Flagellar *flhA*, *flhB* and *flhE* genes, organized in an operon, cluster upstream from the *inv* locus in *Yersinia enterocolitica*

Alan Fauconnier,¹ Abdelmounaïm Allaoui,^{2†} Andrés Campos,^{3‡} Ary Van Elsen,¹ Guy R. Cornelis² and Alex Bollen¹

Author for correspondence: Alan Fauconnier. Tel: +32 67 889 425. Fax: +32 67 889 472.
e-mail: afaucon@sga.ulb.ac.be

¹ Applied Genetics, Université Libre de Bruxelles, Rue de l'Industrie 24, B-1400 Nivelles, Belgium

² Microbial Pathogenesis Unit, International Institute of Cellular and Molecular Pathology and Faculté de Médecine, Université Catholique de Louvain, Avenue Hippocrate 74, B-1200 Bruxelles, Belgium

³ Departamento de Bioenergética, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, México 04510 DF, Mexico

The *inv* gene of *Yersinia enterocolitica* codes for invasin, a member of the invasin/intimin-like protein family, which mediates the internalization of the bacterium into cultured epithelial cells. The putative inclusion of *inv* into a pathogenicity island was tested by investigating its flanking sequences. Indeed, the enteropathogenic *Escherichia coli* (EPEC) intimin, a member of the same family of proteins, is encoded by *eaeA*, a gene which belongs to a pathogenicity island. An ORF located upstream from *inv* was of particular interest since it appeared homologous both to the flagellar *flhA* gene and to *sepA*, an EPEC gene lying inside the same pathogenicity island as *eaeA*. A mutant in this ORF was non-motile and non-flagellated while its invasion phenotype remained unaffected. These data indicated that the ORF corresponded to the *flhA* gene of *Y. enterocolitica*. Subsequently, the *flhB* and *flhE* genes, located respectively upstream and downstream from *flhA*, were identified. The three *flh* genes appear to be transcribed from a single operon called *flhB*, according to the nomenclature used for *Salmonella typhimurium*. Intergenic sequence between *flhE* and *inv* includes a grey hole, with no recognizable function. Downstream from *inv*, we have detected the flagellar *flgM* operon as already reported. Finally, the incongruous localization of *inv* amidst the flagellar cluster is discussed; while transposition could explain this phenomenon, no trace of such an event was detected.

Keywords: *Yersinia enterocolitica*, invasin, LcrD/FlbF, flagellar genes, motility

INTRODUCTION

Pathogenic strains of the genus *Yersinia* harbour a virulence plasmid called pYV which, at 37 °C in the absence of Ca²⁺ ions, restricts bacterial growth and governs secretion of Yop proteins (Cornelis, 1994). Yops lack a conventional signal peptide sequence (Michiels *et al.*, 1990) and are secreted via a new pathway (Michiels

& Cornelis, 1991) called type III by Salmond & Reeves (1993). Various pYV-encoded determinants take part in this secretion mechanism. One of these, LcrD, became the archetype of the recently identified LcrD/FlbF family of proteins (Plano *et al.*, 1991). Members of this family are involved in host-cell invasion (Andrews & Maurelli, 1992; Galán *et al.*, 1992), virulence in several phytopathogenic bacteria (Fenselau *et al.*, 1992; Gough *et al.*, 1993; Huang *et al.*, 1993; Wei & Beer, 1993) or flagellar assembly (Carpenter & Ordal, 1993; Gygi *et al.*, 1995; Kim, 1989; Miller *et al.*, 1993; Minamino *et al.*, 1994; Ramakrishnan *et al.*, 1991).

Chromosomal loci also participate in yersinial virulence. Among these, the *inv* gene encodes invasin, a surface protein that mediates internalization of the bacterium into mammalian cells (Isberg & Falkow, 1985; Isberg *et al.*, 1987). The invasin of *Yersinia enterocolitica* has been shown to play a role in the early steps of the

[†] **Present address:** Department of Microbiology, Medical School, University of Newcastle-upon-Tyne, Framlington Place, Newcastle-upon-Tyne, NE2 4HH, UK.

[‡] **Present address:** Department of Microbiology and Immunology, College of Medicine, E-603 Medical Science Building, 835, South Wolcott Avenue, Chicago, IL 60612-7344, USA.

Abbreviations: EM, electron microscopy; EPEC, enteropathogenic *Escherichia coli*.

The EMBL accession number for the sequence reported in this paper is Z48169.

infection (Pepe & Miller, 1993). Yersinial invasins share similarity with intimin proteins (Jerse *et al.*, 1990), found in various bacterial genera (Frankel *et al.*, 1994; Schauer & Falkow, 1993; Yu & Kaper, 1992) and involved in the development of the so-called attaching and effacing (AE) lesion (Moon *et al.*, 1983). In enteropathogenic *Escherichia coli* (EPEC), the gene encoding intimin, belongs to a pathogenicity island termed LEE (locus of enterocyte effacement) which encodes all known determinants involved in AE lesion formation (McDaniel *et al.*, 1995). One of these, the *sepA* gene product, is a member of the LcrD/FlbF protein family (Jarvis *et al.*, 1995).

The close physical association in EPEC of genes encoding invasin/intimin-like proteins and other virulence genes prompted us to study the sequences around the *inv* locus of *Y. enterocolitica*. Indeed, a partial ORF homologous to *sepA* of EPEC and to the flagellar gene *flhA* from *E. coli* and *Salmonella typhimurium* (Kim, 1989; Minamino *et al.*, 1994) was detected immediately upstream from the *inv* gene. We thus wondered if a pathogenicity island was located in the region, as in EPEC, or if, on the contrary, only flagellar determinants were present.

In the present report, we have described the physical and functional characterization of the ORF located upstream from *inv*. We have shown that it corresponds to the flagellar gene *flhA* and that it forms the second gene of an operon comprising, in addition to *flhA*, the *flhB* and *flhE* genes.

METHODS

Bacterial strains, plasmids and growth conditions. *E. coli* HB101 (Boyer & Roulland-Dussoix, 1969) was used for genetic manipulations with standard plasmid vectors. SY327(λ pir⁺) (Kolter *et al.*, 1978) was the intermediate host for constructs involving the suicide vector pKNG101 (Kaniga *et al.*, 1991). Suicide plasmids were transferred to *Y. enterocolitica* from the donor strain *E. coli* SM10(λ pir⁺) (Simon *et al.*, 1983) which encodes the R6K replication protein pi and the transfer functions of RP4. Recombinant pLAFR3 cosmids of the *Y. enterocolitica* W1024 gene library, kindly provided by I. Delor (Delor *et al.*, 1990), were maintained in the host *E. coli* VCS257 (Stratagene). T7 expression of cloned genes was achieved in *E. coli* LK111 strain (C600 Res⁻ Mod⁺ lacY⁺ lacZAM15 lacI; received from M. Zabeau, Ghent, Belgium) containing pGP1-2 (Tabor & Richardson, 1985) and recombinant plasmids harbouring the genes of interest. *Y. enterocolitica* W1024 is a serotype O:9 strain from the collection of G. Wauters (Brussels, Belgium). KNG1024 is a *bla* mutant isogenic to W1024 (Kaniga *et al.*, 1991). MS22703 is a *Y. enterocolitica* *inv* mutant (Sory & Cornelis, 1994). The yersinia strains used in invasion assays were cured of their pYV plasmids by culture at 37 °C on MOX agar, as described below.

The *E. coli* strains were generally grown in L broth or on L agar. The *Y. enterocolitica* strains were routinely grown in Tryptic Soy Broth (Oxoid) and plated on Tryptic Soy Agar (Oxoid), sometimes supplemented with 20 mM MgCl₂ and 20 mM Na-oxalate (MOX). Antibiotics were used at the following concentrations: 12.5 µg tetracycline ml⁻¹; 35 µg

nalidixic acid, kanamycin and streptomycin ml⁻¹; 40 µg gentamicin ml⁻¹; 150 µg ampicillin ml⁻¹. Plasmids are listed in Table 1.

Invasion assay. Bacterial entry into cultured cells was assayed as described by Isberg & Falkow (1985) and Isberg *et al.* (1987). Briefly, saturated bacterial cultures were washed in PBS and aliquots of 1.5 × 10⁷ bacteria were added to monolayer cultures of 2 × 10⁵ HEP-2 cells per microtitre well in RPMI 1640 medium (Life Technologies). Infected cultures were incubated at 37 °C for 3 h in a 5% CO₂ atmosphere. After allowing entry of bacteria into the animal cells, non-adherent bacteria were removed from the monolayers by washing three times with PBS. Monolayers were then incubated in medium containing gentamicin for one hour at 37 °C and washed three more times in PBS. Internalized bacteria were released from the monolayers by addition of 1% Triton X-100 and then titred on L agar plates. Results are expressed in percentages, as the ratio of viable counts recovered after gentamicin treatment to the original input.

Construction of a *flhA* mutant. To generate the NIVh81 mutant, a 244 bp fragment (codons 512–592) was deleted from the *flhA* segment present on pNIV1907 and replaced by a cassette containing the *aphA*-3 gene which confers kanamycin resistance (Ménard *et al.*, 1993): the cassette was excised from pUC18K by *Sma*I digestion and introduced in the *Pvu*II-filled *Asp*7181 sites present within *flhA* (Fig. 1). This construct generated an early stop in *flhA* translation and allowed in-frame translation of the remaining 3' end of the mutated gene, avoiding possible polar effects on expression of downstream cistrons. The mutated *flhA* with its flanking sequences, from a *Pvu*II site external to *flhA* to *Bsu*36I, was subsequently introduced in the suicide vector pKNG101 yielding pNIV1937 (Fig. 1). *E. coli* SM10(λ pir⁺) was used to replicate and mobilize pNIV1937. Mating with *Y. enterocolitica* KNG1024 was done overnight on plates at 32 °C and the yersinia clones with pNIV1937 integrated in their chromosome were selected on kanamycin and nalidixic acid. The second recombination event was selected for resistance to sucrose and sensitivity to streptomycin (Kaniga *et al.*, 1991). The disruption of *flhA* in NIVh81 was confirmed by sequencing pNIV1940 and pNIV1941, two plasmids displaying *aphA*-3 and its flanking sequences extracted from NIVh81 (see below).

Motility assay. We tested *Y. enterocolitica* strains for motility by spotting colonies onto Tryptic Soy swarm plates, containing 0.3% agar, supplemented with appropriate antibiotics. The plates were incubated for up to 24 h at 24 °C and the colonies were examined each day for zones of motility.

Electron microscopy (EM). The presence of flagella on *Y. enterocolitica* cells was investigated by EM. To this end, bacteria from swarm plates were grown for 24 h at 24 °C without agitation in 10 ml LB broth supplemented with appropriate antibiotics. To observe their ability to swim before examination by EM, cells from swarm plates and LB cultures were observed with a dark-field light microscope (Macnab, 1976). Five-hundred microliters of each culture was harvested, washed and resuspended in 100 µl 10 mM Tris, pH 7.5. Five microlitres of the cells were adsorbed for 2 min to 200 hex mesh copper grid with a carbon-coated support film. Samples were drained against filter paper and, after negative staining for 1 min with 1% (w/v) uranyl acetate, grids were examined with a JEOL 1200 Ex11 transmission microscope operating at 60 kV.

Cloning of the missing part of *flhA*. Two genomic libraries of NIVh81, a *Y. enterocolitica* *flhA* mutant, were constructed by

Table 1. Plasmids used

Plasmid	Relevant properties*	Source/reference
pID1902	pLAFR3 + 22 kb insert containing <i>invA</i>	This study
pNIV1907	pUC19 + 7500 bp <i>EcoRV</i> fragment from pNIV1902	This study
pNIV1909	pUC19 + 8400 bp <i>PvuII</i> fragment from pNIV1902	This study
pNIV1937	pKNG101 + <i>PvuII</i> - <i>Bsu36I</i> fragment from pNIV1907 <i>flhA</i> (Δ 512-592):: <i>aphA</i> -3 (nonpolar mutation)	This study
pNIV1940	pUC18 + 5 kb <i>NsiI</i> fragment containing <i>flhA</i> (Δ 512-592):: <i>aphA</i> -3 and flanking sequences from NIVh81	This study
pNIV1941	pUC18 + 11 kb <i>BglII</i> fragment containing <i>flhA</i> (Δ 512-592):: <i>aphA</i> -3 and flanking sequences from NIVh81	This study
pNIV1943	pBC18R + 1458 bp <i>Asp718I</i> - <i>BsaAI</i> fragment from pNIV1940 and 860 bp <i>BsaAI</i> - <i>Clal</i> fragment from pNIV1907, encompassing together the complete <i>flhA</i> gene	This study
pNIV1945	pTM100 + 2340 bp <i>EcoRI</i> - <i>HindIII</i> fragment from pNIV1943, containing the complete <i>flhA</i> sequence	This study
pLAFR3	Mobilizable cosmid used for the <i>Y. enterocolitica</i> gene library construction	Staskawicz <i>et al.</i> (1987)
pUC18/pUC19	Amp ^R cloning vectors	Norlander <i>et al.</i> (1983)
pTZ18R	pUC18 + T7 RNA polymerase ϕ 10 promoter	Pharmacia
pBC18R	pTZ18R + <i>oriT</i>	China <i>et al.</i> (1990)
pUC18K	850 bp <i>aphA</i> -3 inserted in the <i>SmaI</i> site of pUC18	Ménard <i>et al.</i> (1993)
pTM100	pACYC184 + <i>oriT</i> _{RR2} cloned in the <i>SstII</i> site (co-ordinate 832 bp)	Michiels & Cornelis (1991)
pKNG101	<i>pir</i> - <i>ori</i> _{R6K} <i>ori</i> _{RR2} <i>strAB</i> <i>sacBR</i>	Kaniga <i>et al.</i> (1991)
pGP1-2	pBR322 + cI857 + T7 RNA polymerase from promoter pL	Tabor & Richardson (1985)

*(Δ 512-592), numbers in parentheses are the codons deleted from the *flhA* gene.

ligating either *NsiI*- or *BglII*-digested DNA fragments into *PstI*- or *BamHI*-linearized pUC18. Kanamycin-resistant clones, supposed to harbour the mutated *flhA* locus, were isolated from each DNA library. pNIV1940 from the *NsiI* shared DNA library and pNIV1941 from the *BglII* shared one were chosen for further sequencing (Fig. 1).

Sequencing strategy. Nucleotide sequencing was performed by the dideoxy chain-termination method (Sanger *et al.*, 1977) on denatured dsDNA using Sequenase (USB). Large DNA fragments were sequenced in both orientations according to a step-by-step procedure in which synthetic primers for forward sequencing were designed from previously released sequences. The entire pNIV1907 and pNIV1909 inserts as well as the 5' half insert of pNIV1940, upstream from the *BamHI* site, were sequenced. The *BamHI* site of pNIV1940 corresponded to the *Sau3AI* site of pNIV1907 at which ligation occurred, during the gene library construction (Fig. 1). Only the 3' end of the pNIV1941 insert was sequenced, revealing that the *BglII* site at one end of the cloned fragment belonged to the *inv* gene. This observation confirms that recombination occurred at the expected locus (see above). If necessary, this also would confirm that *flhA* linkage to *inv* was not a consequence of an accidental ligation of *Sau3AI* digests, but no *Sau3AI* sites occurred between *flhA* and *inv* preventing such artefact.

RNA extraction and Northern blot analysis. Total RNA was extracted from cells by hot phenol extraction as described by Lambert de Rouvroit *et al.* (1992). Gels for Northern blot hybridization were 0.8% agarose/5.5% formaldehyde and were blotted onto Hybond-N membranes (Amersham). RNA was probed with four synthetic oligonucleotides complemen-

tary to the *flhB*, *flhA*, *flhE* and *inv* coding sequences. The sequences were TTTGGCGCAGCATTTGCTTATC, CGA-TACCGATGGCGAAGTTA, GTGCGACCGGTGATT-GCAGC and GCGCCGATATTAAGGGTATTG, respectively. Hybridizations were carried out at 37 °C, in 5 × SSC, 50 mM phosphate buffer (pH 7.8), 1 × Denhardt's solution and 0.5% SDS. Membranes were washed three times for 5 min in 2 × SSC, 0.5% SDS at room temperature.

Identification of the *flhA* gene product. Specific labelling of FlhA was achieved in *E. coli* by using the T7 expression system (Tabor & Richardson, 1985). pBC18R (China *et al.*, 1990) is a mobilizable derivative of pTZ18R (Pharmacia) containing the T7 RNA polymerase ϕ 10-promoter. The *flhA* gene was introduced in pBC18R as follows. The 1458 bp *Asp718I*-*BsaAI* fragment from pNIV1940, including the ribosome-binding site and the first 470 codons from *flhA*, together with the 860 bp *BsaAI*-*Clal* fragment from pNIV1907 which contains the 3' end of the gene, were ligated in pBC18R linearized by *Asp718I*/*AccI*, yielding pNIV1943 (Fig. 1). The protein of interest was labelled using *E. coli* LK111 containing both pGP1-2 and either pBC18R or pNIV1943, as described previously (Allaoui *et al.*, 1994).

Mutant complementation. Complementation tests were achieved after the introduction of an intact copy of *flhA* in the *Y. enterocolitica* mutant. For this purpose, the complete *flhA* gene contained within the 2363 bp *EcoRI* protruding-*HindIII* blunt-ended fragment from pNIV1943 was introduced in the *EcoRI*-*ScaI* sites of the low-copy-number plasmid pTM100 (Michiels & Cornelis, 1991) and the resulting construction, pNIV1945 (Fig. 1), was mobilized into the yersinia mutant. In

pNIV1945, *flhA* expression occurs under the control of a pTM100 promoter that normally triggers transcription of a chloramphenicol-resistance gene.

Computer analysis. All algorithms used in this study are part of the Genetics Computer Group (GCG) package (Devereux *et al.*, 1984). Particularly, the coding probabilities were given by the algorithm TESTCODE devised by Fickett (1982), using a window size of 200 bp. Database comparisons were performed with EMBL released 44.

RESULTS

Cloning of the *inv* gene and analysis of flanking sequences

A cosmid library from *Y. enterocolitica* strain W1024, distributed into twenty pools of 48 clones, was screened in an invasion assay. One cosmid, which was able to confer an invasive phenotype to *E. coli* VCS257 strain

Fig. 1. Structure of the cosmid and plasmids carrying *inv*, *flhA* and *aphA-3*. Open bars indicate the sequences which are actually flanking the *inv* gene, itself depicted in thinly hatched box. The DNA fragment of chimerical origin harboured by pID1902, pNIV1907 and pNIV1937 is indicated by a thick hatched box. The *flhA* partial or complete coding sequence is indicated by a black box, while the kanamycin-resistance gene (*aphA-3*) is displayed in squared box. The vector names are indicated in parentheses. The thin arrows represent the $\phi 10$ promoter of bacteriophage T7 (pNIV1943) and the *cat* promoter (pNIV1945); the thick arrow corresponds to the *lac* promoter (pNIV1943). Horizontal arrows indicate the direction of transcription of *flhA*, *inv* and *aphA-3*. The scale is only valid for the expanded area and therefore does not refer to that of pID1902, pNIV1907 and pNIV1909. The position of selected restriction sites are shown: A, *Asp*718I; B, *Bsa*AI; Ba, *Bam*HI; Bg, *Bgl*II; Bs, *Bsu*36I; C, *Cl*AI; E, *Eco*RV; N, *Nsi*I; P, *Pvu*II; S, *Sma*I; Sa, *Sau*3AI.

Table 2. Penetration of *Y. enterocolitica* into HEp-2 cultured cells

Strain	Relevant characteristic	Level of invasion (%) [*]
KNG1024	Wild-type	5.2 ± 2.3
NIVh81	<i>flhA</i> mutant	10.3 ± 7.3
MS22703	<i>inv</i> mutant	0.0150 ± 0.0044

^{*} Values are the mean of three determinations. $P \leq 0.05$ by *t*-test.

that normally does not penetrate cultured cells, was isolated and named pID1902. *Eco*RV and *Pvu*II fragments from pID1902 were subcloned in pUC19. Transformed bacteria presenting an invasive phenotype harboured either pNIV1907 which contained a 7.4 kb *Eco*RV DNA fragment or pNIV1909 bearing a *Pvu*II insert of 8.2 kb. The overlapping region of pNIV1907 and pNIV1909 inserts contained the *inv* gene (Fig. 1). Downstream from *inv*, we identified the *flgM* operon, containing the *flgM* and *flgN* genes, as described by Kapatral *et al.* (1996). Upstream from *inv*, we detected a 823 bp ORF, temporarily noted *orf236*. This ORF was similar to the *sepA* gene of EPEC (Jarvis *et al.*, 1995) which clusters with *eaeA*, a gene partially similar to *inv* (Jerse *et al.*, 1990), to *invA* of *S. typhimurium* and *mxIA* of *Shigella flexneri*, both involved in host-cell invasion (Andrews & Maurelli, 1992; Galán *et al.*, 1992), and finally to the flagellar genes *flhA* from *E. coli* and *S. typhimurium* (Kim, 1989; Minamino *et al.*, 1994).

A mutant affected in *orf236* is fully invasive *in vitro*

To assess the function of *orf236*, we engineered a non-polar mutant by allelic exchange. In the resulting mutant, called NIVh81, *orf236* was disrupted by the *aphA-3* cassette which was inserted such that translation was not interrupted. Its putative involvement in invasion-mediated entry was tested in an invasion assay. No significant difference was observed between the invasion levels of the mutant and that of the parental strain. In contrast, an *inv* mutant was respectively 350- and 700-fold less invasive than the wild-type strain KNG1024 and NIVh81 (Table 2). The invasion assays were performed with pYV⁻ variants for two reasons; first, we wanted to exclude a putative complementation of the mutation by *lcrD*, a pYV homologue to *orf236*, and secondly, because pYV-encoded proteins, expressed under invasion assay conditions, are deleterious to the cultivated cell monolayers (Portnoy *et al.*, 1981).

The NIVh81 mutant is non-motile and non-flagellated

When the NIVh81 mutant was spotted onto a semisolid medium, it produced small, compact colonies (Fig. 2b), indicating that it was non-motile, in contrast to the

Fig. 2. Influence of *flhA* mutation on motility in (a) KNG1024 (wild-type strain), (b) NIVh81 (*flhA* mutant) and (c) NIVh81(pNIV1945).

Fig. 3. Electron micrographs of *Y. enterocolitica* wild-type and mutant strain. The wild-type strain KNG1024 (a–c), the mutant NIVh81 (d, e) and the complemented mutant NIVh81(pNIV1945) (f, g) were grown at 24 °C and examined for flagella. While KNG1024 appears peritrichously flagellated, the mutant NIVh81 shows a general lack of flagella. Complementation of the mutation is not totally achieved in NIVh81(pNIV1945) since this strain displays only short and scattered flagella. The scale bar, shown in (e) and valid for all parts, represents 1 µm.

spreading colonies formed by the parent *Y. enterocolitica* strain, KNG1024 (Fig. 2a).

Y. enterocolitica KNG1024 and the NIVh81 mutant were grown at 24 °C and examined by EM. No flagella could be seen in NIVh81 (Fig. 3d, e), while the wild-type strain appeared peritrichously flagellated (Fig. 3a–c).

Cloning of the missing part of *orf236* and sequence analysis

The homology of *orf236* with the genes encoding LcrD/FlbF protein family is interrupted at a *Sau3AI* recognition site at the 5' end, suggesting that the cloned sequence was incomplete. The putatively missing part of

Table 3. Proteins with sequence homology to Orf1 and Orf236

Protein (organism)	Compared to	Identity (%) [*]	Similarity (%) [*]	Reference
FlhB (<i>Escherichia coli</i>)	Orf1	64	82	Kim (1989)
FlhB (<i>Salmonella typhimurium</i>)	Orf1	65	83	Minamino <i>et al.</i> (1994)
FlhB (<i>Bacillus subtilis</i>)	Orf1	36	59	Carpenter <i>et al.</i> (1993)
YscU (<i>Yersinia</i> spp.)	Orf1	30	58	Allaoui <i>et al.</i> (1994)
SpaS (<i>Salmonella typhimurium</i>)	Orf1	25	51	Groisman & Ochman (1993)
Spa40 (<i>Shigella flexneri</i>)	Orf1	25	53	Sasakawa <i>et al.</i> (1993)
HrpN (<i>Pseudomonas solanacearum</i>)	Orf1	31	56	Gough <i>et al.</i> (1993)
FlhA (<i>Escherichia coli</i>)	Orf236	85	92	Kim (1989)
FlhA (<i>Salmonella typhimurium</i>)	Orf236	85	92	Minamino <i>et al.</i> (1994)
FlhA (<i>Proteus mirabilis</i>)	Orf236	78	88	Gygi <i>et al.</i> (1995)
FlhA (<i>Bacillus subtilis</i>)	Orf236	41	67	Carpenter & Ordal (1993)
FlbF (<i>Caulobacter crescentus</i>)	Orf236	40	66	Ramakrishnan <i>et al.</i> (1991)
FlbA (<i>Campylobacter jejuni</i>)	Orf236	39	66	Miller <i>et al.</i> (1993)
LcrD (<i>Yersinia</i> spp.)	Orf236	35	61	Plano <i>et al.</i> (1991)
InvA (<i>Salmonella typhimurium</i>)	Orf236	35	60	Galán <i>et al.</i> (1992)
MxiA (<i>Shigella flexneri</i>)	Orf236	34	59	Andrews & Maurelli (1992)
SepA (EPEC)	Orf236	33	58	Jarvis <i>et al.</i> (1995)
HrpC2 (<i>Xanthomonas campestris</i>)	Orf236	35	62	Fenselau <i>et al.</i> (1992)
HrpO (<i>Pseudomonas solanacearum</i>)	Orf236	35	59	Gough <i>et al.</i> (1993)
HrpI (<i>Pseudomonas syringae</i>)	Orf236	37	63	Huang <i>et al.</i> (1993)
HrpI (<i>Erwinia amylovora</i>)	Orf236	37	64	Wei & Beer (1993)

^{*} Percentage identity and similarity to Orf1 and Orf236 proteins were determined by the GAP program from the GCG package.

orf236 was cloned directly from chromosomal DNA after digestion with enzymes that cut less frequently than *Sau3AI*. The NIVh81 mutant strain was chosen as the source of *Y. enterocolitica* DNA for the library construction. Taking advantage of the presence of *aphA-3* gene in the locus to be cloned, we selected a kanamycin-resistant clone and called its resident plasmid pNIV1940. In pNIV1940, a *Bam*HI site corresponding to the *Sau3AI* site of pNIV1907, was preceded by a sequence presenting the expected similarity with the genes encoding members of the LcrD/FlbF protein family. The reconstituted sequence of 5' flanking *inv* region revealed the complete ORF which extended over 692 codons, specifying a 74966 Da polypeptide. Despite its overall similarity with LcrD/FlbF homologues (Table 3), the highest level of identity was observed with flagellar proteins.

Besides *orf236*, two other ORFs (*orf1* and *orf3*) were detected. *orf1* encoded a 383 amino acid polypeptide with a predicted molecular mass of 42424 Da and showing similarity to FlhB of *E. coli* and *S. typhimurium* (Kim, 1989; Minamino *et al.*, 1994) and other known FlhB homologues (Table 3). While similarities were detected all along the amino acid sequences of these proteins, conservation of putative cytoplasmic domains was much more apparent, as already reported (Allaoui *et al.*, 1994). *orf3* was homologous to the *flhE* genes of *E. coli* and *S. typhimurium* (Kim, 1989; Minamino *et al.*, 1994). It encoded a putative 140 amino acid polypeptide (14878 Da) which was 37% identical to FlhE from these bacteria.

Transcriptional organization of *orf1*, *orf236* and *orf3*

orf1, *orf236* and *orf3* appeared to belong to a single operon, similar in organization to the *flhB* operon of *S. typhimurium* which contains the *flhB*, *flhA* and *flhE* genes (Macnab, 1992). Indeed, the first ATG codon of *orf236* (position 1662 in the deposited sequence) overlapped the TGA stop codon of *orf1*; likewise the stop codon of *orf236* overlapped the start codon of *orf3* (position 3740). No typical promoter sequences preceded the ribosome-binding sites of *orf236* and *orf3*. However, the sequence GCCGATGG (positions 353–360), lying 154 bp upstream from the putative GTG initiation codon of *orf1*, closely matched the consensus sequence GCCGATAA of flagellar class 2 promoters in *E. coli* (Bartlett *et al.*, 1988).

Northern blot analysis confirmed such organization. A single transcript was revealed by probes that specifically hybridized to *orf1*, *orf236* or *orf3*. On the other hand, the *inv* gene message banded at a different size (Fig. 4a). These data indicated that the three ORFs located upstream from *inv* are transcribed as a single transcriptional unit which is different from that of the *inv* gene.

The *inv* locus is surrounded by flagellar genes

The high level of identity of *orf1*, *orf236* and *orf3* with known flagellar genes, their organization into an operon, the presence of an expected alternative transcription

Fig. 4. Northern blot analysis of *flhBAE* and *inv*. (a). Whole-cell RNA from *Y. enterocolitica* KNG1024 was extracted from late-exponential-phase culture ($OD_{600} = 1$) grown at 24 °C. The DNA probes specifically hybridized to *flhB* (lane 1), *flhA* (lane 2), *flhE* (lane 3) and *inv* (lane 4). (b) RNA from the same strain was extracted from late-exponential-phase ($OD_{600} = 1$) and early-stationary-phase ($OD_{600} = 2$) cultures grown either at 24 °C (lanes 1 and 2, respectively), or at 37 °C (lanes 3 and 4, respectively). The message was revealed by the DNA *flhE*-specific probe.

control signal as well as the phenotype of the NIVh81 mutant indicated that these ORFs correspond to flagellar genes. We propose to name them according to the nomenclature used for *S. typhimurium* and *E. coli* equivalent genes (Iino *et al.*, 1988). *orf1*, *orf236* and *orf3* will therefore be referred to as *flhB*, *flhA* and *flhE*.

The *inv* gene was preceded by an unusually extensive A + T-rich region from the stop codon of *flhE* to the *inv* initiation codon. In this 514 bp segment, the A + T content reached 68 mol% while the *Yersinia* mean A + T content is 54 mol% (Marmur & Doty, 1962). A + T contents of 65 mol% or higher are characteristic of non-coding sequence in the related species *E. coli* and *S. typhimurium* (Haughn *et al.*, 1986). Accordingly, open reading frames in both orientations were very short, mostly less than 150 bp. Moreover, the algorithm TESTCODE (Fickett, 1982) predicted non-coding regions to a 95% level of confidence for all the intervening sequence, whereas surrounding flagellar and *inv* genes scored more than 90% in coding probability.

Sequences downstream from *inv* were also analysed, showing the *flgM* operon consisting of the *flgM* and *flgN* genes. These genes encode respectively the anti- σ^{28} factor and a product required for filament assembly (Kapatral *et al.*, 1996).

The overall picture which emerged from the collected data was the singular localization of the *inv* gene within a cluster of flagellar operons (Fig. 5).

flhBAE expression is not regulated by temperature

In *Y. enterocolitica*, the transcription of several flagellar genes is downregulated at 37 °C, reaching complete extinction within a few minutes (Kapatral & Minnich, 1995; Kapatral *et al.*, 1996). The effect of temperature on *flhBAE* transcription was assessed by Northern blot

Fig. 5. Genetic organization and partial restriction map of the *inv* flanking sequences. White boxes indicate the flagellar genes while *inv* is symbolized by the hatched box. The direction of transcription of these different genes is shown by arrows. A, *Asp718I*; Ba, *BamHI*; Bg, *BglII*; E, *EcoRV*; N, *NsiI*; P, *PvuII*.

Fig. 6. Expression of *flhA* with T7 RNA polymerase. *E. coli* LK111 harbouring pGP1-2 and either, lane 1, pBC18R (the vector alone) or, lane 2, pNIV1943 (*flhA*). The positions of molecular size markers (kDa) are indicated. The arrow points to *FlhA*.

analysis. In contrast to other *Yersinia* flagellar genes, the level of *flhB* message transcribed at 24 °C did not show any significant difference with that observed at 37 °C (Fig. 4b).

Identification of the *flhA* gene product

E. coli strain LK111 carrying plasmid pNIV1943, which contained the whole *flhA* sequence, and a control strain with the pBC18R vector alone, were grown at 32 °C and then induced at 42 °C in the presence of [35 S]methionine. Total proteins were resolved by SDS-PAGE and autoradiographed. As shown in Fig. 6, a labelled protein corresponding to the product of *flhA* banded at the expected molecular mass, while no such signal was observed in the control lane.

NIVh81 mutant complementation

Complementation tests were achieved after introduction of an intact *flhA* copy in the *Y. enterocolitica* mutant. For this purpose, the complete *flhA* gene was introduced in the low-copy-number plasmid pTM100 and the

resulting construction, pNIV1945 (Fig. 1), was mobilized into the *Yersinia* mutant. When NIVh81(pNIV1945) was spotted onto swarm plates, motility was restored (Fig. 2c). In the same way, when observed by EM, NIVh81(pNIV1945) presented flagella (Fig. 3f, g). Nevertheless, in both cases, the complementation of the *flhA* mutation was only partial. On swarming plates, the complemented mutant appeared less motile than the wild-type while electron microscopic examination of the complemented mutant revealed bacteria with less and generally shorter flagella. We attributed such a distinct phenotype to the deregulated expression of *flhA*. Indeed, flagellar genes, including *flhA*, are tightly regulated. In contrast, *flhA* expression in pNIV1945 occurs under the control of the pTM100 promoter that normally triggers transcription of the chloramphenicol-resistance gene. Moreover, this plasmid exists in more than one copy per NIVh81 chromosome.

DISCUSSION

Out of the two initial working hypotheses, i.e. the presence of a pathogenicity island upstream from the *inv* locus or that of flagellar genes, the latter proved to be correct. Indeed, the full *orf236*, located upstream to *inv*, was found to share much more identity with flagellar genes from various bacteria than with virulence or cell invasion genes. This conclusion was confirmed by the observation that partial deletion of *orf236* resulted in non-motile, non-flagellated bacteria, still capable of invading mammalian cells *in vitro*. Furthermore, *orf236* clustered with two other ORFs, strikingly similar to flagellar genes from *E. coli* and *S. typhimurium*. Together, these data indicate that the three ORFs identified here correspond to the *Y. enterocolitica* flagellar genes *flhA*, *B* and *E*. Northern blot analysis showed that the *Y. enterocolitica* *flhA*, *B* and *E* genes are organized into an operon which was named *flhB*, according to the nomenclature established for *S. typhimurium* and *E. coli*. In addition to this operon, the *Y. enterocolitica* *flgM* and *flgN* were identified, mapping immediately downstream from *inv*, as described by Kapatral *et al.* (1996).

The genetics of flagellar, motility and chemotaxis functions has been thoroughly studied in *S. typhimurium*. These functions turned out to be complex processes resulting from the highly regulated expression of more than 40 flagellar genes assembled in at least 15 operons, themselves organized in a regulon (Macnab, 1992). These operons are divided into three main classes, 1, 2 and 3, with respect to their transcription hierarchy (Kutsukake *et al.*, 1990). At the top of the hierarchy, class 1 *flhC* and *flhD* genes are grouped in a single operon which is required for expression of class 2 operons. Class 2 includes genes which encode structural components of the basal body and the hook as well as the *fliA* gene (Macnab, 1992). The *fliA* gene product is the alternative σ^{28} factor that directs the transcription of flagellin genes included in the third class of hierarchy along with genes for chemotaxis and *flgM* (Arnosi &

Chamberlin, 1989). *flgM* is a regulatory locus responsible for downregulating flagellin gene transcription when one or more class 2 gene products are missing or defective (Gillen & Hughes, 1991). It encodes an anti-sigma factor which binds FliA to prevent its association with RNA polymerase core enzyme (Ohnishi *et al.*, 1992). The presence of both *fliA* and *flgM* in the *Y. enterocolitica* chromosome (Iriarte *et al.*, 1995; Kapatral *et al.*, 1996) suggests that, in this species, the regulation of class 3 flagellar genes follows the *Salmonella* paradigm. The *Salmonella* model is probably also valid for *Y. enterocolitica* class 2 promoters, given the close match of the putative *Y. enterocolitica* *flhB* promoter with the canonical *S. typhimurium* class 2 promoters. Nevertheless, unlike *S. typhimurium*, *Y. enterocolitica* is not motile at 37 °C. Kapatral *et al.* (1996) have suggested that this temperature-dependence of class 3 flagellin genes is due to the conditional expression of *fliA* itself. In *S. typhimurium*, *fliA* is a class 2 gene and accordingly, it appears preceded by the consensus class 2 promoter (Kutsukake *et al.*, 1990). However, no such sequence could be found upstream from the *Y. enterocolitica* *fliA* gene. The contrast between the temperature-insensitive *flhB* operon which owns a putative class 2 promoter and the thermoregulated *fliA* gene which lacks such a promoter, suggests that in *Y. enterocolitica*, *fliA* could be a non-class 2 gene. *fliA* could therefore appear as the knot of the thermoregulation intrigue.

Besides flagellar genes, sequence analysis revealed the presence of a probably non-coding sequence between the *flhE* and *inv* genes. This DNA stretch was longer than the 50–300 bp of non-coding sequence typically found between adjacent operons in bacterial genomes. Neither a repetitive extragenic palindromic (REP) sequence nor an intergenic repeat unit (IRU) could be detected in this intervening sequence, whereas a partial IRU is present upstream from the *inv* gene of *Y. pseudotuberculosis* (Sharples & Lloyd, 1990). The term 'grey hole' was coined by Daniels *et al.* (1992) for similar DNA regions, devoid of any identifiable function. Grey holes neighbouring flagellar genes have been reported previously by Raha *et al.* (1993), who uncovered such stretches inserted between flagellar regions IIIa and IIIb of *E. coli* and *S. typhimurium*, adjacent to the flagellar gene *fliE*.

These data contribute to a better knowledge of the genetics of flagellum functions in the genus *Yersinia*. It appears that the nucleotide sequence, the organization and the overall regulation of flagellar genes in *Y. enterocolitica* are well-conserved as compared to the situation in *E. coli* and *S. typhimurium*. Such a high level of similarity was expected in view of the close phylogenetic and ecological relationships shared by these species. More surprising was the clustering of the *flgM* and the *flhB* operons. The former operon maps at 23 and 24 min in *S. typhimurium* and *E. coli*, respectively, while the *flhB* operon maps around 40 min in *S. typhimurium* and 41–42 min in *E. coli* (Macnab, 1992). In *Y. enterocolitica*, the *flhB* operon belongs to the same cluster as the *flgM* operon, separated only by the *inv* gene. Neither

the two ORFs occurring next to the *flgN* gene in *S. typhimurium* (Kutsukake *et al.*, 1994) nor the salmonella *cheZ* gene, mapping upstream to the *flhB* operon (Stock & Stock, 1987), could be detected in the cloned locus of *Y. enterocolitica*. It should be noted that other organisms such as *Bacillus subtilis* or *Caulobacter crescentus* harbour flagellar genes homologous to those, although with markedly different organizations. The linkage of *flhB* with *flgM* probably reflects a divergent evolutionary history between yersinia and salmonella rather than relevant physiological difference. Interestingly, in *Proteus mirabilis*, the *flhB* operon clusters with the *flgN* gene, just separated by an ORF of unknown function (D. Gygi & C. Hughes, personal communication).

Although related to virulence functions, the *inv* gene of *Y. enterocolitica* does not belong to a pathogenicity island but is isolated among flagellar genes. Since horizontal transfer of virulence loci has been evoked for various bacteria (Altmeyer *et al.*, 1993; Falkow *et al.*, 1987; Knapp *et al.*, 1986; Low *et al.*, 1984; Mekalanos, 1983; So *et al.*, 1979; Waldor & Mekalanos, 1996), it was tempting to suggest that *inv* could have undergone the same fate. *inv* homologous sequences were found in other bacterial genera, such as pathogenic strains of *E. coli* (Jerse *et al.*, 1990; Yu & Kaper, 1992), *S. typhimurium* (unpublished observation), *Hafnia alvei* (Frankel *et al.*, 1994) and *Citrobacter freundii* (Schauer & Falkow, 1993), but also in a laboratory strain of *E. coli* (unpublished observation). In the latter case, the *inv*-like ORF is flanked by the insertion sequence IS3a (Umeda, 1993). Nevertheless, the presumed import of the *inv* gene and its insertion between flagellar operons was difficult to argue for. No flanking repeat or transposon-, phage- or plasmid-like elements could be found in the *inv* flanking sequences. The codon usage of *inv* was roughly equivalent to that of the surrounding flagellar genes. Moreover, the G+C content of *inv* (45.9 mol%) was identical to the mean G+C content of yersinia (46 mol%) (Marmur & Doty, 1962). Therefore, instead of being acquired by virulent *Y. enterocolitica* and *Y. pseudotuberculosis* strains, the invasion-encoding genes appeared as genuine yersinia sequences, lost or inactivated in non-enteropathogenic species of the genus. Accordingly, it should be emphasized that *Y. pestis* possesses an *inv* gene but the absence of invasive phenotype associated with this species can be explained by the inactivating insertion of an IS200-like element in the *inv* sequence (Simonet *et al.*, 1996).

ACKNOWLEDGEMENTS

We thank Philip Matsumura for critical reading of the manuscript. We are grateful to I. Delor for the gift of the *Y. enterocolitica* gene library and to M. P. Sory for providing the *inv* mutant strain MS22703. We thank D. Gygi and C. Hughes for providing unpublished information on *Proteus mirabilis* *flhA* gene. We thank also M. Iriarte, K. Hughes, T. Minamino and K. Kutsukake as well as M. Umeda and E. Ohtsubo for providing, before publication, data and sequences on *Y. enterocolitica* *fliA*, on *S. typhimurium* *flgA/M/N* and

flhB/A/E, and on *E. coli* IS3a, respectively. Jorge Sepúlveda Sánchez and Rodolfo Paredes Díaz (Unidad de microscopia, Instituto de Fisiología Celular, UNAM) provided excellent technical assistance. Finally, we thank P. Installé and the BEN (Belgian Embnet Node) for computer facilities, and I. Stainier for her great help in the RNA manipulations. A.A. was the recipient of a fellowship from ICP.

REFERENCES

- Allaoui, A., Woestyn, S., Sluiters, C. & Cornelis, G. R. (1994). YscU, a *Yersinia enterocolitica* inner membrane protein involved in Yop secretion. *J Bacteriol* 176, 4534–4542.
- Altmeyer, R. M., McNern, J. K., Bossio, J. C., Rosenshine, I., Brett Finlay, B. & Gálan, J. E. (1993). Cloning and molecular characterization of a gene involved in *Salmonella* adherence and invasion of cultured epithelial cells. *Mol Microbiol* 7, 89–98.
- Andrews, G. P. & Maurelli, A. T. (1992). *mxiA* of *Shigella flexneri* 2a, which facilitates export of invasion plasmid antigens, encodes a homolog of the low-calcium-response protein, LcrD, of *Yersinia pestis*. *Infect Immun* 60, 3287–3295.
- Arnosti, D. N. & Chamberlin, M. J. (1989). Secondary σ factor controls transcription of flagellar and chemotaxis genes in *Escherichia coli*. *Proc Natl Acad Sci USA* 86, 830–834.
- Bartlett, D. H., Frantz, B. B. & Matsumura, P. (1988). Flagellar transcriptional activators FlhB and FlaI: gene sequences and 5' consensus sequences of operons under FlhB and FlaI control. *J Bacteriol* 170, 1575–1581.
- Boyer, H. W. & Roulland-Dussoix, D. (1969). A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J Mol Biol* 41, 459–472.
- Carpenter, P. B. & Ordal, G. W. (1993). *Bacillus subtilis* FlhA: a flagellar protein related to a new family of signal-transducing receptors. *Mol Microbiol* 7, 735–743.
- China, B., Michiels, T. & Cornelis, G. R. (1990). The pYV plasmid of *Yersinia* encodes a lipoprotein, YlpA, related to TraT. *Mol Microbiol* 4, 1585–1593.
- Cornelis, G. R. (1994). *Yersinia* pathogenicity factors. *Curr Top Microbiol Immunol* 192, 243–263.
- Daniels, D. L., Plunkett, G., III, Burland, V. & Blattner, F. R. (1992). Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. *Science* 257, 771–778.
- Delor, I., Kaeckenbeeck, A., Wauters, G. & Cornelis, G. R. (1990). Nucleotide sequence of *yst*, the *Yersinia enterocolitica* gene encoding the heat-stable enterotoxin, and prevalence of the gene among pathogenic and nonpathogenic yersiniae. *Infect Immun* 58, 2983–2988.
- Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12, 387–395.
- Falkow, S., Small, P., Isberg, R., Hayes, S. F. & Corwin, D. (1987). A molecular strategy for the study of bacterial invasion. *Rev Infect Dis* 9, S450–S455.
- Fenselau, S., Balbo, I. & Bonas, U. (1992). Determinants of pathogenicity in *Xanthomonas campestris* pv. *vesicatoria* are related to proteins involved in secretion in bacterial pathogens of animals. *Mol Plant-Microbe Interact* 5, 390–396.
- Fickett, J. W. (1982). Recognition of protein coding regions in DNA sequences. *Nucleic Acids Res* 10, 5303–5318.
- Frankel, G., Candy, D. C., Everest, P. & Dougan, G. (1994). Characterization of the C-terminal domains of intimin-like proteins of enteropathogenic and enterohemorrhagic *Escherichia*

- coli*, *Citrobacter freundii*, and *Hafnia alvei*. *Infect Immun* 62, 1835–1842.
- Galán, J. E., Ginocchio, C. & Costeas, P. (1992). Molecular and functional characterization of the *Salmonella* invasion gene *invA*: homology of *InvA* to members of a new protein family. *J Bacteriol* 174, 4338–4349.
- Gillen, K. L. & Hughes, K. T. (1991). Negative regulatory loci coupling flagellin synthesis to flagellar assembly in *Salmonella typhimurium*. *J Bacteriol* 173, 2301–2310.
- Gough, C. L., Genin, S., Lopes, V. & Boucher, C. A. (1993). Homology between the HrpO protein of *Pseudomonas solanacearum* and bacterial proteins implicated in a signal peptide-independent secretion mechanism. *Mol Gen Genet* 239, 378–392.
- Groisman, E. A. & Ochman, H. (1993). Cognate gene clusters govern invasion of host epithelial cells by *Salmonella typhimurium* and *Shigella flexneri*. *EMBO J* 12, 3779–3787.
- Gygi, D., Bailey, M. J., Allison, C. & Hughes, C. (1995). Requirement for FlhA in flagella assembly and swarm-cell differentiation by *Proteus mirabilis*. *Mol Microbiol* 15, 761–769.
- Haughn, G. W., Wessler, S. R., Gemmill, R. M. & Calvo, J. M. (1986). High A+T content conserved in DNA sequences upstream of *leuABCD* in *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol* 166, 1113–1117.
- Huang, H.-C., Xiao, Y., Lin, R.-H., Lu, Y., Hutcheson, S. W. & Collmer, A. (1993). Characterization of the *Pseudomonas syringae* pv. *syringae* 61 *hrpJ* and *hrpI* genes: homology of *HrpI* to a superfamily of proteins associated with protein translocation. *Mol Plant-Microbe Interact* 6, 515–520.
- Iino, T., Komeda, Y., Kutsukake, K., Macnab, R. M., Matsumura, P., Parkinson, J. S., Simon, M. I. & Yamaguchi, S. (1988). New unified nomenclature for the flagellar genes of *Escherichia coli* and *Salmonella typhimurium*. *Microbiol Rev* 52, 533–535.
- Iriarte, M., Stainier, I., Mikulskis, A. & Cornelis, G. R. (1995). *fliA* encoding sigma 28 in *Yersinia enterocolitica*. *J Bacteriol* 177, 2299–2304.
- Isberg, R. R. & Falkow, S. (1985). A single genetic locus encoded by *Yersinia pseudotuberculosis* permits invasion of cultured animal cells by *Escherichia coli* K-12. *Nature* 317, 262–264.
- Isberg, R. R., Voorhis, D. L. & Falkow, S. (1987). Identification of invasins: a protein that allows enteric bacteria to penetrate cultured mammalian cells. *Cell* 50, 769–778.
- Jarvis, K. G., Giron, J. A., Jerse, A. E., McDaniel, T. K., Donnenberg, M. S. & Kaper, J. B. (1995). Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proc Natl Acad Sci USA* 92, 7996–8000.
- Jerse, A. E., Yu, J., Tall, B. D. & Kaper, J. B. (1990). A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proc Natl Acad Sci USA* 87, 7839–7843.
- Kaniga, K., Delor, I. & Cornelis, G. R. (1991). A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria: inactivation of the *blaA* gene of *Yersinia enterocolitica*. *Gene* 109, 137–141.
- Kapatral, V. & Minnich, S. A. (1995). Co-ordinate, temperature-sensitive regulation of the three *Yersinia enterocolitica* flagellin genes. *Mol Microbiol* 17, 49–56.
- Kapatral, V., Olson, J. W., Pepe, J. C., Miller, V. L. & Minnich, S. A. (1996). Temperature-dependent regulation of *Yersinia enterocolitica* Class III flagellar genes. *Mol Microbiol* 19, 1061–1071.
- Kim, M. C. (1989). Molecular characterization and regulation of the gene expression of *flhB* operon in *Escherichia coli*. PhD thesis, University of Illinois.
- Knapp, S., Hacker, J., Jarchau, T. & Goebel, W. (1986). Large, unstable inserts in the chromosome affect virulence properties of uropathogenic *Escherichia coli* O6 strain 536. *J Bacteriol* 168, 22–30.
- Kolter, R., Inuzuka, M. & Helinski, D. R. (1978). Trans-complementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. *Cell* 15, 1199–1208.
- Kutsukake, K., Ohya, Y. & Iino, T. (1990). Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. *J Bacteriol* 172, 741–747.
- Kutsukake, K., Okada, T., Yokoseki, T. & Iino, T. (1994). Sequence analysis of the *flgA* gene and its adjacent region in *Salmonella typhimurium*, and identification of another flagellar gene, *flgN*. *Gene* 143, 49–54.
- Lambert de Rouvroit, C., Sluiter, C. & Cornelis, G. R. (1992). Role of the transcriptional activator, VirF, and temperature in the expression of the pYV plasmid genes of *Yersinia enterocolitica*. *Mol Microbiol* 6, 395–409.
- Low, D., David, V., Lark, D., Schoolnik, G. & Falkow, S. (1984). Gene clusters governing the production of hemolysin and mannose-resistant hemagglutination are closely linked in *Escherichia coli* serotype O4 and O6 isolates from urinary tract infections. *Infect Immun* 43, 353–358.
- McDaniel, T. K., Jarvis, K. G., Donnenberg, M. S. & Kaper, J. B. (1995). A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci USA* 92, 1664–1668.
- Macnab, R. M. (1976). Examination of bacterial flagellation by dark-field microscopy. *J Clin Microbiol* 4, 258–265.
- Macnab, R. M. (1992). Genetics and biogenesis of bacterial flagella. *Annu Rev Genet* 26, 131–158.
- Marmur, J. & Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J Mol Biol* 5, 109–118.
- Mekalanos, J. J. (1983). Duplication and amplification of toxin genes in *Vibrio cholerae*. *Cell* 35, 253–263.
- Ménard, R., Sansonetti, P. J. & Parsot, C. (1993). Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *J Bacteriol* 175, 5899–5906.
- Michiels, T. & Cornelis, G. R. (1991). Secretion of hybrid proteins by the *Yersinia* Yop export system. *J Bacteriol* 173, 1677–1685.
- Michiels, T., Wattiau, P., Brasseur, R., Ruyschaert, J.-M. & Cornelis, G. (1990). Secretion of Yop proteins by *Yersinia*. *Infect Immun* 58, 2840–2849.
- Miller, S., Pesci, E. C. & Pickett, C. L. (1993). A *Campylobacter jejuni* homolog of the LcrD/FlbF family of proteins is necessary for flagellar biogenesis. *Infect Immun* 61, 2930–2936.
- Minamino, T., Iino, T. & Kutsukake, K. (1994). Molecular characterization of the *Salmonella typhimurium flhB* operon and its protein products. *J Bacteriol* 176, 7630–7637.
- Moon, H. W., Whipp, S. C., Argenzio, R. A., Levine, M. M. & Giannella, R. A. (1983). Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infect Immun* 41, 1340–1351.
- Norlander, J., Kempe, T. & Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* 26, 101–106.
- Ohnishi, K., Kutsukake, K., Suzuki, H. & Iino, T. (1992). A novel

transcriptional regulation mechanism in the flagellar regulon of *Salmonella typhimurium*: an anti-sigma factor inhibits the activity of the flagellum-specific sigma factor, σ^F . *Mol Microbiol* **6**, 3149–3157.

Pepe, J. C. & Miller, V. L. (1993). *Yersinia enterocolitica* invasin: a primary role in the initiation of infection. *Proc Natl Acad Sci USA* **90**, 6473–6477.

Plano, G. V., Barve, S. S. & Straley, S. C. (1991). LcrD, a membrane-bound regulator of the *Yersinia pestis* low-calcium response. *J Bacteriol* **173**, 7293–7303.

Portnoy, D. A., Moseley, S. L. & Falkow, S. (1981). Characterization of plasmids and plasmid-associated determinants of *Yersinia enterocolitica* pathogenesis. *Infect Immun* **31**, 775–782.

Raha, M., Kihara, M., Kawagishi, I. & Macnab, R. M. (1993). Organization of the *Escherichia coli* and *Salmonella typhimurium* chromosomes between flagellar regions IIIa and IIIb, including a large non-coding region. *J Gen Microbiol* **139**, 1401–1407.

Ramakrishnan, G., Zhao, J.-L. & Newton, A. (1991). The cell cycle-regulated flagellar gene *flbF* of *Caulobacter crescentus* is homologous to a virulence locus (*lcrD*) of *Yersinia pestis*. *J Bacteriol* **173**, 7283–7292.

Salmond, G. P. C. & Reeves, P. J. (1993). Membrane traffic wardens and protein secretion in Gram-negative bacteria. *Trends Biochem Sci* **18**, 7–12.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.

Sasakawa, C., Komatsu, K., Tobe, T., Suzuki, T. & Yoshikawa, M. (1993). Eight genes in region 5 that form an operon are essential for invasion of epithelial cells by *Shigella flexneri* 2a. *J Bacteriol* **175**, 2334–2346.

Schauer, D. B. & Falkow, S. (1993). Attaching and effacing locus of a *Citrobacter freundii* biotype that causes transmissible murine colonic hyperplasia. *Infect Immun* **61**, 2486–2492.

Sharples, G. J. & Lloyd, R. G. (1990). A novel repeated DNA sequence located in the intergenic regions of bacterial chromosomes. *Nucleic Acids Res* **18**, 6503–6508.

Simon, R., Priefer, U. & Puhler, A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/Technology* **1**, 784–791.

Simonet, M., Riot, B., Fortineau, N. & Berche, P. (1996). Invasin production by *Yersinia pestis* is abolished by insertion of an IS200-like element within the *inv* gene. *Infect Immun* **64**, 375–379.

So, M., Heffron, F. & McCarthy, B. J. (1979). The *E. coli* gene encoding heat stable toxin is a bacterial transposon flanked by inverted repeats of IS1. *Nature* **277**, 453–456.

Sory, M.-P. & Cornelis, G. R. (1994). Translocation of a hybrid YopE-adenylate cyclase from *Yersinia enterocolitica* into HeLa cells. *Mol Microbiol* **14**, 583–594.

Staskawicz, B., Dahlbeck, D., Keen, N. & Napoli, C. (1987). Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. *J Bacteriol* **169**, 5789–5794.

Stock, A. M. & Stock, J. B. (1987). Purification and characterization of the CheZ protein of bacterial chemotaxis. *J Bacteriol* **169**, 3301–3311.

Tabor, S. & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* **82**, 1074–1078.

Umeda, M. (1993). *Genomic rearrangements caused by repetitive DNA sequences in bacteria and rice plant*. PhD thesis, University of Tokyo.

Waldor, M. K. & Mekalanos, J. J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**, 1910–1914.

Wei, Z.-M. & Beer, S. V. (1993). HrpI of *Erwinia amylovora* functions in secretion of harpin and is a member of a new protein family. *J Bacteriol* **175**, 7958–7967.

Yu, J. & Kaper, J. B. (1992). Cloning and characterization of the *eae* gene of enterohaemorrhagic *Escherichia coli* O157:H7. *Mol Microbiol* **6**, 411–417.

Received 4 April 1997; revised 30 June 1997; accepted 9 July 1997.