the *flaB* gene and linked genes, and expression ratio of the relevant genes in the HP0906 mutant. (b) Results from RT-PCR analysis of the *flaB* operon structure. Primer pairs were designed to amplify the intergenic region between the genes *hp0114* and *flaB* (region 1) and a fragment of the *flaB* transcript (region 2).

lighting that the diversity between strains is also the result of plasmid integration and horizontal transfer (Alm *et al.*, 1999). The majority of genes identified as divergent in CCUG17874 relative to the NCTC26695 and J99 reference

Fig. 4. qRT-PCR analysis of transcription of selected flagellar genes in *H. pylori fliK* and *flgE* mutants. Fold-changes and sds (error bars) were calculated relative to abundance of the *era* transcript. qRT-PCRs were performed on at least two biological replicates.

genomes encode DNA restriction/modification enzymes. Interestingly, CCUG17874 also lacked some genes encoding proteins associated with the cell envelope and pathogenesis that are likely to be involved in interaction with the host cells. It is noteworthy that the CCUG17874 genome may also contain additional genes that are strain-specific.

Global transcript analysis of the *fliK* mutant indicated that insertional inactivation of this gene increased the expression of most RpoN-dependent genes, and it did not directly modulate any other key regulators of the flagellar regulon at a transcriptional level. The class I genes encoding RpoN and the FlgR/FlgS system were transcribed at wild-type levels in the HP0906 mutant. The most affected genes in the RpoN regulon were two essential structural genes, HP0870/*flgE* (flagellar hook) and HP0115/*flaB* (minor flagellin), and also HP1076, encoding a hypothetical protein. Confirmation of the inclusion of HP1076 in the *H. pylori* RpoN regulon highlights the motivation for functional investigation of the role of this gene in motility.

HP0114 is an essential gene for motility (Schirm *et al.*, 2003). It was initially assigned to class II (RpoN-dependent genes) (Niehus *et al.*, 2004). Our array analysis suggested that HP0114 and *flaB* are not co-transcribed. Using an algorithm to predict stress-induced DNA duplex destabilization, we identified a potential regulatory region located downstream of the *flaB* gene, further supporting the suggestion that the two genes are not in fact co-transcribed. Transcription termination predictions also identified a potential intrinsic terminator downstream of *flaB*. We concluded that HP0114 is required in the flagellar regulon, but that it is not in class II (RpoN-dependent genes), as suggested in an earlier study (Niehus *et al.*, 2004), and is more likely to be in the intermediate class.

The upregulation of the RpoN-dependent genes in the *fliK* mutant suggests a transcriptional activation mechanism for RpoN-dependent genes which is possibly triggered by the FlgS/FlgR system in response to the completion of the MS ring and/or export apparatus. Presumably, this activation is normally transient in wild-type cells. After completion of the hook and progression from RpoN-dependent to FlhA-dependent transcription, a downregulation or termination mechanism should also exist to terminate the activation of the RpoN regulon. The lack of such a mechanism in the *fliK* mutant apparently results in a sustained upregulation of the RpoN-dependent genes [exemplified phenotypically by the polyhook structure in a *fliK* mutant (Ryan *et al.*, 2005b)]. We hypothesize that the signal inducing the FlgR/FlgS activation system stops upon completion of the rod and the hook. In a wild-type cell, expression of the RpoN-dependent genes, including the hook component, is thus prevented from occurring at a later stage of the flagellar assembly.

The coupling between hook assembly completion and turning off RpoN-dependent gene transcription is mecha-

nistically unclear. It is unlikely that the turn-off of the RpoN regulon is controlled by a FliA-dependent gene. In a *Campylobacter jejuni* *fliA* mutant, a poly-hook phenotype has not been reported, in contrast to that previously observed in FliK mutants (Jagannathan *et al.*, 2001). In theory, the FliK protein could be directly responsible for switching. In *S. enterica*, the FliK protein is secreted during hook polymerization (Minamino *et al.*, 1999). We tentatively suggest that the accumulation of FliK upon completion of the hook affects RpoN transcription activity by interacting directly or indirectly with the FlgR/FlgS system or the RpoN sigma factor. This is indicated in the model in Fig. 5, which builds on our previous work (Douillard *et al.*, 2008). However, no protein–protein interactions between FlgR or FlgS and FliK have so far been identified (Rain *et al.*, 2001). This suggests that the signal controlling the transcription of the class II genes does not directly involve FliK itself, but more likely another component of the flagellar apparatus, such as the flagellar protein FlhB (not shown in the model). This protein controls the switch of substrate specificity export from rod–hook–basal body class genes to filament genes in *S. enterica* serovar Typhimurium (Williams *et al.*, 1996). The fragment FlhB_C resulting from the autocatalytic cleavage of FlhB shares similarities with a novel flagellar protein, FlhX (Wand *et al.*, 2006), which may also be implicated. Upon

completion of the hook, the RpoN regulon would be turned off by a FliK-dependent mechanism, illustrated by dotted lines in Fig. 5. However, the mechanisms for this proposed onwards relay of a FliK–FlhB-transmitted signal to RpoN are not clear. The elevation of RpoN-dependent gene expression in the *flgE* mutant is compatible with this model, since the transition to FliA-mediated gene expression cannot occur in the absence of a completed hook–basal body.

We have recently shown that the HP0958 protein controls the translation of the *flaA* mRNA (Douillard *et al.*, 2008), as well as being an RpoN chaperone (Pereira & Hoover, 2005). Significantly, HP0958 also binds to FliH (Rain *et al.*, 2001), the inhibitor of the flagellum export ATPase FliI (Lane *et al.*, 2006). In the termination stages of this model (Fig. 5), the HP0958 protein becomes detached from its association with the sigma factor RpoN. In parallel, the anti-sigma factor FlgM is secreted from the cell. Thus, the *fliA*-dependent genes, including *flaA*, are then transcribed at high levels, consistent with our previously reported repression of *flaA* transcription and translation in a *fliK* mutant (Ryan *et al.*, 2005b). [Curiously, however, mutation of *flgE* does not abolish FlaA protein production (O'Toole *et al.*, 1994), though it should be noted that *flaA* transcription was not analysed in that study.] As has been

Fig. 5. Proposed and expanded model of action of FliK (HP0906) and HP0958 in *H. pylori* flagellar biogenesis. Known events or interactions are indicated by unbroken arrows, proposed or unclear events by dotted arrows. The potential roles of the membrane-bound FlhB protein, or FlhX, are omitted for clarity. The lower boxed section indicates late-stage events after flagellin gene transcription has commenced.

previously proposed, HP0958 acts as a post-transcriptional regulator on flagellin synthesis by targeting *flaA* mRNA to the export apparatus and promoting coupled FlaA translation/secretion (Douillard *et al.*, 2008). Experiments are in progress to test these hypotheses.

ACKNOWLEDGEMENTS

H. pylori flagellum research in P. W. O'T.'s lab was supported by a Science Foundation Ireland grant from the Research Frontiers Programme. We acknowledge the Wellcome Trust for supporting BpUG@S (Bacterial Microarray Group at St George's, University of London). We thank A. Zomer for valuable discussions on the manuscript and M. J. Claesson for the generation of the circular genome map.

REFERENCES

- Alm, R. A. & Trust, T. J. (1999). Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. *J Mol Med* **77**, 834–846.
- 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. & other authors (1999). Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* **397**, 176–180.
- Beier, D. & Frank, R. (2000). Molecular characterization of two-component systems of *Helicobacter pylori*. *J Bacteriol* **182**, 2068–2076.
- Bi, C. & Benham, C. J. (2004). WebSIDD: server for predicting stress-induced duplex destabilized (SIDD) sites in superhelical DNA. *Bioinformatics* **20**, 1477–1479.
- Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193.
- Brahmachary, P., Dashti, M. G., Olson, J. W. & Hoover, T. R. (2004). *Helicobacter pylori* FlgR is an enhancer-independent activator of σ^{54} -RNA polymerase holoenzyme. *J Bacteriol* **186**, 4535–4542.
- Colland, F., Rain, J.-C., Gounon, P., Labigne, A., Legrain, P. & De Reuse, H. (2001). Identification of the *Helicobacter pylori* anti- σ^{28} factor. *Mol Microbiol* **41**, 477–487.
- Douillard, F. P., Ryan, K. A., Caly, D. L., Hinds, J., Witney, A. A., Husain, S. E. & O'Toole, P. W. (2008). Posttranscriptional regulation of flagellin synthesis in *Helicobacter pylori* by the RpoN chaperone HP0958. *J Bacteriol* **190**, 7975–7984.
- Eaton, K. A., Brooks, C. L., Morgan, D. R. & Krakowka, S. (1991). Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infect Immun* **59**, 2470–2475.
- Eaton, K. A., Morgan, D. R. & Krakowka, S. (1992). Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori*. *J Med Microbiol* **37**, 123–127.
- Eaton, K. A., Suerbaum, S., Josenhans, C. & Krakowka, S. (1996). Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. *Infect Immun* **64**, 2445–2448.
- EUROGAST (1993). An international association between *Helicobacter pylori* infection and gastric cancer. *Lancet* **341**, 1359–1362.
- Ferris, H. U. & Minamino, T. (2006). Flipping the switch: bringing order to flagellar assembly. *Trends Microbiol* **14**, 519–526.
- Frenck, R. W., Jr & Clemens, J. (2003). *Helicobacter* in the developing world. *Microbes Infect* **5**, 705–713.
- Galkin, V. E., Yu, X., Bielnicki, J., Heuser, J., Ewing, C. P., Guerry, P. & Egelman, E. H. (2008). Divergence of quaternary structures among bacterial flagellar filaments. *Science* **320**, 382–385.
- Hinds, J., Laing, K. G., Mangan, J. A. & Butcher, P. D. (2002a). Glass slide microarray for bacterial genomes. In *Methods in Microbiology: Functional Microbial Genomics*, pp. 83–99. Edited by B. W. Wren & N. Dorell. London: Academic Press.
- Hinds, J., Witney, A. A. & Vass, J. K. (2002b). Microarray design for bacterial genomes. In *Methods in Microbiology: Functional Microbial Genomics*, pp. 67–82. Edited by B. W. Wren & N. Dorell. London: Academic Press.
- Jagannathan, A., Constantinidou, C. & Penn, C. W. (2001). Roles of *rpoN*, *fliA*, and *flgR* in expression of flagella in *Campylobacter jejuni*. *J Bacteriol* **183**, 2937–2942.
- Josenhans, C., Niehus, E., Amersbach, S., Horster, A., Betz, C., Drescher, B., Hughes, K. T. & Suerbaum, S. (2002). Functional characterization of the antagonistic flagellar late regulators FliA and FlgM of *Helicobacter pylori* and their effects on the *H. pylori* transcriptome. *Mol Microbiol* **43**, 307–322.
- Kikuchi, S. & Dore, M. P. (2005). Epidemiology of *Helicobacter pylori* infection. *Helicobacter* **10**, 1–4.
- Kim, J. S., Chang, J. H., Seo, W. Y., Yu, G. J., Chung, S. I. & Yum, J. S. (2000). Cloning and characterization of a 22 kDa outer-membrane protein (Omp22) from *Helicobacter pylori*. *Mol Cells* **10**, 633–641.
- Kim, C. C., Joyce, E. A., Chan, K. & Falkow, S. (2002). Improved analytical methods for microarray-based genome-composition analysis. *Genome Biol* **3**, research0065.
- Kingsford, C. L., Ayanbule, K. & Salzberg, S. L. (2007). Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* **8**, R22.
- Lane, M. C., O'Toole, P. W. & Moore, S. A. (2006). Molecular basis of the interaction between the flagellar export proteins FliH and FliI from *Helicobacter pylori*. *J Biol Chem* **281**, 508–517.
- Logan, R. P. & Berg, D. E. (1996). Genetic diversity of *Helicobacter pylori*. *Lancet* **348**, 1462–1463.
- Minamino, T., Gonzalez-Pedrajo, B., Yamaguchi, K., Aizawa, S. I. & Macnab, R. M. (1999). FliK, the protein responsible for flagellar hook length control in *Salmonella*, is exported during hook assembly. *Mol Microbiol* **34**, 295–304.
- Niehus, E., Gressmann, H., Ye, F., Schlapbach, R., Dehio, M., Dehio, C., Stack, A., Meyer, T. F., Suerbaum, S. & Josenhans, C. (2004). Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori*. *Mol Microbiol* **52**, 947–961.
- O'Toole, P. W., Kostrzynska, M. & Trust, T. J. (1994). Non-motile mutants of *Helicobacter pylori* and *Helicobacter mustelae* defective in flagellar hook production. *Mol Microbiol* **14**, 691–703.
- Parsonnet, J., Hansen, S., Rodriguez, L., Gelb, A. B., Warnke, R. A., Jellum, E., Orentreich, N., Vogelman, J. H. & Friedman, G. D. (1994). *Helicobacter pylori* infection and gastric lymphoma. *N Engl J Med* **330**, 1267–1271.
- Pereira, L. & Hoover, T. R. (2005). Stable accumulation of σ^{54} in *Helicobacter pylori* requires the novel protein HP0958. *J Bacteriol* **187**, 4463–4469.
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45.
- Rain, J. C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J. & other authors (2001). The protein-protein interaction map of *Helicobacter pylori*. *Nature* **409**, 211–215.

- Rozen, S. & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365–386.
- Ryan, K. A., Karim, N., Worku, M., Moore, S. A., Penn, C. W. & O'Toole, P. W. (2005a). HP0958 is an essential motility gene in *Helicobacter pylori*. *FEMS Microbiol Lett* **248**, 47–55.
- Ryan, K. A., Karim, N., Worku, M., Penn, C. W. & O'Toole, P. W. (2005b). *Helicobacter pylori* flagellar hook-filament transition is controlled by a FliK functional homolog encoded by the gene HP0906. *J Bacteriol* **187**, 5742–5750.
- Scarlato, V., Delany, I., Spohn, G. & Beier, D. (2001). Regulation of transcription in *Helicobacter pylori*: simple systems or complex circuits? *Int J Med Microbiol* **291**, 107–117.
- Schirm, M., Soo, E. C., Aubry, A. J., Austin, J., Thibault, P. & Logan, S. M. (2003). Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Mol Microbiol* **48**, 1579–1592.
- Sebert, M. E., Palmer, L. M., Rosenberg, M. & Weiser, J. N. (2002). Microarray-based identification of *htrA*, a *Streptococcus pneumoniae* gene that is regulated by the CiaRH two-component system and contributes to nasopharyngeal colonization. *Infect Immun* **70**, 4059–4067.
- 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. & other authors (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539–547.
- van Hijum, S. A., Baerends, R. J., Zomer, A. L., Karsens, H. A., Martin-Requena, V., Trelles, O., Kok, J. & Kuipers, O. P. (2008). Supervised Lowess normalization of comparative genome hybridization data – application to lactococcal strain comparisons. *BMC Bioinformatics* **9**, 93.
- Veldhuyzen van Zanten, S. J. & Sherman, P. M. (1994). *Helicobacter pylori* infection as a cause of gastritis, duodenal ulcer, gastric cancer and nonulcer dyspepsia: a systematic overview. *CMAJ* **150**, 177–185.
- Wand, M. E., Sockett, R. E., Evans, K. J., Doherty, N., Sharp, P. M., Hardie, K. R. & Winzer, K. (2006). *Helicobacter pylori* FlhB function: the FlhB C-terminal homologue HP1575 acts as a “spare part” to permit flagellar export when the HP0770 FlhBCC domain is deleted. *J Bacteriol* **188**, 7531–7541.
- Williams, A. W., Yamaguchi, S., Togashi, F., Aizawa, S.-I., Kawagishi, I. & Macnab, R. M. (1996). Mutations in *fliK* and *flhB* affecting flagellar hook and filament assembly in *Salmonella typhimurium*. *J Bacteriol* **178**, 2960–2970.

Edited by: R. J. Maier