## Identification of a novel type IV pilus gene cluster required for gastrointestinal colonization of Citrobacter rodentium

Molecular Microbiology (2003) **48**(3), 795–809

# Identification of a novel type IV pilus gene cluster required for gastrointestinal colonization of *Citrobacter rodentium*

**Rosanna Mundy, Derek Pickard, Rebecca K. Wilson, Cameron P. Simmons, Gordon Dougan and Gad Frankel***

*Centre for Molecular Microbiology and Infection, Department of Biological Sciences Flowers Building, Imperial College London, London SW7 2AZ, UK.*

## Summary

***Citrobacter rodentium*** **is used as an** ***in vivo*** **model system for clinically significant enteric pathogens such as enterohaemorrhagic** ***Escherichia coli*** **(EHEC) and enteropathogenic** ***E. coli*** **(EPEC). These pathogens all colonize the lumen side of the host gastrointestinal tract via attaching and effacing (A/E) lesion formation. In order to identify genes required for the colonization of A/E-forming pathogens, a library of signature-tagged transposon mutants of** ***C. rodentium*** **was constructed and screened in mice. Of the 576 mutants tested, 14 were attenuated in their ability to colonize the descending colon. Of these, eight mapped to the locus of enterocyte effacement (LEE), which is required for the formation of A/E lesions, underlying the importance of this mechanism for pathogenesis. Another mutant, P5H2, was found to have a transposon insertion in an open reading frame that has strong similarity to type IV pilus nucleotide-binding proteins. The region flanking the transposon insertion was sequenced, identifying a cluster of 12 genes that encode the first described pilus of** ***C. rodentium*** **(named colonization factor** ***Citrobacter*****, CFC). The proteins encoded by** ***cfc*** **genes have identity to proteins of the type IV COF pilus of enterotoxigenic** ***E. coli*** **(ETEC), the toxin co-regulated pilus of** ***Vibrio cholerae*** **and the bundle-forming pilus of EPEC. A non-polar mutation in** ***cfcI*****, complementation of this strain with wild-type** ***cfcI*** **and complementation of strain P5H2 with wild-type** ***cfcH*** **confirmed that these genes are required for colonization of the gastrointestinal tract by** ***C. rodentium*****. Thus, CFC provides a convenient model to study type IV pilus-mediated pathogen–host interactions under physiological conditions in the natural colonic environment.**

Accepted 20 January, 2003. *For correspondence. E-mail g.frankel@ic.ac.uk; Tel. (+44) 20 7594 5253; Fax (+44) 7594 3069.



## Introduction

*Citrobacter rodentium* belongs to a family of human and animal enteric pathogens that use attaching and effacing (A/E) lesions to colonize the host gastrointestinal tract. A/E lesions are characterized by localized destruction of brush border microvilli and intimate attachment of the bacteria to the plasma membrane of the host epithelial cells (reviewed by Frankel *et al*., 1998). Enteropathogenic (EPEC) and enterohaemorrhagic (EHEC) *Escherichia coli* are clinically important A/E pathogens for humans. EPEC is a significant cause of infantile diarrhoea in developing countries (Nataro and Kaper, 1998), whereas EHEC is more of a concern in the developed world, with around 73 000 cases reported annually in the USA (Mead *et al*., 1999). In addition to causing diarrhoea, EHEC infection can result in the life-threatening complications of haemorrhagic colitis and haemolytic uraemic syndrome (HUS) resulting from the production of verocytotoxin (Nataro and Kaper, 1998).

A/E lesions were first described in EPEC strains (Moon *et al*., 1983), and similar lesions have also now been reported in EHEC (Tzipori *et al*., 1995; Phillips *et al*., 2000) and *C. rodentium* (Schauer and Falkow, 1993a). A/E lesion formation has been shown to be required for pathogenicity in all animal models tested (Donnenberg *et al*., 1993; Schauer and Falkow, 1993b; McKee *et al*., 1995; Dean-Nystrom *et al*., 1998; Marches *et al*., 2000). The genes encoding the A/E phenotype map to a pathogenicity island termed the locus of enterocyte effacement (LEE) (McDaniel *et al*., 1995; Perna *et al*., 1998; Deng *et al*., 2001). This pathogenicity island encodes an adhesion molecule called intimin (*eae*) that mediates intimate attachment to epithelial cells through binding to a second LEE-encoded protein, Tir, which is delivered to the host cell membrane by a type III secretion system (reviewed by Frankel *et al*., 1998; Vallance and Finlay, 2000). In addition, the LEE type III secretion system also translocates other effector proteins (Esps) into host cells, which subvert host signal transduction and cause cytoskeletal rearrangements that are necessary for the formation of

A/E lesions (Frankel *et al*., 1998; Vallance and Finlay, 2000).

It is highly likely that colonization factors distinct from the LEE also play a role in A/E pathogens. Putative virulence factors of EHEC include a pilus (Ashkenazi *et al*., 1992), an intimin-independent adhesin (Sherman *et al*., 1991) and a *Vibrio cholerae* iron-regulated gene A homologue (IrgA adhesin) that conferred cell adhesion to *E. coli* K-12 (Tarr *et al*., 2000). Recently, a large toxin has been identified in both EHEC (*efa1*) and EPEC (*lifA*) which may play a role in adhesion and/or in the inhibition of lymphocyte activation (Klapproth *et al*., 2000; Nicholls *et al*., 2000). Significantly, it was shown recently that Efa1 of EHEC serogroups O5 and O111 plays a direct role in colonization using the calf infection model *in vivo* (Stevens *et al*., 2002).

Neither EPEC nor EHEC is able to colonize mice; in fact, the only known A/E pathogen to do so is *C. rodentium*. This pathogen causes transmissible colonic hyperplasia in mice (Barthold *et al*., 1976; reviewed by Luperchio and Schauer, 2001), which results in weight loss, soft stools and enlargement of the descending colon (through hyperplasia), sometimes resulting in rectal prolapse. *C. rodentium* is therefore an ideal small animal model in which to study luminal microbial pathogens relying on A/E formation for colonization of the host.

The aim of this study was to identify novel virulence genes in *C. rodentium*, using signature-tagged mutagenesis (STM), in order to provide new insights into the pathogenesis of A/E bacteria. STM enables large numbers of mutants to be screened in a small number of animals and the rapid identification of those with attenuated virulence. Mutants with transposon insertions in genes required for colonization fail to establish sustained infection in a host and are identified by negative selection (Hensel *et al*., 1995). This method was first used in the enteric pathogen *Salmonella typhimurium*, where it led to the identification of a novel type III secretion system required for systemic pathogenesis (Shea *et al*., 1996). The STM screen of *C. rodentium* mutants presented in this study has resulted in the identification of a novel gene cluster encoding a type IV pilus that is necessary for colonization of mice.

## Results

### *Construction and characterization of the mutant library*

Forty-eight preselected signature-tagged mini-Tn*5*Km2 transposons in plasmid pUT, which harboured tags giving consistent and specific signals, were used to construct the mutant library. Forty-seven of the preselected plasmids were conjugated into the *C. rodentium* recipient DBS100 nal$^r$. Examination of representative exconjugants revealed that over 98% were genuine, random, single-transposon insertion mutants (data not shown). The 48th signature tag (H6) was conjugated into the avirulent *C. rodentium* intimin mutant strain DBS255. As the recipient strain was already nal$^r$ and kan$^r$, exconjugants were checked for the presence of the signature tag by colony polymerase chain reaction (PCR) using tag primers P2 and P4 (Hensel *et al*., 1995). Approximately 9% contained the tag, and one such colony was selected (strain ICC174) and placed into each pool as a positive control for attenuation. A total of 2400 nal$^r$, kan$^r$, pip$^s$ mutants were assembled into 50 pools each containing 48 mutants.

### *Screening the library for mutants unable to colonize the colon*

With the *C. rodentium* infection model used, an inoculum of $>10^8$ colony-forming units (cfu) per mouse by oral gavage is required in order to ensure that 100% of mice become colonized. In preliminary experiments, a pool of 48 *C. rodentium* mutants was inoculated into each of three mice by oral gavage at a dose of $10^{10}$ cfu (the maximum dose that can be used in this model). All 48 mutants were found to be present in the input pool and in stool samples taken 6 h after inoculation. However, only 17–22 of the original 48 mutants were present in stool samples taken 2 days after inoculation, and even more were missing from day 5 colon samples taken from each mouse (data not shown). The same pool of 48 mutants was then split into two groups of 24 strains and given to mice at a dose of $10^{10}$ cfu. The majority of the 24 mutants were reproducibly present in output pools recovered from mice stools at days 2 and 3 and from day 5 colons. Therefore, the mutants were screened in pools of 24–28 in each of two mice. Figure 1 shows an example of the hybridization results; the mutants in positions H2, H4 and H6 give good signals in the input blot (Fig. 1A) but are absent from the day 3 output blot (Fig. 1B). H6 is the signature-tagged intimin mutant ICC174; this strain was not recovered after day 2 in any of the pools screened.

Mutants with strong hybridization signals in the input

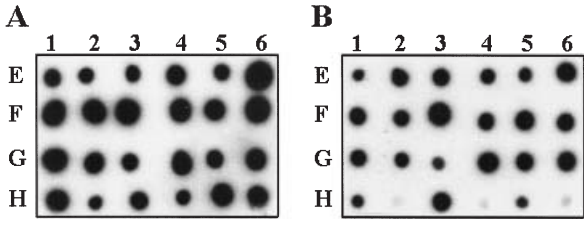


**Fig. 1.** Representative hybridization results from the STM screen for colonization-defective mutants. Plasmid dot blots were probed with radiolabelled tags amplified from bacteria in the inoculum (A) and from bacteria recovered from stools of mice 3 days after inoculation (B). The mutants in position H2, H4 and H6 were present in the inoculum but failed to be isolated from infected mice.

**Table 1.** Transposon insertion sites in colonization-defective mutants.

| Strain | Disrupted gene or homologue | Amino acid identity/ organism[a] | Function | *In vivo* CI[b] |
|---|---|---|---|---|
| LEE mutants | | | | |
| P1H4 | *espD* | 100%, Cr | Type III secretion | 0 |
| P3D4, P8C6 | *tir* | 100%, Cr | Receptor for intimin | 0.00007 |
| P4H2, P8A2 | *eae* | 100%, Cr | Adhesin | 0.00085 |
| P5A3 | orf4 | 100%, Cr | Type III secretion | 0 |
| P6C6 | 24 bp upstream of *map* | 100%, Cr | Type III secretion | 0.055 |
| P10G3 | *escD* | 100%, Cr | Type III secretion | 0.038 |
| Pilus mutant | | | | |
| P5H2 | *cofH* | 62%, ETEC | Assembly of pilus | 0.011 |
| Metabolic | | | | |
| P10D1 | *pepN* | 88%, Ec<br>83%, EHEC | Glutathione utilization | 0.24 |
| Function unknown | | | | |
| P3G5 | Z0656 | 90%, EHEC<br>90%, Ec | Hypothetical protein | 0.0065 |
| P3F1 | | No homology | | 0.092 |
| P4C4 | | No homology | | 0.018 |
| P5H5 | | No homology | | 0.072 |

**a.** The homologue with the highest sequence identity is shown. Cr, *C. rodentium*; Ec, *E. coli* K-12; ETEC, enterotoxigenic *E. coli*, EHEC, enterohaemorrhagic *E. coli*.
**b.** *In vivo* competitive indexes (CIs) are the average of results from two or three animals.

pool but weak or no hybridization signals in day 3 stools and day 5 colons were rescreened in the next pool of 24, increasing the pool size to a maximum of 28 mutants. A total of 576 mutants were screened in this way and, of these, 33 gave consistently negative hybridization signals in both rounds of screening. The colonization potential of each of these mutants was established using mixed infection experiments, with inoculum containing ≈ 50% wild-type *C. rodentium* and 50% mutant. The exact ratio of mutant to wild-type cfu was calculated for both the inocula and the stool contents at day 3. The competitive index (CI) was defined as the output ratio (mutant/wild-type cfu) divided by the input ratio (mutant/wild-type cfu). Only 14 mutants with a significant colonization defect compared with the wild type (CI < 1) were characterized further (Table 1).

### *Characterization of the attenuated strains identified by STM*

Southern hybridization analysis of the 14 mutants determined that each strain harboured only one transposon insertion in the genome (data not shown). DNA flanking each mini-Tn*5*Km2 insertion was cloned, sequenced and used to search the GenBank database for homologous genes or proteins using programs BLASTN and BLASTX (http://www.ncbi.nih/gov). The results are shown in Table 1. The mutations fall into five categories with transposon insertions in genes encoding type III secretion, the adhesin intimin, enzymes in metabolic pathways, proteins with pilus assembly function and proteins of unknown function. Around half (8/14) the attenuated mutants have transposon insertions into the LEE of *C. rodentium*. Of the other six strains, one has an insertion into a homologue of a gene in enterotoxigenic *E. coli* (ETEC; P5H2), and two have insertions into putative genes found in *E. coli* K-12 as well as in EHEC 0157 (P3G5, P10D1). For the other three mutants (P3F1, P4C4 and P5H5), no gene or protein was found with significant similarity to sequences flanking the transposon insertion sites.

To determine the level of virulence attenuation of the *C. rodentium* mutants, single infection experiments were carried out. Mice were challenged orally with $\approx 2 \times 10^8$ cfu of wild type, DBS255 or signature-tagged mutants. Thirteen days after inoculation, measurement of pathogen burden in the colons of these mice revealed marked differences in the ability of the strains to colonize the colonic epithelium and cause hyperplasia (see Table 2). To confirm whether the reduction in colonization for any of these 14 mutants was caused by a general growth defect, the growth curves of these strains were compared with those of the wild type in rich media (LB) and minimal media (VB). The analysis of sequence characteristics, growth and virulence of each of these mutants is summarized below.

### *Insertion into LEE genes*

The majority of the transposon insertions were found in genes located within the *C. rodentium* LEE (full sequence and analysis of which is found in Deng *et al*., 2001). The identification of a number of LEE genes in this study

**Table 2.** Virulence of signature-tagged *C. rodentium* strains in mice.[a]

| Bacterial strain | No. of mice | Average colon weight[b] (g ± SE) | Bacterial colonization[c] (cfu/colon ± SE) |
|---|---|---|---|
| Uninfected | 4 | 0.166 ± 0.002 | 0 |
| Wild-type DBS100 | 12 | 0.327 ± 0.011** | $(1.76 \pm 0.57) \times 10^8$ |
| DBS255 (Δ*eae*) | 4 | 0.167 ± 0.003 | 0 |
| P1H4 (Tn::*espD*) | 4 | 0.160 ± 0.004 | 0 |
| P10G3 (Tn::*escD*) | 4 | 0.162 ± 0.004 | 0 |
| P6C6 (Tn::*map*) | 3 | 0.190 ± 0.015 | $(5.0 \pm 3.1) \times 10^7$ |
| P5H2 (Tn::*cofH*) | 4 | 0.162 ± 0.005 | 0 |
| P5H5 | 4 | 0.174 ± 0.004 | 0 |
| P3F1 | 4 | 0.249 ± 0.027* | $(2.01 \pm 1.09) \times 10^8$ |
| P3G5 | 4 | 0.305 ± 0.033** | $(1.56 \pm 1.07) \times 10^8$ |
| P4C4 | 4 | 0.233 ± 0.017* | $(1.11 \pm 0.63) \times 10^8$ |

**a.** C3H/Hej mice were infected with various *C. rodentium* strains and the colons collected at day 13 after inoculation.
**b.** SE, standard error. 8 cm of distal colon was weighed after removal of stool pellets. The colons removed from mice infected with strains DBS255, P1H4, P10G3, P6C6, P5H2 and P5H5 are all statistically in the same weight range as the group of uninfected mice ($P < 0.001$). Significantly increased colon weight compared with uninfected mice is indicated by $^{**}P < 0.01$ and $^{*}P < 0.05$ as analysed by one-way analysis of variance (ANOVA) test.
**c.** cfu, colony-forming units. The lower limit of detection is 50 cfu/colon.

underlines the importance of the A/E mechanism in the colonization and survival of *C. rodentium* in the mouse gastrointestinal tract. Two independent insertions were found in the genes encoding the adhesin intimin (P4H2 and P8A2) and the translocated intimin receptor Tir (P3D4 and P8C6). In both cases, the insertions arose from independent matings with different tags, and the insertions are in different locations within each gene. The role of intimin-mediated intimate attachment has been demonstrated previously in a number of infection models, including EHEC in calves (Dean-Nystrom *et al*., 1998) and pigs (Donnenberg *et al*., 1993; McKee *et al*., 1995), rabbit (R)EPEC in rabbits (Marches *et al*., 2000) and *C. rodentium* in mice (Schauer and Falkow, 1993b). The isolation of several signature-tagged strains with mutations in the known virulence factor intimin therefore validates the STM screening of *C. rodentium* mutants in this mouse model.

Intimin has been shown to attach intimately to epithelial cells through binding to Tir (Kenny *et al*., 1997; Deibel *et al*., 1998). Tir has been shown to be necessary for the A/E lesion formation of EPEC, EHEC and REPEC in cell culture (Kenny *et al*., 1997; Deibel *et al*., 1998; DeVinney *et al*., 1999; Marches *et al*., 2000) and for pathogenesis in the REPEC model of infection (Marches *et al*., 2000). The identification in this study of two strains with independent transposon insertions in Tir suggests that this protein is also required for *C. rodentium* colonization of the mouse colon. However, as *tir* is the first gene in an operon comprising *tir*, *cesT* and *eae* (Deng *et al*., 2001), polar effects on downstream genes cannot be ruled out as the cause of failure to colonize in strains P3D4 and P8C6.

Mutant P5A3 has a mutation in *orf4* of *C. rodentium*, which does not yet have an ascribed function, whereas strain P1H4 has a mutation in *espD*, a translocator protein and a major component of the putative type III secretion translocation pore (Wachter *et al*., 1999). When P1H4 was inoculated into mice as a single infection, it was unable to colonize and failed to induce hyperplasia as measured by colon weight (Table 2). These results are consistent with those reported in mice infected with an *espB* mutant strain (Newman *et al*., 1999). However, as both *orf4* and e*spD* genes are part of the *LEE1* and *LEE4* operons, respectively, we cannot rule out the possibility that the transposon insertions in these strains might have polar effects on the expression of downstream genes.

The final two LEE mutants identified in this study, P6C6 and P10G3, both contain mutations affecting monocistronic genes. P6C6 contains an insertion at position 21606 of the *C. rodentium* LEE sequence (accession number AF311901 in GenBank; Deng *et al*., 2001). As the effector protein Map is encoded by nucleotides 21630–22241, the transposon has inserted 24 bp upstream of the start codon. In EPEC, the transcriptional start site of *map* was identified as 133 bp upstream of the translational start codon, and analysis of the region 5′ of this revealed putative promoter sequences (at −35 and −10; Sanchez-SanMartin *et al*., 2001). The sequence upstream of *map* in *C. rodentium* has 85% identity at the nucleotide level to the same region in EPEC; therefore, it is likely that the transposon insertion interferes with the expression of this gene. The precise function of Map is not yet known, but it appears, after translocation, to interfere with the maintenance of membrane potential in mitochondria (Kenny and Jepson, 2000) and Cdc42-mediated signal transduction pathways (Kenny *et al*., 2002).

Strain P6C6 was tested in a single infection to determine whether it had a typical LEE mutant phenotype in mice. However, P6C6 was able to colonize the colon as effectively as wild type, with comparable numbers of bac-

teria recovered from the colonic tissue (Table 2) and from stool samples taken at days 1, 2, 5, 7 and 9 after inoculation (data not shown). These results suggest that Map has a *cis* activity that contributes directly to competitiveness. However, the mutation did appear to affect the level of hyperplasia (as measured by colon weight), as mice infected with this strain had colon weights not statistically different from those of uninfected mice (Table 2).

Strain P10G3 contains a mutation 191 bp downstream of the start codon of *escD*. This strain was highly attenuated when put into mice in a single infection, with a complete inability to colonize the colon or cause hyperplasia, as measured by colon weight (Table 2). EscD was shown to be part of the type III secretion apparatus (Kresse *et al*., 1998), so bacteria lacking EscD are deficient in protein translocation and A/E lesion formation.

### *Non-LEE insertions*

Strain P10D1 contains a mutation in a gene with high identity (83–88% identity at the nucleotide level) to *pepN* of *E. coli* K-12 and EHEC 0157:H7 (Blattner *et al*., 1997; Makino *et al*., 1999; Perna *et al*., 2001) (Table 1). Strain P3G5 has an insertion into an open reading frame (ORF) with 90% identity to ORF Z0656 of EHEC 0157 (Perna *et al*., 2001). This strain was outcompeted by the wild type in mixed infection experiments, with a competitive index of 0.0065, explaining why it was identified in the STM screen in the first place. It had a growth curve in LB and VB comparable with that of wild-type *C. rodentium* (data not shown). However, when put into mice as a single infection, strain P3G5 had no colonization defect, with a colon weight and number of cfu recovered from the colon statistically in the same range as the wild-type strain (Table 2).

Of the three strains with a transposon insertion into DNA with no similarity to sequences in the databases, two strains (P3F1 and P4C4) have no colonization defect when tested in single infection (Table 2). However, they both have colon weights that are statistically heavier than the uninfected/avirulent groups ($P < 0.05$) but lighter than the wild-type groups (P3F1, $P < 0.05$; P4C4, $P < 0.001$). So these mutations may have some effect on the ability of *C. rodentium* to cause hyperplasia in the mouse gut. In contrast, strain P5H5 is completely colonization and hyperplasia deficient when put into mice as a single infection (Table 2). However, there is as yet no clue as to what the transposon in this strain has inserted into to affect virulence so severely.

### *Identification of a novel type IV pilus biogenesis gene in* C. rodentium

Of the mutations identified by STM that did not map to the LEE region, P5H2 is the most interesting. This mutant strain has a transposon insertion in an ORF with similarity to *cofH*, which encodes a nucleotide-binding protein required for assembly of the COF pilus in ETEC (Taniguchi *et al*., 2001). There is 62% identity at the amino acid level between the ORF identified in strain P5H2 and that of CofH. There is also similarity to TcpT of the toxin co-regulated pilus (TCP) of *V. cholerae* (44% identity) and BfpD of the bundle-forming pilus (BFP) of EPEC (29% identity) (Fig. 2). These proteins all have a number of motifs in common, including conserved Walker boxes A and B, Asp boxes and two pairs of cysteine residues (CXXC motifs) all associated with a nucleotide-binding activity (Walker *et al*., 1982) (Fig. 2). It has been speculated that these nucleotide-binding proteins may be involved in the production of energy required for the biogenesis and secretion of type IV pili (Sohel *et al*., 1996; Taniguchi *et al*., 2001).

Strain P5H2 is completely unable to colonize the mouse colon when put in as a single infection, with no bacteria recoverable from the colon at day 13 after inoculation (Table 2) or from stools 3 days after inoculation (data not shown). However, this strain is able to grow as well as the wild-type strain in LB and VB (data not shown), so does not have an observable growth defect in culture.

### *The type IV pilus gene is part of a large gene cluster*

The original *Pst*I fragment cloned from mutant P5H2 contained only 844 bp of *C. rodentium* DNA flanking the left-hand side of the transposon insertion site. So larger regions of DNA flanking the transposon were cloned and sequenced to determine whether further type IV pilus genes were present in this region of *C. rodentium*. Southern hybridization analysis revealed two restriction enzymes that could be used to clone DNA containing the transposon insertion: *Kpn*I (≈ 7.3 kb) and *Nde*I (≈ 6.2 kb) (data not shown). *Kpn*I and *Nde*I transposon-containing fragments were cloned into pUC18 restricted with the appropriate enzyme. The *Kpn*I fragment was sequenced by primer walking to produce 5.6 kb of *C. rodentium* DNA flanking the right-hand side of the transposon. The *Nde*I fragment contained 0.9 kb of DNA flanking the right-hand side of the transposon, leaving 3.5 kb of DNA flanking the left-hand side of the transposon. Sequence analysis of the *Nde*I–*Kpn*I region revealed nine ORFs, eight with homology to COF pilus genes (Fig. 3). The rest of the predicted ORFs were identified and sequenced by PCR of a 4 kb fragment using primers to *orfD* (similar to *cofD*) and to *E. coli* gene *mrsA* predicted to flank the 5′ end of the pilus gene cluster. The 13 858 bp fragment has been deposited into GenBank (accession number AY158001).

Examination of the 13.8 kb DNA sequence indicated that it contained 14 ORFs and one partial ORF, shown in

```
                    Walker box A                    Asp box
CfcH    181  LSNPYGAYVIAGTTGSGKSTTLKNLMEWLQINRYDDRGCFLTVEDPVEYQIYGAKQSSVV 240
CofH    173  LSNPYGVYFIAGTTGSGKSTTLKNMMEWMQINRYDDKGCFLTVEDPVEYQIYGAKQSSVL 232
TcpT    225  LNTSYGLFIVSGTTGSGKSTSLKKYIELLFFNKYKGKGCFVTVEDPVEYLISGAQQSSIV 284
BfpD    250  RNLPIGINIISGPTGSGKSTTLKNLLELLYIEKKKKVNII-SIEDPPEYEIDGTAQLPIT 308

                         Asp box
CfcH    241  SVDGGGFHAAIKSALRRDPDVLMVGEIRDNVSANALAGAVESGHYCFTTVHAGNIVSLLQ 300
CofH    233  DGENGGFHSAIKSSLRRDPDVLMVGEIRDPVSSNALAGAVESGHYCFTTVHAGNIVTLLQ 292
TcpT    289  DKTKNPFADAVRSAMRRDPDVIMIGEIRDKPTVEALSSAVESGHYCLTTIHAGSVVSVLQ 348
BfpD    314  AQRGEEYRKAITAALRSDPDIIMPGEARDAEVINLLFTAAMTGHQVWTSLHANNALAIFD 373

                                    * *                 *
CfcH    301  RMSALGISSDKLSTPGFIAGLQCQKLVPVLCDACKQLYAEKVIAGRTFTLYQSSEQGCVK 360
CofH    293  RLSALGISSDKLSTPGFIAGLQCQKLIPVLCDNCKLEYAVN-ILGRRFTLCKVNESGCER 351
TcpT    349  RLSGLGMKADKIASPGFLAGITSQKLIPELCPSCKVSFVDERYQRAVFSANENG---CEA 405
BfpD    374  RLKDQGVDEFKLTDPELITGLVAQRLVRKLCAQCSITLTEYIASGGEISDTDRKIISGHE 433

         *                  Walker box B
CfcH    361  CRHSGIR----------GRQLAMEYFLPTYDELAAISRQEWLNVYTLWREKRKSVAGL 408
CofH    352  CKNTGVK----------SRQLVMEYLIPTQRELEALANQHWLNVYSIWREKRFTSSGI 400
TcpT    406  CNHSGFK----------GRLLLLETLVPTVEDLELVASENWVSLYRKYRERRFIKTGK 453
BfpD    434  TSVRFPNPRAKKCCRDGYNGRTILAEVIEPDSKLLRLVAEGKREDAQH            480

CfcH    409  AEGFEIREKVMAYVLQGRIDARWFAMEFGVLPVEDLEMIVEKIQ                452
CofH    401  SEGFSIKEKVFMHVLRKRACYQWFIMEFGEVPDTDMEVLHEKIQ                444
TcpT    454  KGLGEGFSIKDKAYYNVLKGKVCHEYF                                 480
```

**Fig. 2.** Alignment of the deduced amino acid sequence of CfcH with those of known type IV pilus nucleotide-binding proteins. Identical amino acids are indicated by a grey background. Gaps introduced for alignment are represented by dashes. The region of CfcH marking the nucleotide-binding domain is aligned with ETEC CofH (BAB62904), *V. cholerae* TcpT (AAK20762) and EPEC BfpD (BAA84843). The Walker box A, Asp boxes and Walker box B are underlined. The conserved CXXC motifs are marked by asterisks.

Fig. 3. The features of these genes and their predicted proteins are listed in Table 3. The region commenced with the end of an ORF with a high degree of similarity to MrsA, a 47.5 kDa putative phosphoglucomutase/phosphomannomutase of *E. coli* K-12 that is thought to play a role in the metabolism of glucose (Blattner *et al*., 1997). Homologues of the MrsA protein are also found in the genomes of EHEC 0157:H7 (Makino *et al*., 1999; Perna *et al*., 2001), *Salmonella typhimurium* (McClelland *et al*., 2001), *Salmonella typhi* (Parkhill *et al*., 2001), *Yersinia pestis* (Deng *et al*., 2002) and *V. cholerae* (Heidelberg *et al*., 2000). The final two ORFs in the 13.8 kb sequence have high identity to SecG, a putative protein involved in protein secretion, and to *leuU*, a tRNA-Leu of *E. coli* K-12 (Blattner *et al*., 1997). Homologues of these proteins are also found in EHEC 0157:H7 (Makino *et al*., 1999; Perna

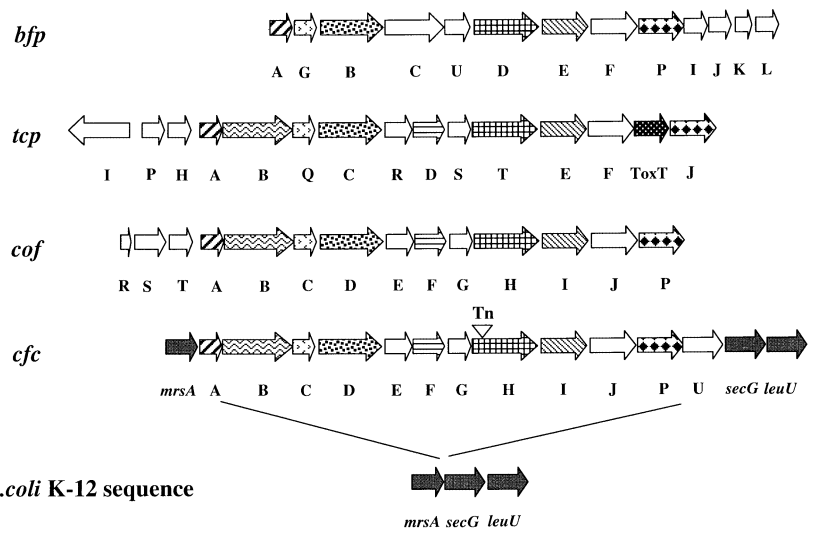


**Fig. 3.** Genetic organizations of *cfc*, *cof*, *tcp* and *bfp* gene clusters. The homologous genes are indicated by the same shading patterns.

**Table 3.** Features of the *cfc* genes and deduced proteins.

| Gene or ORF | | | Deduced protein | | | | |
|---|---|---|---|---|---|---|---|
| Name | Position (start–stop) | %G+C | Length (amino acids) | Size (kDa) | Similar protein(s); source (accession no.) | % Identical/% similar (no. of amino acids) | Predicted function or location |
| *mrsA* | 1–579 | 55.6 | NA | NA | MrsA; *E. coli* K-12 (AAC76208) | 89/90 192 | Phosphoglucomutase |
| *cfcA* | 979–1695 | 44.5 | 238 | 24.9 | LgnA; ETEC (AAC33154) | 42/61 (233) | Type IV pilin |
| | | | | | CofA; ETEC (BAB62897) | 38/61 (233) | |
| *cfcB* | 1814–3697 | 45.7 | 627 | 69.5 | CofB; ETEC (BAB62898) | 25/39 (626) | Type IV pilin-like protein |
| *cfcC* | 3728–4159 | 43.1 | 143 | 16.0 | CofC; ETEC (BAB62899) | 41/60 (128) | Unknown |
| | | | | | TcpQ: *V. cholerae* (AAF93993) | 25/47 (86) | |
| | | | | | BfpG: EPEC (BAA84839) | 21/43 (118) | |
| *cfcD* | 4167–5636 | 43.5 | 489 | 54.1 | CofD; ETEC (BAB62900) | 46/66 (489) | Outer membrane lipoprotein |
| | | | | | TcpC; *V. cholerae* (AAF93994) | 25/49 (450) | |
| *cfcE* | 5649–6191 | 48.4 | 180 | 20.5 | CofE; ETEC (BAB62901) | 31/55 (138) | Inner membrane |
| *cfcF* | 6195–7091 | 41.5 | 298 | 33.0 | CofF; ETEC (BAB62902) | 23/40 (262) | Inner membrane |
| *cfcG* | 7079–7540 | 53.2 | 153 | 16.9 | CofG; ETEC (BAB62903) | 33/63 (56) | Unknown |
| *cfcH* | 7832–9190 | 50.9 | 452 | 50.6 | CofH; ETEC (BAB62904) | 62/76 (445) | Inner membrane nucleotide-binding protein |
| | | | | | TcpT; *V. cholerae* (AAK20762) | 44/65 (404) | |
| | | | | | BfpD; EPEC (BAA84843) | 29/44 (305) | |
| *cfcI* | 9228–10202 | 49.7 | 325 | 36.0 | CofI; ETEC (BAB62905) | 54/74 (338) | Inner membrane |
| | | | | | TcpE; *V. cholerae* (AAK20763) | 36/60 (339) | |
| | | | | | BfpE; EPEC (BAA848844) | 25/50 (347) | |
| *cfcJ* | 10202–11065 | 43.9 | 287 | 32.3 | | | Unknown |
| *cfcP* | 11245–12090 | 55.8 | 846 | 32.0 | CofP; ETEC (BAB62907) | 41/55 214 | Prepilin peptidase |
| | | | | | PilD; *N. gonorrhoeae* (P33566) | 40/52 235 | |
| *cfcV* | 12087–13217 | 54.8 | 376 | 40.9 | Putative M23/M37 peptidase motif (pfam01551) | 33/46 (80) | Peptidase |
| *secG* | 13431–13763 | 53.0 | 110 | 11.4 | | | Protein export |
| tRNA-leu | 13779–13858 | 55.7 | NA | NA | | | tRNA-leu |

*et al*., 2001), *S. typhimurium* and *S. typhi* (McClelland *et al*., 2001; Parkhill *et al*., 2001). In these four genomes, the MrsA, SecG and tRNA-Leu are all encoded consecutively and in the same gene order (Fig. 3). The difference in the *C. rodentium* region sequenced is that a cluster of 12 genes has inserted between *mrsA* and *secG*. Sequence analysis of this novel region revealed similarity to the COF pilus operon of ETEC, in both gene order and similarity at the amino acid level (Fig. 3; Table 3). This new gene cluster has been designated *cfc* (for colonization factor *Citrobacter*). The *cof* nomenclature has been followed when naming individual genes within the *cfc* operon.

### *Properties of* cfc *genes and deduced proteins*

The 12 tandemly arranged genes (*cfcA*, *B*, *C*, *D*, *E*, *F*, *G*, *H*, *I*, *J*, *P and V*) within the *cfc* gene cluster are all encoded in the same translational orientation and may constitute an operon. The *cfcA* gene is preceded by a ribosome binding site, and the region immediately upstream of this contains a putative promoter sequence (not shown), as predicted by the NNPP program (http://dot.imgen.bcm.tmc.edu) (Reese *et al*., 1996). The major features of the *cfc* genes and deduced proteins are summarized in Table 3. The *cfc* gene cluster encodes proteins similar to those involved in the production of class B type IV pili including the COF and longus pili of ETEC (Giron *et al*., 1994; Taniguchi *et al*., 1995), TCP of *V. cholerae* (Taylor *et al*., 1987) and BFP of EPEC (Giron *et al*., 1991).

The first gene in the pilus cluster, *cfcA*, encodes a 238-amino-acid protein (24.9 kDa) that is a putative pilin belonging to the type IV pilin family. Alignments of the conserved N-terminal sequence of CfcA with other type IV pilin proteins is shown in Fig. 4. The type IV pilins have a number of features in common including a characteristic leader peptide sequence of 25–30 amino acids that is usually highly conserved, a hydrophobic N-terminal domain and also two cysteine residues near the C-terminus. These proteins are produced initially as a prepilin and are then processed to a mature pilin by cleavage at a highly conserved site (QXG↓F[M]T[S]LXE) located close to their N-termini (Taniguchi *et al*., 2001). CfcA has a leader sequence of 27 amino acids and possesses the cleavage consensus site (QRG↓VTLLE). Therefore, this protein may well be processed to a mature 211-amino-acid pilin by cleavage between Gly-27 and Val-28 by a prepilin peptidase (predicted to be encoded by *cfcP*). CfcA also possesses two cysteine residues, Cys-135 and Cys-197 (of the mature pilin).

*cfcB* encodes a 627-amino-acid protein, and *cfcC* encodes a 143-amino-acid protein, both with unknown function. *cfcD* encodes a 489-amino-acid protein that has similarity to putative outer membrane proteins CofD (46%

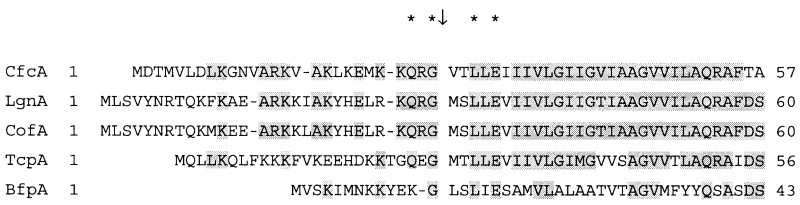


**Fig. 4.** Alignment of the deduced amino acid sequence of CfcA with those of known type IV pilins. Identical amino acids are indicated by a grey background. Gaps introduced for alignment are represented by dashes. The type IV pilin cleavage site is shown by an arrow. The conserved glycine, leucine and glutamic acid residues are marked by asterisks.

identity), TcpC (25% identity) (Ogierman and Manning, 1992), BfpB (22% identity) (Ramer *et al*., 1996) and PilN (20% identity) of EHEC strain 0113:H21 (Srimanote *et al*., 2002). CfcE and CfcF have similarity at the amino acid level to putative inner membrane proteins CofE and CofF, respectively (Taniguchi *et al*., 2001), and CfcG has similarity to CofG, with unknown function. *cfcI* encodes a 325-amino-acid protein that has similarity at the protein level to CofI (54% identity), TcpE (36% identity) and BfpE (25% identity). All these proteins belong to the bacterial type II secretion system protein F family, are all around 400 amino acids in length, are highly hydrophobic and are thought to be integral inner membrane proteins (InterPro, IPR003004).

*cfcJ* encodes a 287-amino-acid protein that has no similarity to any sequence in the GenBank DNA or protein databases. In contrast, *cfcP*, which encodes an 846-amino-acid protein, has strong similarity at the amino acid level to type IV prepillin peptidases. CfcP also contains two pairs of cysteine residues (CXXC) starting at amino acids 60 and 85 that are present in all type IV prepillin peptidases and have been shown to be necessary for the enzymatic activity of *Pseudomonas aeruginosa* PilD (Strom *et al*., 1993a,b). The final ORF of the *cfc* gene cluster, *cfcV*, encodes a 376-amino-acid protein, of which amino acids 203–283 have similarity to another peptidase family, M23/M37, which includes a number of endopeptidases with a range of specificities (pFam01151, GenBank). So it is possible that CfcV, in addition to the putative prepillin peptidase CfcP, cleaves proteins encoded within the *cfc* gene cluster.

### *CfcH and CfcI are required for colonization of* C. rodentium *in mice*

Strain P5H2 (Tn::*cfcH*) is avirulent in mice, with a phenotype similar to that displayed by the intimin mutant DBS255 and LEE mutants P1H4 (Tn::*espD*) and P10G3 (Tn::*escD*). However, it is possible that the transposon is also having polar effects on downstream genes in the putative operon (*cfcI*, *cfcJ*, *cfcP* and *cfcV*). To exclude this possibility, strain P5H2 was transformed with plasmid pICC250 (wild-type *cfcH* carried on plasmid pWSK29). This strain P5H2(pICC250) was then tested in mice as a single infection along with wild-type strain DBS100 and mutant strain P5H2. The results are shown in Table 4. In agreement with previous experiments, strain P5H2 was unable to colonize the colon, with no bacteria isolated from stools 3 days after inoculation. However strain P5H2(pICC250) was able to colonize the colon, with cfu isolated from the colon at a comparable level to the wild-type strain and colon weights that were significantly higher than those taken from mice infected with strain P5H2 ($P < 0.01$) (Table 4).

To determine whether other genes in the *cfc* operon are also required for colonization, a non-polar *aphT* kanamycin cassette that lacks a transcriptional terminator (Galan *et al*., 1992) was introduced into *cfcI* as described in *Experimental procedures*, producing strain ICC175. The ability of strain ICC175 to colonize mice was investigated by carrying out a single infection. Twelve days after inoculation, strain ICC175 was present in the colons of mice but at 5 logs lower than the wild-type strain (Table 4). By day 14 after inoculation, strain ICC175 was no longer recoverable from colons (data not shown). The wild-type colonization phenotype was restored when *cfcI* was introduced on a low-copy-number plasmid, pICC251 (Table 4). These results indicate that CfcH and CfcI proteins are both essential for efficient colonization of the mouse colon.

## Discussion

In this study, we have used STM to isolate genes involved

**Table 4.** Virulence of *C. rodentium* strains in mice.[a]

| Bacterial strain | No. of mice | Average colon weight[b] (g ± SE) | Bacterial colonization[c] (cfu per colon ± SE) |
|---|---|---|---|
| Wild-type DBS100 | 6 | 0.363 ± 0.013 | $(3.05 \pm 0.82) \times 10^8$ |
| P5H2(Tn::*cfcH*) | 3 | 0.196 ± 0.012*** | 0 |
| P5H2(pICC250) | 3 | 0.304 ± 0.024 | $(2.29 \pm 1.08) \times 10^8$ |
| ICC175(Δ*cfcI*) | 4 | 0.362 ± 0.013 | $(2.45 \pm 1.52) \times 10^3$ |
| ICC175(pICC251) | 3 | 0.348 ± 0.013 | $(4.69 \pm 1.44) \times 10^8$ |

**a.** C3H/Hej mice were infected with various *C. rodentium* strains and the colons collected at day 12 after inoculation.
**b.** SE, standard error. 8 cm of distal colon was weighed after removal of stool pellets. Significantly decreased colon weight compared with mice infected by wild-type strain DBS100 is indicated by ***$P < 0.001$, as analysed by one-way analysis of variance (ANOVA) test.
**c.** cfu, colony-forming units. The lower limit of detection is 50 cfu per colon.

in *C. rodentium* colonization of mice. The screen has identified both known and new virulence factors and has led to the discovery of a novel type IV pilus gene cluster of *C. rodentium*, which is absent from the sequenced A/E pathogen EHEC 0157:H7.

Most STM screens have used an optimal pool size of 96 mutants. However, oral models of infection often require a smaller pool size of 48 mutants for the majority of strains to colonize the gastrointestinal (GI) tract and be recovered successfully (Chiang and Mekalanos, 1998; Mecsas *et al*., 2001; Maroncle *et al*., 2002). In our *C. rodentium* model of infection, the pool size had to be reduced even further to 24–28 mutants. However, such a small pool size is not unprecedented, for example Tsolis *et al*. (1999) used a similar pool size of 24–30 mutants to screen *S. typhimurium* strains orally in mice and calves. In addition, Darwin and Miller (1999) found that oral infection with $10^{10}$ cfu of signature-tagged *Yersinia enterocolitica* strains only allowed up to 30 bacterial cells to establish infection. In this example, they chose to change the route of infection to intraperitoneal inoculation, which allowed them to use a pool size of 96 mutants (Darwin and Miller, 1999). These studies all suggest that there are several significant barriers to infection that limit the number of bacteria that can establish and colonize the GI tract successfully when introduced orally. These barriers could include the acidity of the stomach, along with the mucus and bile salts found in the small intestine and peristalsis plus epithelial cell shedding and resident microflora already established in the colon. The average transit time of inert particles through the gut from oral inoculation to isolation in stools has been determined to be ≈ 4–6 h in mice (Koopman *et al*., 1978). When we took stool samples 6 h after inoculation from mice infected with 48 signature-tagged *C. rodentium* mutants, all 48 strains were recoverable ($\approx 5 \times 10^8$ cfu $g^{-1}$ stool). Therefore, the majority of the cells were surviving transit through the GI tract. The main barriers limiting the number of *C. rodentium* cells able to establish infection may well occur in the colon itself and could include peristalsis and competition from resident microflora.

The smaller pool size used in our model meant that it would have been too time consuming to screen all 2400 mutants generated. Instead, 576 mutants were screened initially, resulting in the identification of 14 attenuated mutants. Analysis of these mutants revealed that around half contained transposon insertions in LEE-encoded genes. The animal infection experiments demonstrate just how important the *C. rodentium* LEE genes are for both colonization and formation of hyperplasia in mice. Strains that contain mutations in the *LEE1*, *LEE4* and *tir* operons, as well as in the *map* and *escD* genes, were all outcompeted at 72 h after inoculation by wild-type strains expressing the LEE type III secretion system and intimin. This was seen in both the pool screening and mixed infections. The fact that, in a single infection, the *map* mutant colonized the gut at a comparable level to the wild type suggests that, although not essential, Map gives the bacterium some competitive advantages. Whether this is related to the involvement of Map in the induction of colonic hyperplasia is at present not known. However, Map was shown recently to play a direct role in cell signalling during infection *in vitro* (Kenny *et al*., 2002). It has been reported previously that, after mixed infection *in vitro*, intimin mutants can *trans*-complement type III secretion mutants (Rosenshine *et al*., 1996; Shaw *et al*., 2002). The fact that both intimin and type III secretion mutants were outcompeted in the screen suggests that infection with *C. rodentium* is initiated at discrete sites and by single organisms.

Insertions in non-LEE genes, which also have a vital role to play in virulence, were also isolated. These included two strains that were completely avirulent, P5H2 (Tn::*cfcH*) and P5H5 (no similarity to sequences in the databases). The other four strains (P3F1, P3G5, P4C4 and P10D1) contained mutations that had only a subtle affect on virulence, with normal levels of colonization in single infection experiments, but were outcompeted in mice in mixed infection experiments with the wild type. Of note is the fact that the transposon insertion in strain P3G5 was found to be 736 bp downstream of the *C. rodentium* homologue of gene *lifA/efa1*.

The most interesting non-LEE gene identified was the pilus biogenesis nucleotide-binding protein homologue, *cfcH*. Further sequencing of this region revealed that *cfcH* is encoded within a novel gene cluster of *C. rodentium*. The 12 genes in this cluster have identity at the amino acid level to type IV class B pilus biogenesis genes of ETEC, EPEC and *V. cholerae*. These type IV pili share some structural, biochemical and functional features in common (reviewed by Strom and Lory, 1993). They have been grouped into two classes, with class B pilins being associated with intestinal infections (reviewed by Giron *et al*., 1997).

Thus far, we have been unable to visualize CFC using negative staining of *C. rodentium* grown in defined laboratory media. Moreover, we never observed piliated *C. rodentium in vivo* during infection of the colon where the bacterium is associated with the mucosa via A/E lesions. This might suggest that expression of the *cfc* operon is highly regulated.

This study is not the first occasion on which an STM screen has identified type IV pilus genes that are essential for colonization. Chiang and Mekalanos (1998) screened a library of *V. cholerae* mutants in mice and isolated five strains with transposon insertions in the known virulence operon TCP (toxin co-regulated pilus). The mutations were in genes *tcpA* (pilin), *tcpE*, *tcpF*, *tcpT* and *toxT*.

Interestingly, *tcpT* encodes the pilus nucleotide-binding protein, which is the homologue of *cofH* and *cfcH*. This *tcpT* mutant strain was still able to produce both cholera toxin and the TcpA pilin, but was highly attenuated in mixed infections, with a competitive index of 0.0004 (Chiang and Mekalanos, 1998). Other studies on *tcpT* mutants have shown that these bacteria are unable to translocate TcpA from the inner membrane and do not produce pili on the cell surface (Iredell and Manning, 1997).

Our study has shown that two putative pilus biogenesis genes are required for colonization in mice by *C. rodentium*. The function of the *cfcI* homologue has not yet been fully elucidated in the other type IV pili, and this mutation does not affect colonization as severely as the mutation in *cfcH*. This situation is mirrored in TCP, where the *cfcH* homologue *tcpT* is outcompeted by the wild-type strain more severely (CI = 0.0004) compared with the *cfcI* homologue *tcpE* (CI = 0.11). It is not clear why the *tcpT* mutant is more attenuated than the *tcpE* mutant, as a mutation in *tcpE* results in bacteria that have been shown to be unable to express cell surface TCP, to autoagglutinate or to maintain virulence in the infant mouse model (Taylor *et al*., 1988). Homologues of *cfcH* and *cfcI* are also present in BFP of EPEC (*bfpD* and *bfpE* respectively) (Sohel *et al*., 1996; Stone *et al*., 1996). *BfpD* mutants cannot produce BFP, autoaggregate or cause localized adherence (Bieber *et al*., 1998; Anantha *et al*., 2000) and, more recently, *bfpE* mutants have been shown to have a similar phenotype (Blank and Donnenberg, 2001).

Homologues of nucleotide-binding protein *cfcH* have been speculated to be involved in producing energy via ATP hydrolysis, which is required for assembly and secretion of the pilus (Ogierman *et al*., 1996; Iredell and Manning, 1997; Anantha *et al*., 2000; Taniguchi *et al*., 2001). Therefore, by disrupting the *cfcH* gene with a transposon insertion, the bacteria may well be unable to assemble and/or secrete CFC pili. It is therefore likely that this results in *C. rodentium* cells that are unable to adhere to colonic epithelia and so are unable to establish an infection. The *cfcI* mutation still allows colonization of the colon, but at a much lower level than for wild-type bacteria. Strain ICC175 (Δ*cfcI*) also still causes hyperplasia (as measured by colon weight), suggesting that this mutation does not prevent expression of intimin, the type III secretion system and its effectors.

Donnenberg and Kaper (1992) proposed a three-stage model of EPEC adherence to cultured human epithelial cells. Step one is initial binding to intestinal cells mediated by the type IV bundle-forming pili (BFP). The second step is the translocation of LEE effector proteins into host cells via the LEE-encoded type III secretion system. The final stage is the more intimate attachment of EPEC to epithelial cells via the adhesin, intimin, and its translocated receptor, Tir. This induces actin-rich pedestal formation by the host cell, which the bacterium sits on. A BFP homologue has not been found previously in *C. rodentium*, and it is not known how this pathogen manages to adhere efficiently to mouse colonic epithelial cells before the LEE proteins are expressed and translocated. It is tempting to speculate that the type IV CFC pilus (which shares low-level identity to the BFP pilus), is involved in this initial adherence to mouse epithelial cells. Recent studies have shown that BFP appears to mediate adherence in a species-specific manner only to human-derived cells (Tobe and Sasakawa, 2002). Could the CFC pilus mediate species-specific adhesion of *C. rodentium* to mouse cells? If this proves to be the case, then by transferring the *cfc* cluster into EHEC or EPEC strains, it may prove possible to produce a clinically important human A/E strain that is able to colonize mice.

## Experimental procedures

### *Bacterial strains, plasmids and growth conditions*

The bacterial strains and plasmids used in this study are shown in Table 5. Bacteria were grown on Luria–Bertani (LB) medium unless otherwise specified, with additional antibiotics where appropriate at the following concentrations: nalidixic acid (nal), 50 μg ml$^{-1}$; kanamycin (kan), 100 μg ml$^{-1}$; chloramphenicol (chl), 50 μg ml$^{-1}$; ampicillin (amp), 100 μg ml$^{-1}$; pipericillin (pip), 75 μg ml$^{-1}$. Minimal media consisted of M9 salts or Vogel–Bonner (VB) salts supplemented with 1% glucose.

### *Generation of the transposon mutant library*

A pool of 48 preselected signature-tagged mini-Tn*5*Km2 transposons in plasmid pUT was used to generate 48 sets of mutants. These tagged transposons, which were preselected on the basis that they amplify and hybridize well but do not cross-hybridize with other tags, were obtained from Microscience. Plasmid DNA containing each tagged kan$^r$ transposon was transformed into competent *E. coli* S17-1λpir cells by heat shock. Transformants were used to transfer each of the 48 tagged transposons to the recipient strain *C. rodentium* nal$^r$ by conjugation. Briefly, 400 μl of donor and recipient strains was mixed in 5 ml of 10 mM $MgSO_4$ and filtered through a 25 mm 0.45 μm pore size membrane filter (Millipore). The filters were placed bacteria side up on M9 minimal media and incubated for 16 h. The filters were then placed in LB broth and incubated at 37°C for 40 min, then plated onto LB, nal and kan to select for *C. rodentium* exconjugants. The exconjugants were checked for failure to grow on LB and pip to identify any mutants that had retained the pUT plasmid (pipericillin was used as *C. rodentium* is partially resistant to ampicillin). The resulting nal$^r$, kan$^r$, pip$^s$ exconjugants were stored as pools of 48 mutants in 96-well microtitre dishes containing LB, nal and kan. For long-term storage at –80°C, 15% glycerol was incorporated into the medium.

**Table 5.** Bacterial strains and plasmids used in this study.

| Strain or plasmid | Description | Source |
|---|---|---|
| *Escherichia coli* | | |
| DH5α | F′endA1 supE44 thi-1 hsdR17($r_k^-m_k^+$)recA1 gyrA relA1 Δ(lacIZYA-argF)U169 deoR (ø80 dLacΔ(lacZ)M15) | Invitrogen |
| S17-1 λpir | RecA thi pro hsdR$^-$M$^+$ RP4::2-Tc::Mu::Km Tn*7* lysogenized with λpir phage | de Lorenzo *et al.* (1990) |
| *Citrobacter rodentium* | | |
| DBS100 | Wild-type strain | Schauer and Falkow (1993a) |
| DBS255 | *eae* deletion mutant | Schauer and Falkow (1993b) |
| ICC174 | DBS255 carrying signature tag H6 | This study |
| ICC175 | *CfcI* deletion mutant | This study |
| Plasmids | | |
| pUC18 | Amp$^r$; high-copy-number cloning vector | Invitrogen |
| PCR 2.1 | Amp$^r$, kan$^r$; high-copy-number TA cloning vector | Invitrogen |
| pWSK29 | Amp$^r$; low-copy-number cloning vector | Wang and Kushner (1991) |
| pACYC184 | Chl$^r$, tet$^r$, cloning vector | New England Biolabs |
| pKD46 | Amp$^r$; expressing λ red recombinase | Datsenko and Wanner (2000) |
| pICC247 | *Kpn*I P5H2 transposon-containing fragment in pUC18 | This study |
| pICC248 | *Bam*HI fragment containing *cfcI* from plasmid pICC247 | This study |
| pICC249 | AphT cassette inserted into *Sna*BI site of *cfcI* in pICC248 | This study |
| pICC250 | *Citrobacter cfcH* cloned in pWSK29, Amp$^r$ | This study |
| pICC251 | *Citrobacter cfcI* cloned in pWSK29, Amp$^r$ | This study |

### *Infections of mice*

Male, specific pathogen-free, C3H/Hej mice (4–5 weeks old) were purchased from Harlan Olac. All mice were housed in individual ventilated cages with free access to food and water. Unanaesthetized mice were orally gavaged with 200 μl of bacterial suspension using a gavage needle. The viable count of the inoculum was determined by retrospective plating on LB agar containing appropriate antibiotics.

For the screening of pools, each of the mutants was grown individually in the well of a microtitre dish in 200 μl of LB, nal and kan and incubated for 18 h at 37°C (static). The growth of mutants was assessed by determining the optical density (OD) for each well with a microtitre plate reader. All mutants should have reached stationary phase and have a similar OD. Any mutant with a significant growth defect in LB was therefore identified at this stage. Bacterial cells (100 ml) were then pooled, spun at 4000 r.p.m. for 10 min and resuspended in 1 ml of PBS, resulting in $\approx 5 \times 10^{10}$ cfu ml$^{-1}$, which was confirmed by plating dilutions onto LB agar containing nal and kan. Three mice received $\approx 1 \times 10^{10}$ cfu consisting of 24–28 different mutants by oral gavage. The remaining inoculum was used to harvest DNA for preparation of the input probe. Stool samples were taken from the mice at days 2 and 3 after inoculation, and the mice were killed and the colons removed at day 5 after inoculation. Stool samples and colons were mashed up in PBS and plated onto LB agar containing nal and kan. More than 10 000 colonies were used to isolate DNA at each time point. Signature tags were amplified and radiolabelled, using primers P2 and P4, as described previously (Hensel *et al.*, 1995). The input and recovered probes were used in hybridizations with replica dot blots generated using plasmid DNA from each of the 48 tagged transposons cloned in the high-copy-number vector pCR2.1 TOPO (constructs supplied by Microscience). Mutants with tags that hybridize with the input probe but not with the recovered probe from day 3 stools and day 5 colon samples were considered to have a colonization defect and were characterized further.

In mixed infection experiments, mutant (nal$^r$, kan$^r$) and wild-type (nal$^r$) bacteria were grown to stationary phase in LB broth, and equal amounts of bacteria ($10^9$ each in 100 μl of PBS) were mixed then orally gavaged into mice. To determine the proportion of wild-type to mutant bacteria, dilutions of the inoculum were plated onto media containing nal only and onto LB, nal and kan. After 5 days, the mice were killed by cervical dislocation, and bacteria were recovered by plating dilutions of homogenized colon to media containing nal and containing nal and kan. The virulence of each mutant was analysed in two or more animals, and the results are given as an average. The competitive index (CI) was calculated as the proportion of mutant to wild-type bacteria recovered from animals divided by the proportion of mutant to wild type in the inoculum.

In single infection experiments, bacteria were grown to stationary phase in LB broth plus appropriate antibiotic. The bacteria were pelleted by centrifugation, resuspended in an equal volume of PBS and gavaged into mice ($\approx 2 \times 10^8$ cfu).

### *Cloning and sequencing transposon insertion sites*

Southern analysis was performed on genomic DNA from colonization-defective mutants digested with each of six restriction enzymes (*Bgl*II, *Eco*RI, *Kpn*I, *Pst*I, *Sal*I and *Sph*I). DNA was electrophoresed through 0.8% agarose gels and transferred to nylon membranes (Hybond N+ Amersham) by standard methods. The membranes were probed with the kan resistance cassette from mini-Tn*5* labelled by the megaprime DNA labelling system (Amersham). An enzyme that gave hybridizing fragments of between 2.5 and 7.5 kb was chosen for each mutant and used to digest 40 μg of genomic DNA, which was size-fractionated on a 0.8% agarose gel and the appropriate fragments excised and purified (Qiaquick;

Qiagen). The fragments were then ligated into pUC18 digested with the same enzyme, the resulting ligation mix was used to transform *E. coli* DH5α, and kanamycin-resistant colonies were selected. The DNA adjacent to the tagged mini-Tn*5*Km2 insert was sequenced using transposon primers I and O (Hensel et al., 1995). Sequence homologies were performed using GenBank databases and BLAST programs found at http://www.ncbi.nih/gov. A computer-assisted ORF search was performed using the following criteria: an ORF would encode a minimum of 100 amino acids; ATG as the translational start codon and an *E. coli* consensus ribosome binding site (RBS) that was located at an optimal distance upstream of the ATG.

### *Cloning and sequencing of the 12.7 kb* cfc *gene cluster*

Two restriction enzymes were used to clone large fragments of the Cfc operon: *Kpn*I (7.3 kb including Tn and chromosomal DNA flanking the 5′ end) and *Nde*I (6.2 kb including transposon and chromosomal DNA flanking both ends). Kan$^r$ *Kpn*I and *Nde*I fragments were cloned into pUC18 digested with the appropriate enzyme and sequenced (MWG, primer walking service). The final 4 kb of the Cfc operon was amplified using primers from *cfcD* and the predicted *E. coli* sequence flanking the *cfc* pilus. PCR was carried out using mutant P5H2 genomic DNA as template; a product of ≈4 kb was identified, cloned into vector pCR2.1 TOPO (Invitrogen) and sequenced.

### *Construction of the non-polar mutation in* cfcI

A 4.1 kb *Bam*HI fragment was excised from plasmid pICC247 (Table 5) and cloned into pACYC184 generating plasmid pICC248. A non-polar *aphT* cassette, which conferred kanamycin resistance, was incorporated into the *Sna*BI site of pICC248 (255 bp downstream of the *cfcI* start codon) creating plasmid pICC249. In order to enhance allelic exchange, plasmid pKD46, an easily curable low-copy-number plasmid that expressed lambda red recombinase, described by Datsenko and Wanner (2000), was transformed into DBS100 by electroporation, generating strain DBS100 (pKD46). pKD46 has been shown to aid chromosomal recombination of foreign DNA in *E. coli* K-12 (Datsenko and Wanner, 2000).

A 4.9 kb *Bam*HI fragment was excised from plasmid pICC249; this contained the *cfcI* gene with an in frame insertion of the *aphT* cassette and flanking 1 kb of DNA upstream of *cfcI* and 2 kb of DNA downstream of *cfcI*. This fragment was transformed by electroporation into strain DBS100 (pKD46). Clones were grown on LB and kan to select for kanamycin resistance. pKD46 was cured by growth at 37°C. The mutation in *cfcI*, strain ICC175, was verified by PCR and DNA sequencing.

### *Cloning of wild-type* cfcH *and* cfcI *and complementation of strains P5H2 and ICC175*

*CfcH* was amplified by PCR and cloned into pWSK29 (Wang and Kushner, 1991), a low-copy-number expression vector with a T3 promoter, generating plasmid pICC250. The plasmid was transformed into strain P5H2 (Tn::*cfcH*) using electroporation creating strain P5H2(pICC250). *CfcI* was also amplified and cloned into pWSK29, generating plasmid pICC251, which was then transformed into strain ICC175. Complementation of the *cfcH* and *cfcI* mutants was verified by restoration of a wild-type phenotype, shown by single infection experiments in mice.

## Acknowledgements

We thank Microscience Ltd for making their preselected signature-tagged transposons available for this screen. This work was supported by the BBSRC.

## References

Anantha, R.P., Stone, K.D., and Donnenberg, M.S. (2000) Effects of bfp mutations on biogenesis of functional enteropathogenic *Escherichia coli* type IV pili. *J Bacteriol* **182:** 2498–2506.

Ashkenazi, S., Larocco, M., Murray, B.E., and Cleary, T.G. (1992) The adherence of verocytotoxin-producing *Escherichia coli* to rabbit intestinal cells. *J Med Microbiol* **37:** 304–309.

Barthold, S.W., Coleman, G.L., Bhatt, P.N., Osbaldiston, G.W., and Jonas, A.M. (1976) The etiology of transmissible murine colonic hyperplasia. *Lab Anim Sci* **26:** 889–894.

Bieber, D., Ramer, S.W., Wu, C.Y., Murray, W.J., Tobe, T., Fernandez, R., *et al.* (1998) Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science* **280:** 2114–2118.

Blank, T.E., and Donnenberg, M.S. (2001) Novel topology of BfpE, a cytoplasmic membrane protein required for type IV fimbrial biogenesis in enteropathogenic *Escherichia* coli. *J Bacteriol* **183:** 4435–4450.

Blattner, F.R., Plunkett, G., III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277:** 1453–1474.

Chiang, S.L., and Mekalanos, J.J. (1998) Use of signature-tagged transposon mutagenesis to identify *Vibrio cholerae* genes critical for colonization. *Mol Microbiol* **27:** 797–805.

Darwin, A.J., and Miller, V.L. (1999) Identification of *Yersinia enterocolitica* genes affecting survival in an animal host using signature-tagged transposon mutagenesis. *Mol Microbiol* **32:** 51–62.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97:** 6640–6645.

Dean-Nystrom, E.A., Bosworth, B.T., Moon, H.W., and O'Brien, A.D. (1998) *Escherichia coli* O157:H7 requires intimin for enteropathogenicity in calves. *Infect Immun* **66:** 4560–4563.

Deibel, C., Kramer, S., Chakraborty, T., and Ebel, F. (1998) EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells, where it appears as a tyrosine-phosphorylated 90 kDa protein. *Mol Microbiol* **28:** 463–474.

Deng, W., Li, Y., Vallance, B.A., and Finlay, B.B. (2001)

Locus of enterocyte effacement from *Citrobacter rodentium*: sequence analysis and evidence for horizontal transfer among attaching and effacing pathogens. *Infect Immun* **69:** 6323–6335.

Deng, W., Burland, V., Plunkett, G., III, Boutin, A., Mayhew, G.F., Liss, P., *et al.* (2002) Genome sequence of *Yersinia pestis* KIM. *J Bacteriol* **184:** 4601–4611.

DeVinney, R., Stein, M., Reinscheid, D., Abe, A., Ruschkowski, S., and Finlay, B.B. (1999) Enterohemorrhagic *Escherichia coli* O157:H7 produces Tir, which is translocated to the host cell membrane but is not tyrosine phosphorylated. *Infect Immun* **67:** 2389–2398.

Donnenberg, M.S., and Kaper, J.B. (1992) Enteropathogenic *Escherichia coli*. *Infect Immun* **60:** 3953–3961.

Donnenberg, M.S., Tzipori, S., McKee, M.L., O'Brien, A.D., Alroy, J., and Kaper, J.B. (1993) The role of the *eae* gene of enterohemorrhagic *Escherichia coli*. intimate attachment *in vitro* and in a porcine model. *J Clin Invest* **92:** 1418–1424.

Frankel, G., Phillips, A.D., Rosenshine, I., Dougan, G., Kaper, J.B., and Knutton, S. (1998) Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements. *Mol Microbiol* **30:** 911–921.

Galan, J.E., Ginocchio, C., and 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.

Giron, J.A., Ho, A.S., and Schoolnik, G.K. (1991) An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* **254:** 710–713.

Giron, J.A., Levine, M.M., and Kaper, J.B. (1994) Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli*. *Mol Microbiol* **12:** 71–82.

Giron, J.A., Gomez-Duarte, O.G., Jarvis, K.G., and Kaper, J.B. (1997) Longus pilus of enterotoxigenic *Escherichia coli* and its relatedness to other type-4 pili – a minireview. *Gene* **192:** 39–43.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., *et al.* (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406:** 477–483.

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., and Holden, D.W. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* **269:** 400–4033.

Iredell, J.R., and Manning, P.A. (1997) Translocation failure in a type-4 pilin operon: *rfb* and *tcpT* mutants in *Vibrio cholerae*. *Gene* **192:** 71–77.

Kenny, B., and Jepson, M. (2000) Targeting of an enteropathogenic *Escherichia coli* (EPEC) effector protein to host mitochondria. *Cell Microbiol* **2:** 579–590.

Kenny, B., DeVinney, R., Stein, M., Reinscheid, D.J., Frey, E.A., and Finlay, B.B. (1997) Enteropathogenic *E. coli* (EPEC). transfers its receptor for intimate adherence into mammalian cells. *Cell* **91:** 511–520.

Kenny, B., Ellis, S., Leard, A.D., Warawa, J., Mellor, H., and Jepson, M.A. (2002) Co-ordinate regulation of distinct host cell signalling pathways by multifunctional enteropathogenic *Escherichia coli* effector molecules. *Mol Microbiol* **44:** 1095–1107.

Klapproth, J.M., Scaletsky, I.C., McNamara, B.P., Lai, L.C., Malstrom, C., James, S.P., *et al.* (2000) A large toxin from pathogenic *Escherichia coli* strains that inhibits lymphocyte activation. *Infect Immun* **68:** 2148–2155.

Koopman, J.P., Kennis, H.M., and van Druten, J.A. (1978) Colonization resistance of the digestive tract and gastrointestinal transit time in SPF mice. *Lab Anim* **12:** 223–226.

Kresse, A.K., Schulze, K., Deibel, C., Ebel, F., Rohde, M., Chakraborty, T., and Guzmán, C.A. (1998) Pas, a novel protein required for protein secretion and attaching and effacing activities of enterohemorrhagic *Escherichia coli*. *J Bacteriol* **180:** 4370–4379.

de Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K.N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in Gram-negative eubacteria. *J Bacteriol* **172:** 6568–6572.

Luperchio, S.A., and Schauer, D.B. (2001) Molecular pathogenesis of *Citrobacter rodentium* and transmissible murine colonic hyperplasia. *Microbes Infect* **3:** 333–340.

McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., *et al.* (2001) Complete genome sequence of *Salmonella* enterica serovar *Typhimurium* LT2. *Nature* **413:** 852–856.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., and Kaper, J.B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci USA* **92:** 1664–1668.

McKee, M.L., Melton-Celsa, A.R., Moxley, R.A., Francis, D.H., and O'Brien, A.D. (1995) Enterohemorrhagic *Escherichia coli* O157:H7 requires intimin to colonize the gnotobiotic pig intestine and to adhere to HEp-2 cells. *Infect Immun* **63:** 3739–3744.

Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S., Kurokawa, K., *et al.* (1999) Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak. *Genes Genet Syst* **74:** 227–239.

Marches, O., Nougayrede, J.P., Boullier, S., Mainil, J., Charlier, G., Raymond, I., *et al.* (2000) Role of tir and intimin in the virulence of rabbit enteropathogenic *Escherichia coli* serotype O103:H2. *Infect Immun* **68:** 2171–2182.

Maroncle, N., Balestrino, D., Rich, C., and Forestier, C. (2002) Identification of *Klebsiella pneumoniae* genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis. *Infect Immun* **70:** 4729–4734.

Mead, P.S., Slutsker, L., Dietz, V., McCaig, L.F., Bresee, J.S., Shapiro, C., *et al.* (1999) Food-related illness and death in the United States. *Emerg Infect Dis* **5:** 607–625.

Mecsas, J., Bilis, I., and Falkow, S. (2001) Identification of attenuated *Yersinia pseudotuberculosis* strains and characterization of an orogastric infection in BALB/c mice on day 5 postinfection by signature-tagged mutagenesis. *Infect Immun* **69:** 2779–2787.

Moon, H.W., Whipp, S.C., Argenzio, R.A., Levine, M.M., and Giannella, R.A. (1983) Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli*. pig and rabbit intestines. *Infect Immun* **41:** 1340–1351.

Nataro, J.P., and Kaper, J.B. (1998) Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11:** 142–201.

Newman, J.V., Zabel, B.A., Jha, S.S., and Schauer, D.B. (1999) *Citrobacter rodentium espB* is necessary for signal transduction and for infection of laboratory mice. *Infect Immun* **67:** 6019–6025.

Nicholls, L., Grant, T.H., and Robins-Browne, R.M. (2000) Identification of a novel genetic locus that is required for in vitro adhesion of a clinical isolate of enterohaemorrhagic *Escherichia coli* to epithelial cells. *Mol Microbiol* **35:** 275–288.

Ogierman, M.A., and Manning, P.A. (1992) TCP pilus biosynthesis in *Vibrio cholerae* O1: gene sequence of tcpC encoding an outer membrane lipoprotein. *FEMS Microbiol Lett* **76:** 179–184.

Ogierman, M.A., Voss, E., Meaney, C., Faast, R., Attridge, S.R., and Manning, P.A. (1996) Comparison of the promoter proximal regions of the toxin-co-regulated tcp gene cluster in classical and El Tor strains of *Vibrio cholerae* O1. *Gene* **170:** 9–16.

Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., *et al.* (2001) Complete genome sequence of a multiple drug resistant *Salmonella* enterica serovar *Typhi* CT18. *Nature* **413:** 848–852.

Perna, N.T., Mayhew, G.F., Posfai, G., Elliott, S., Donnenberg, M.S., Kaper, J.B., *et al.* (1998) Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **66:** 3810–3817.

Perna, N.T., Plunkett, G., III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., *et al.* (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157: H7. *Nature* **409:** 529–533.

Phillips, A.D., Navabpour, S., Hicks, S., Dougan, G., Wallis, T., and Frankel, G. (2000) Enterohaemorrhagic *Escherichia coli* O157:H7 target Peyer's patches in humans and cause attaching/effacing lesions in both human and bovine intestine. *Gut* **47:** 377–381.

Ramer, S.W., Bieber, D., and Schoolnik, G.K. (1996) BfpB, an outer membrane lipoprotein required for the biogenesis of bundle-forming pili in enteropathogenic *Escherichia coli*. *J Bacteriol* **178:** 6555–6563.

Reese, M.G., Harris, N.L., and Eeckman, F.H. (1996) Large scale sequencing specific neural networks for promoter and splice site recognition. In *Biocomputing. Proceedings of the 1996 Pacific Symposium.* Hunter, L., and Klein, T.E. (eds). Singapore: World Scientific Publishing Co.

Rosenshine, I., Ruschkowski, S., Stein, M., Reinscheid, D.J., Mills, S.D., and Finlay, B.B. (1996) A pathogenic bacterium triggers epithelial signals to form a functional bacterial receptor that mediates actin pseudopod formation. *EMBO J* **15:** 2613–2624.

Sanchez-SanMartin, C., Bustamante, V.H., Calva, E., and Puente, J.L. (2001) Transcriptional regulation of the orf19 gene and the *tir-cesT-eae* operon of enteropathogenic *Escherichia coli*. *J Bacteriol* **183:** 2823–2833.

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

Schauer, D.B., and Falkow, S. (1993b) The *eae* gene of *Citrobacter freundii* biotype 4280 is necessary for colonization in transmissible murine colonic hyperplasia. *Infect Immun* **61:** 4654–4661.

Shaw, R.K., Daniell, S., Frankel, G., and Knutton, S. (2002) Enteropathogenic *Escherichia coli* translocate Tir and form an intimin-Tir intimate attachment to red blood cell membranes. *Microbiology* **148:** 1355–1365.

Shea, J.E., Hensel, M., Gleeson, C., and Holden, D.W. (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. *Proc Natl Acad Sci USA* **93:** 2593–2597.

Sherman, P., Cockerill, F., III, Soni, R., and Brunton, J. (1991) Outer membranes are competitive inhibitors of *Escherichia coli* O157:H7 adherence to epithelial cells. *Infect Immun* **59:** 890–899.

Sohel, I., Puente, J.L., Ramer, S.W., Bieber, D., Wu, C.Y., and Schoolnik, G.K. (1996) Enteropathogenic *Escherichia coli*: identification of a gene cluster coding for bundle-forming pilus morphogenesis. *J Bacteriol* **178:** 2613–2628.

Srimanote, P., Paton, A.W., and Paton, J.C. (2002) Characterization of a novel type IV pilus locus encoded on the large plasmid of locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect Immun* **70:** 3094–3100.

Stevens, M.P., van Diemen, P.M., Frankel, G., Phillips, A.D., and Wallis, T.S. (2002) Efa1 influences colonization of the bovine intestine by shiga toxin-producing *Escherichia coli* serotypes O5 and O111. *Infect Immun* **70:** 5158–5166.

Stone, K.D., Zhang, H.Z., Carlson, L.K., and Donnenberg, M.S. (1996) A cluster of fourteen genes from enteropathogenic *Escherichia coli* is sufficient for the biogenesis of a type IV pilus. *Mol Microbiol* **20:** 325–337.

Strom, M.S., and Lory, S. (1993) Structure–function and biogenesis of the type IV pili. *Annu Rev Microbiol* **47:** 565–596.

Strom, M.S., Bergman, P., and Lory, S. (1993a) Identification of active-site cysteines in the conserved domain of PilD, the bifunctional type IV pilin leader peptidase/N-methyltransferase of *Pseudomonas aeruginosa*. *J Biol Chem* **268:** 15788–15794.

Strom, M.S., Nunn, D.N., and Lory, S. (1993b) A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proc Natl Acad Sci USA* **90:** 2404–2408.

Taniguchi, T., Fujino, Y., Yamamoto, K., Miwatani, T., and Honda, T. (1995) Sequencing of the gene encoding the major pilin of pilus colonization factor antigen III (CFA/III) of human enterotoxigenic *Escherichia coli* and evidence that CFA/III is related to type IV pili. *Infect Immun* **63:** 724–728.

Taniguchi, T., Akeda, Y., Haba, A., Yasuda, Y., Yamamoto, K., Honda, T., *et al.* (2001) Gene cluster for assembly of pilus colonization factor antigen III of enterotoxigenic *Escherichia coli*. *Infect Immun* **69:** 5864–5873.

Tarr, P.I., Bilge, S.S., Vary, J.C., Jr, Jelacic, S., Habeeb, R.L., Ward, T.R., *et al.* (2000) Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. *Infect Immun* **68:** 1400–1407.

Taylor, R.K., Miller, V.L., Furlong, D.B., and Mekalanos, J.J. (1987) Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc Natl Acad Sci USA* **84:** 2833–2837.

Taylor, R., Shaw, C., Peterson, K., Spears, P., and Mekalanos, J. (1988) Safe, live *Vibrio cholerae* vaccines? *Vaccine* **6:** 151–154.

Tobe, T., and Sasakawa, C. (2002) Species-specific cell adhesion of enteropathogenic *Escherichia coli* is mediated by type IV bundle-forming pili. *Cell Microbiol* **4:** 29–42.

Tsolis, R.M., Townsend, S.M., Miao, E.A., Miller, S.I., Ficht, T.A., Adams, L.G., *et al.* (1999) Identification of a putative *Salmonella* enterica serotype *typhimurium* host range factor with homology to IpaH and YopM by signature-tagged mutagenesis. *Infect Immun* **67:** 6385–6393.

Tzipori, S., Gunzer, F., Donnenberg, M.S., de Montigny, L., Kaper, J.B., and Donohue-Rolfe, A. (1995) The role of the *eaeA* gene in diarrhea and neurological complications in a gnotobiotic piglet model of enterohemorrhagic *Escherichia coli* infection. *Infect Immun* **63:** 3621–3627.

Vallance, B.A., and Finlay, B.B. (2000) Exploitation of host cells by enteropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* **97:** 8799–8806.

Wachter, C., Beinke, C., Mattes, M., and Schmidt, M.A. (1999) Insertion of EspD into epithelial target cell membranes by infecting enteropathogenic *Escherichia coli*. *Mol Microbiol* **31:** 1695–1707.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1:** 945–951.

Wang, R.F., and Kushner, S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100:** 195–199.