DNA Sequence and Mutational Analyses of the pVir Plasmid of *Campylobacter jejuni* 81-176

David J. Bacon,¹ Richard A. Alm,² Lan Hu,¹ Thomas E. Hickey,¹ Cheryl P. Ewing,¹ Roger A. Batchelor,¹ Trevor J. Trust,² and Patricia Guerry¹*

*Naval Medical Research Center, Silver Spring, Maryland 20910,*¹ *and AstraZenecaR&D Boston, Waltham, Massachusetts 02451*²

Received 25 February 2002/Returned for modification 23 April 2002/Accepted 29 July 2002

The circular pVir plasmid of *Campylobacter jejuni* **strain 81-176 was determined to be 37,468 nucleotides in length with a GC content of 26%. A total of 83% of the plasmid represented coding information, and all but 2 of the 54 predicted open reading frames were encoded on the same DNA strand. There were seven genes on the plasmid in a continguous region of 8.9 kb that encoded orthologs of type IV secretion proteins found in** *Helicobacter pylori,* **including four that have been described previously (D. J. Bacon, R. A. Alm, D. H. Burr, L. Hu, D. J. Kopecko, C. P. Ewing, T. J. Trust, and P. Guerry, Infect. Immun. 68:4384–4390, 2000). There were seven other pVir-encoded proteins that showed significant similarities to proteins encoded by the plasticity zones of either** *H. pylori* **J99 or 26695. Mutational analyses of 19 plasmid genes identified 5 additional genes that affect in vitro invasion of intestinal epithelial cells. These included one additional gene encoding a component of a type IV secretion system, an ortholog of Cj0041 from the chromosome of** *C. jejuni* **NCTC 11168, two** *Campylobacter* **plasmid-specific genes, and an ortholog of HP0996 from the plasticity zone of** *H. pylori* **26695.**

Although *Campylobacter jejuni* is among the major causes of diarrheal disease worldwide (27) and was the first enteropathogen to have its complete genome sequence published (30), the mechanisms by which it causes disease remain unclear. All strains of *C. jejuni* produce cytolethal distending toxin, a cytotoxin that arrests eukaryotic cells in the $G₂$ phase of the cell cycle and induces interleukin-8 (IL-8) release from intestinal epithelial cells (17, 23, 32, 39). Motility and chemotaxis are required for *C. jejuni* to colonize the intestinal tracts of animals and humans (5, 31) and for invasion in vitro (12, 37, 42), perhaps because the flagella structure may secrete virulence determinants in the absence of another type III secretion system (30, 33). Although invasion of intestinal epithelial cells is thought to play a role in campylobacter virulence, there is a tremendous range in the in vitro invasion level among strains of *C. jejuni* (3, 20, 29). Strain 81-176, isolated from an outbreak of diarrheal disease (22), is one of the best-characterized strains of *C. jejuni* and one which has been shown to be virulent in two human volunteer studies (5; D. Tribble, unpublished data). Studies of *C. jejuni* strain 81-176 have indicated that this strain is internalized into intestinal epithelial cells at relatively high levels by a microtubule-dependent invasion system $(20, 100)$ 29) and that this in vitro invasion appears to correlate with disease in vivo in the ferret diarrhea model (3, 4, 42). Additionally, adherence and/or invasion in vitro results in the release of IL-8 from intestinal epithelial cells, independently of cytolethal distending toxin (17, 18). Adherence and invasion of strain 81-176 are modulated by phase variation of both the polysaccharide capsule (4) and lipooligosaccharide (LOS)

(16). Moreover, we have recently shown that a plasmid is involved in invasion of 81-176 (3). Partial DNA sequencing of this plasmid, called pVir, revealed the presence of four genes encoding homologs of a type IV secretion system. Type IV secretion systems have been described in numerous bacteria and are involved in DNA export, bacterial conjugation, and protein secretion (6, 7). Site-specific mutation of two of these putative type IV secretion genes on the 81-176 pVir plasmid resulted in a significant decrease in invasion of intestinal epithelial cells in vitro, as well as a reduction in natural transformation frequency (3). One of the mutants, with a mutation in a *virB11* homolog, was tested in the ferret diarrhea model and caused significantly less disease than 81-176 did (3). The presence of pVir genes in a subset of fresh clinical isolates (3) and the absence of this plasmid from other strains of *C. jejuni*, notably the genomic strain, NCTC 11168, led us to speculate that there may be differences in the pathogenic mechanisms among strains of *C. jejuni.* Herein, we report on the complete DNA sequence of the 37.5-kb pVir plasmid of *C. jejuni* 81-176. This sequence analysis reveals the presence of additional genes encoding putative components of type IV secretion systems, as well as orthologs of genes found in the hypervariable plasticity zones of *Helicobacter pylori* (1). Preliminary mutational analyses have identified five other genes that contribute to the ability of 81-176 to invade intestinal epithelial cells in vitro.

MATERIALS AND METHODS

Strains, plasmids, and culture conditions. *C. jejuni* 81-176 was isolated from an outbreak of campylobacteriosis and shown to cause disease in human volunteer studies (5, 22; D. Tribble, unpublished). *C. jejuni* strains were grown on Mueller-Hinton (MH) medium under microaerobic conditions. Plasmid pBluescript was used as the cloning vector, and *Escherichia coli* DH5 α was the host for cloning experiments. *E. coli* strains were grown on Luria broth supplemented with $50 \mu g$ of ampicillin per ml or 20 μg of chloramphenicol per ml when appropriate.

^{*} Corresponding author. Mailing address: Enteric Diseases Program, Naval Medical Research Center, 503 Robert Grant Ave., Silver Spring, MD 20910-7500. Phone: (301) 319-7662. Fax: (301) 319-7679. E-mail: guerryp@nmrc.navy.mil.

Cloning and sequence analyses. Plasmid DNAs from *C. jejuni* were isolated using mini-Qiagen columns (Qiagen, Chatsworth, Calif.) as previously described (3). Restriction enzymes were purchased from New England Biolabs (Beverly, Mass.) and used as recommended by the supplier. Plasmid DNAs for sequence analyses were isolated using QIAprep spin miniprep columns (Qiagen). DNA sequencing was done using Big Dye sequencing kits (Perkin Elmer-Applied Biosystems, Foster City, Calif.) and an Applied Biosystems 373A or 3100 DNA sequencer. Sequencing of pVir inserts was performed by primer walking and by sequencing in both directions from within a chloramphenicol resistance cassette (41) following in vitro transposition, as described previously (15, 16). Custom DNA-sequencing primers were synthesized on an Applied Biosystems model 392 DNA synthesizer. The sequences of individual plasmid clones were assembled by either sequencing of overlapping cloned restriction fragments or PCR. The assembled sequence was analyzed and annotated manually. Open reading frames (ORFs) of greater than 30 residues were evaluated based on the presence of a suitable initiation codon with appropriate spacing to a ribosome-binding site as well as physical location to other ORFs. Similarity searches to the predicted proteins were performed using the BLAST-WU algorithm against all available public databases. Where appropriate, full-length proteins were aligned and analyzed using CLUSTALX.

Mutagenesis of pVir. Site-specific mutagenesis of pVir was performed using selected clones generated from the in vitro transposition outlined above. Plasmid DNAs containing transposon insertions were electroporated into *C. jejuni* 81-176 or NCTC 11168 as previously described with selection on MH agar containing 20 μ g of chloramphenicol per ml (15, 16). Transformants were analyzed by PCR using primers flanking the insertion point to confirm that a double crossover event had occurred. To determine if the Cm^r cassette exerted a polar effect on downstream genes, reverse transcriptase PCR (RT-PCR) was performed on the genes that are downstream of each of the insertionally inactivated genes in the five mutants that showed an invasion defect (see Table 2). RNA was isolated using RNAeasy kits (Qiagen). RT-PCR was done with the Prostar kit (Stratagene, La Jolla, Calif.) using Moloney murine leukemia virus RT as specified by the manufacturer. Denaturation was performed at 94°C and extension was performed at 72°C for 30s for all primers sets. The specific primers, annealing temperatures, and products were as follows. Expression of Cjp3 in the Cjp2 mutant was examined using primers 5-CGACAACAAAATAGGACAAAATC GC-3' and 5'-CGTAAGGCATTCCGTATCTTTCAAG-3' at an annealing temperature of 51°C; the expected product size was 170 bp. Expression of Cjp16 in the Cjp15 mutant was examined using primers 5'-GAGCAATTTAATGATGG AATTTGC-3' and 5'-CGTAAGGCATTCCGTATCTTTCAAG-3' at an annealing temperature of 48°C; the expected product size was 453 bp. Expression of Cjp30 in a Cjp29 mutant was examined using primers 5'-CAGCAGGTATT CCTGTAATCGAC-3' and 5'-GACTTTATCTGTGTATTTCGTGCC-3' at an annealing temperature of 50°C; the expected product size was 341 bp. Expression of Cjp33 in a Cjp32 mutant was done with primers 5-GGATAAGTGCAGTC GTAATTTTATTACTATGAG-3 and 5-GCTTCAACCGTAGCCTTATTCA TAG-3' at an annealing temperature of 49°C; the expected product size was 246 bp. Expression of Cjp48 in a Cjp49 mutant was done with primers 5-GTGCTT TGGCTTTAGGAGCTATGAC-3' and 5'-GTCTTTATCGCAATAAACACTC ACAAC-3' at an annealing temperature of 50°C; the expected product size was 317 bp. RT-PCR products were analyzed on 1.2% agarose. Postive controls were done using DNA as template and RNA from wild-type 81-176. The negative controls were the corresponding assays done without RT.

Invasion assays. Invasion assays were performed essentially as described previously (3, 4, 20, 29). The human embryonic intestinal (INT407) cell line obtained from the American Type Culture Collection was maintained in liquid nitrogen and cultivated in minimal essential medium with 10% heat-inactivated fetal calf serum (Gibco-BRL, Gaithersburg, Md.), 0.2 mM D-glutamine, and nonessential amino acids as suggested by the American Type Culture Collection. To a semiconfluent monolayer of about 10^5 INT407 cells per well of a 24-well plate, mid-log-phase bacteria at a multiplicity of infection of 20 were added, and the mixture was incubated for 2 h at 37°C under $5\%CO_{2}$ -95% air. The infected monolayer was washed three times with Hanks' balanced salt solution and incubated for another 2 h in fresh culture medium containing 100μ g of gentamicin per ml to kill extracellular bacteria. Subsequently, the infected monolayers were washed three times in Hanks' balanced salt solution and lysed with 0.1% Triton X-100 in phosphate-buffered saline for 15 min on an orbital shaker. Following serial dilution in phosphate-buffered saline, internalized bacteria were enumerated by plate count on MH agar cultured under microaerobic condition. All invasion assays were conducted in duplicate wells, and the data are presented as the experimental mean of three independent experiments \pm one standard deviation.

Motility. The motility of all mutants defective in invasion was determined in motility agar (MH broth $+$ 0.4% agar) as described previously (13, 14).

Statistical analyses. Percent invasion values of bacterial strains were compared using two-tailed *t* tests. Sample variance was assumed to be equal as determined by F-test analysis.

Nucleotide sequence accession number. The sequence of the pVir plasmid has been deposited in GenBank under accession no. AF472533.

RESULTS AND DISCUSSION

General description. The entire sequence of the circular pVir plasmid was determined to be 37,468 bp. Annotation of the sequence identified 54 ORFs likely to represent functional genes. A total of 83% of the plasmid represents coding sequence, which is significantly lower than for the densely packed *C. jejuni* genome, which, at 94% coding density, is one of the most efficient genomes sequenced to date. All but two of these ORFs, Cjp28 (ORFs are named for "*C. jejuni* plasmid") and Cjp49, were transcribed in the clockwise direction on the DNA plus strand. A linear map of plasmid ORFs is shown in Fig. 1, and the sequence and ORF annotation are shown in Table 1. The circular plasmid genome was arbitrarily broken at the translational start of *virB8/comB1* (3). A total of 35 ORFs encoded predicted proteins with no significant orthologs in the current sequence databases and to date appear to be *C. jejuni* specific. In addition, eight proteins had existing orthologs (two in *C. jejuni* NCTC 11168 and six in *H. pylori*) whose function is currently unknown.

The overall G+C content of the plasmid was 26% , which is lower than in either the NCTC 11168 chromosome (30.6%) (30) or plasmid sequences previously identified from *C. jejuni* and *Campylobacter coli*, which have $G + C$ levels of 33.5% (25) and 31.2 to 31.4% (GenBank accession no. X82079 and X82080), respectively. However, a small cryptic plasmid from *Campylobacter hyointestinalis* has a similarly low $G + C$ content of 28.3% (38). Plasmid sequences from the related species *H. pylori* have $G + C$ levels that range from 33 to 37% $(9, 24, 26)$ (GenBank accession no. AF056496 and AF275307, respectively).

There are two large noncoding regions present on the pVir plasmid, the 1,147-bp region between Cjp28 and Cjp29 and the 1,049-bp intergenic region between Cjp49 and Cjp50. The region between Cjp28 and Cjp29, which has a $G+C$ content of 19.8%, shows no homology to known DNA molecules. This stretch of DNA includes a complex region of direct repeats that is somewhat similar to the repeats that have been suggested as origins of replication of small cryptic campylobacter plasmids (25). The intergenic repetitive region spans 260 bp between Cjp28 and Cjp29 and is bracketed by perfect 52-bp direct repeats (5-CCTATTTTACCATTACTTTAAGCTTTT TATAAGTCGCAAAATGTCGGTATAA-3). The terminal 36 bases of the 52-mer are repeated once more within the 260-bp region, and the terminal 24 bases are repeated one more time within the same region. Interestingly, the 52-bp repeat is also found in the noncoding region between Cjp48 and Cjp49 (bp 31845 to 31896). The 36-bp part of the repeat is found again in the long noncoding intergenic region between Cjp49 and Cjp50 (bp 34216 to 34251), which has a similarly low G+C content of 18.3% . Thus, Cjp49 is bracketed by 36-bp direct repeats, an organization reminiscent of a transposable element (Fig. 1).

6244 BACON ET AL.

Gene	Name	Start position	End position	Length $(aa)^a$	% Identity/similarity to sequence with highest probability score (length [aa] of alignment)	Annotation
Cjp01	virB8	$\mathbf{1}$	678	225	100(225)	Type IV secretion/competence protein
Cjp02	virB9	678	1748	356	100(356)	Type IV secretion/competence protein
Cjp03	virB10	1745	2881	378	100(378)	Type IV secretion/competence protein
Cjp04		2892	3092	66	NA^b	Unknown
Cjp05	virB11	3085	4038	317	100(317)	Type IV secretion protein
Cjp06	virD4	4452	6338	628	36/55 (557)	Homolog of H. pylori VirD4 (traG conjugational transfer protein); best hit found in ATCC 55679
Cjp07		6457	6627	56	NA	Unknown
Cjp08		6702	7583	293	19/42 (269)	Homolog of <i>H. pylori</i> JHP0926 (unknown function)
Cjp09		7852	7998	48	NA	Unknown
Cjp10		8008	9156	382	21/37 (339)	Homolog of <i>H. pylori</i> JHP0937 (unknown function)
Cjp11	topA	9178	11145	655	43/60 (652)	Homolog of <i>H. pylori</i> topoisomerase 1 proteins (plasticity zone copies)
Cjp12		11352	11717	121	NA	Unknown
Cjp13	ssb	11733	12191	152	34/57 (122)	Single-stranded binding protein
Cjp14		12327	12500	57	NA	Unknown
Cjp15		12488	13123	211	NA	Unknown
Cjp16		13071	13565	164	NA.	Unknown
Cjp17		13590	13943	117	NA.	Unknown
Cjp18		13933	14058	41	NA.	Unknown
Cjp19		14067	14267	66	NA.	Unknown
Cjp20		14158	14703	181	NA	Unknown
Cjp21		14700	15113	137	NA.	Unknown
Cjp22		15110	15523	137	NA.	Unknown
Cjp23		15527	15850	107	$\rm NA$	Unknown
		15919	16152	77	NA.	Unknown
Cjp24 Cjp25		16294	16398	34	NA	Unknown
Cjp26		16642	17310	222	38/54 (209)	Homolog of <i>H. pylori</i> HP1000 and JHP0935 (plasmid partitioning protein)
Cjp27		17320	17949	209	NA	Unknown
Cjp28		18781c	17987	263^d	27/50 (225)	Homolog of plasmid replication protein
Cjp29	repA	19929	21050	373	NA	Unknown (glutamine-rich protein; see the text)
Cjp30		21047	21925	292	22/41 (233)	Unknown
Cjp31		22029	22811	260	32/49 (167)	Homolog of TrbM protein (plasmid RP4)
Cjp32		22842	23561	239	21/40 (236)	Homolog of <i>C. jejuni</i> CJ0041 (unknown function)
		23561	24814	417	28/48 (405)	Homolog of <i>H. pylori</i> HP0444 (unknown function)
Cjp33 Cjp34		24798	25226	142	NA	Unknown
				114		
Cjp35		25701	26045 26367	102	NA.	Unknown
Cjp36		26059 26377	26496	39	NA.	Unknown Unknown
Cjp37			27005		NA.	
Cjp38		26610 27078	27488	131	NA. NA	Unknown Unknown
Cjp39				136		
Cjp40		27485	27916	143	NA.	Unknown
Cjp41		28124	28309	61	NA	Unknown
Cjp42		28321	28491	56	NA	Unknown
Cjp43		28571	28699	42	NA	Unknown
Cjp44		29101	29466	121	NA	Unknown
Cjp45		29503	29838	111	NA	Unknown
Cjp46		30035	30397	120	NA.	Unknown
Cjp47		30630	31034	134	NA.	Unknown
Cjp48		31280	31672	130	90/94 (104)	Homolog of C. jejuni CJ1456c (unknown function)
Cjp49		33251^c	31899	450	27/47 (413)	Homolog of <i>H. pylori</i> HP0996 (JHP0942) and related paralogs
Cjp50		34301	34393	30	NA.	Unknown
Cjp51		34468	34605	45	NA	Unknown
Cjp52		34605	34847	80	32/51(78)	Homolog of <i>H. pylori</i> HP0442 (unknown function, paralogs of HP0016 and JHP0014)
Cjp53	virB4	34857	37325	822	33/54 (771)	Homolog of the <i>H. pylori</i> VirB4 proteins
Cip54	virB7	37325	37453	42	NA	Homolog of VirB7 proteins

TABLE 1. Predicted protein coding sequences on the pVir plasmid

^a aa, amino acids.

^b NA, Not applicable.

^c ORF is encoded by the complementary strand. *^d* pseudogene, contains in-frame termination codon.

All of the predicted ORFs on pVir had suitable ribosomebinding sites (RBS) upstream of their predicted initiation codons. Although the spacing varied slightly (between 6 and 11 bases), the consensus RBS was $AAGGA-N₇$ -initiation codon. Only two of the predicted ORFs possessed an alternative start codon, Cjp31 with GUG and Cjp28 with UUG, although Cjp28 contains an authentic mutation causing an in-frame termination codon early in the protein (Table 1). This frequency of AUG initiation codons (96.3%) is higher than that seen for the NCTC 11168 chromosome (86.2%). There are 16 putative proteins encoded by pVir that are smaller than 100 residues and a further 19 that are between 100 and 200 residues. This represents an above-average proportion of low-molecularweight proteins. However, the sequence quality of pVir was very high, and suitable RBSs preceded all these small proteins. The biological significance of the coding capacity of pVir being biased toward low-molecular-weight proteins is unknown.

The pVir proteins contain an amino acid bias that is driven by their low- $G+C$ coding DNA content. There are high levels of Lys (11.8%) and Phe (6.6%) residues in the proteins, with 85% and 91.2% of the Lys and Phe residues being encoded by the AAA and UUU triplets, respectively. As expected, there is a strong bias in the third wobble position of the coding triplet, with this position being filled by an A or U base 81.5% of the time.

Genes encoding orthologs of NCTC 11168 genes. There were two genes that encoded proteins with homology to those encoded by the chromosome of *C. jejuni* NCTC 11168 (30). Cjp32 encodes a predicted protein of 239 amino acids which showed 21% identity and 40% similarity to the Cj0041 product (598 amino acids) over 98.7% of the total length of the Cjp32 protein, as well as 22% identity and 43% similarity over 209 amino acids to the Cj0320/FliH protein (276 amino acids), a protein involved in flagellin export. Cjp48 encodes a predicted protein of 130 amino acids which showed 90% identity and 94% similarity to the Cj1456c protein (104 amino acids) over 80% of the total length of the Cjp48 protein. Neither the Cj0041 nor Cj1456c protein show any significant homology to any other known protein. There were two possible initiation codons for Cjp48, one that matched Cj1456c and one that would have extended the NH₂-terminal region of the protein by 14 residues. The internal Met residue was selected for Cjp48 due to the presence of a preferred RBS, as well as the fact that the homology to the NCTC 11168 chromosome deteriorated in the upstream residues including lacking the alternative Met residue. However, alignment of Cjp48 with the region in NCTC11168 demonstrated that the full-length homolog is present in the sequenced chromosome but is truncated due to a frameshift in a string of A residues (six in Cjp48 but only five in Cj1456c), suggesting that this gene is likely to be subject to slipped-strand regulation (30). Taking this frameshift into account, the identity of Cjp48 and the chromosomal ortholog increases to 92.3%. Further, the DNA homology to the NCTC chromosome extends for 66 bp downstream of the Cjp48 coding region.

Genes encoding type IV secretion system homologs. There are seven genes on pVir that encode homologs of type IV secretion proteins and are clustered in a region spanning 8.9 kb with an overall $G+C$ content of 29.1%. The type IV genes in this region include the four genes described above (*virB*8/ *comB1*, *virB9/comB2*, *virB10/comB3*, and *virB11*, ORFs Cjp1, Cjp2, Cjp3, and Cjp5, respectively). Orthologs of VirB8, VirB9, and VirB10 have been shown in other bacterial systems to encode pore-forming proteins, and VirB11 orthologs function as ATPases (6, 7). Just downstream of *virB11*, Cjp6 encodes an ortholog of VirD4 of *H. pylori* ATCC 55679 (36% identity and 55% similarity over 557 amino acids), another putative ATPase (6, 7). Cjp53 encodes an ortholog of VirB4 of *H. pylori* (33% identical and 54% similar over 842 amino acids), also a putative ATPase.

Located just upstream of *virB8/comB1* on the circular map, Cjp54 encodes a small protein $(M_r 4789)$ containing a lipobox (40). Although this protein showed no significant matches by BLASTP analyses, probably owing to its small size (42 amino acids), it showed 12% identity and 32% similarity to the lipoprotein VirB7 of *Agrobacterium tumefaciens* (11) by CLUSTAL analysis. In *A. tumefaciens*, a cysteine residue of VirB7 interacts with a cysteine of VirB9 to stabilize the complex of proteins that form the type IV transmembrane channel (2, 8, 35). Interestingly, Cjp54 contains three Cys residues across its 42-amino-acid length.

Additionally, there is one other gene encoding an ortholog of a type IV secretion system that is not linked to the others. This is Cjp31, which encodes a predicted protein with 32% identity and 49% similarity over 167 amino acids to *trbM* of plasmid RP4 (199 amino acids), a protein involved in conjugal transfer of this broad-host-range plasmid. The homology includes a stretch of 63 residues with 58.7% identity and 77.8% similarity to TrbM.

Genes encoding orthologs of *H. pylori* **genes.** The type IV secretion system encoded by pVir shows its greatest similarities to type IV secretion proteins found in *H. pylori.* The pVir plasmid also encodes seven additional proteins whose best (and often only) homology is to other proteins identified in *H. pylori* (Table 1; Fig. 1). The Cjp11 protein displays a high level of homology to DNA topoisomerase I proteins, involved in DNA replication. However, it is significant that the topoisomerase I proteins with the highest level of similarity ($\geq 60\%$) are the "strain-specific" paralogs that have been identified in the plasticity zones of both sequenced *H. pylori* isolates, which are present in addition to the orthologous genes located in the conserved backbone of the *H. pylori* chromosome (1, 36).

The remaining six pVir proteins with orthologs in *H. pylori* are the Cjp8, Cjp10, Cjp26, Cjp33, Cjp49, and Cjp52 proteins (Table 1). Importantly, all of these *H. pylori* orthologs are found in the strain-specific plasticity zones of either *H. pylori* J99 (Cjp8 and Cjp10), *H. pylori* 26695 (Cjp33 and Cjp52), or both strains (Cjp26 and Cjp49). The Cjp26 protein shows similarity to HP1000 in strain 26695 and to JHP0935 in strain J99, which are homologs of ParA (partition) proteins that affect the proper distribution of newly replicated plasmids to daughter cells. The proteins with homology to the Cjp49 protein in strains 26695 and J99 are HP0996 and JHP0942, respectively, which are proteins of unknown function. However, in strain 26695 there is a second paralog (HP1004), which is also located in the plasticity zone. Analysis of 43 *H. pylori* isolates from Costa Rica demonstrated that at least one ortholog of this gene family was present in 19 of the strains (28). The two *H. pylori* J99 orthologs that were also present on pVir (Cjp8/JHP0926 and Cjp10/JHP0937) were looked at for their presence in the

panel of 43 *H. pylori* isolated from Costa Rica (17 strains from cancer patients and 26 strains from gastritis patients) (28). Orthologs of JHP0926 (Cjp8) and JHP0937 (Cjp10) were found in 9 and 25 of the 43 strains, respectively (28).

It has been hypothesized that the variability of the *H. pylori* plasticity zone may be due, in part, to acquisition of sequences from self-replicating plasmids, since some plasticity zone genes have been identified on plasmids found in other *H. pylori* strains (1). Recently, proteins with high similarity to both chromosomal and plasticity zone-encoded proteins have also been found to be encoded by plasmids in certain *H. pylori* isolates (19), supporting the original theory that plasmid acquisition and genetic exchange plays a major role in the macrodiversity observed within the *H. pylori* population.

The presence of orthologs, from both plasticity zones, on a *C. jejuni* plasmid further strengthens this hypothesis and suggests the possibility that sequences may be exchanged between these two highly related species. Indeed, the presence of plasticity zone genes on a plasmid involved in pathogenesis of *C. jejuni* suggests that these regions in *H. pylori* may play a role in the varying level of disease caused by infection with different strains of this organism.

The plasticity zone is one of several regions in the *H. pylori* genome that is characterized by having a $G+C$ content (35%) significantly lower than the remainder of the genome (39%), which is consistent with this region being introduced by horizontal transfer from another species. The plasticity zone orthologs on the pVir plasmid range in their $G+C$ content from 25.3% (Cjp26) to 33.9% (Cjp10). Interestingly, the orthologs of the *H. pylori* J99 plasticity zone have significantly higher G+C content (Cjp8, 32.2%; Cjp10, 33.9%) than do the orthologs of the *H. pylori* 26695 plasticity zone (Cjp26, 25.3%; Cjp33, 26.6%; Cjp52, 28.8%). Whether the ancestral lineage of the pVir plasmid is married to the lineage of the *C. jejuni* species is unknown, but the low $G+C$ content makes it unlikely and the evidence reported here strongly suggests that the pVir plasmid of *C. jejuni* and the plasticity zone of *H. pylori* have their origins from an as yet undiscovered source. There is considerable evidence suggesting that the diversity of the plasticity zone(s) of *H. pylori* strains is influenced by plasmid sequences (1, 19) and that each strain contains its own unique complement of genes, even though some of them are in common (28). Furthermore, recent evidence has demonstrated that *H. pylori* isolates that were taken from the J99 patient 6 years after the initial isolation of J99 and shown by sequence analysis of several regions totaling over 4 kb to be 100% identical to J99 have undergone several significant changes in the content of their plasticity zones (21). While some colonies had lost large regions of the plasticity zone, many had actually acquired DNA previously seen only in the plasticity zone of *H. pylori* 26695 (21). This confirmed that this region is hypervariable, and the presence of some of these orthologs on a plasmid in *C. jejuni* strongly suggests that both species have acquired DNA from a common ancestor or that they have exchanged DNA in the past.

Genes encoding other proteins of known function. There are only two other genes on pVir that encode proteins with similarity to proteins of known function. These are Cjp13, whose product shows similarity to single-stranded DNA-binding proteins, and Cjp28, which, although its product shows similarity to plasmid replication proteins from a number of bacterial species, contains an authentic in-frame stop codon and appears to be a pseudogene, suggesting that Cjp28 is not involved in plasmid replication. Interestingly, the only other gene encoding a protein predicted to be involved in plasmid replication was Cjp26 (see above). However, as shown below, a plasmid with a mutation in Cjp26 was stably maintained. Thus, the sequence does not reveal any clear information about plasmid replication or maintenance. The Cjp26 and Cjp28 proteins are the only pVir products that show any similarity to proteins encoded by the two large *H. pylori* plasmids, pHel4 and pHel5, that were recently sequenced (19).

Mutational analyses. Preliminary mutational analyses were done using selected transposon insertions generated during sequence analyses. These insertions were electroporated into 81-176 as described previously (15, 16), and putative mutants were confirmed by PCR analysis as described in Materials and Methods. These mutants, all of which were fully motile, were examined phenotypically in terms of invasion of INT407 cells to screen for potential virulence genes. We have previously described the effect of mutations in *virB11* and *virB10/comB3* on invasion (3). Mutations were characterized in four additional genes encoding putative components of a type IV secretion system, *virB8/comB1*, *virB9/comB2*, *virD4*, and *virB4.* Of these, only the mutant in *virB9/comB2* showed a significant reduction in invasion (Table 2). It was surprising that mutation of some, but not all, of the genes encoding the components of the type IV system affected invasion in vitro. This observation may reflect, in part, the considerable variability known to exist among type IV secretion systems in diverse bacteria. Thus, the type IV secretion systems of *H. pylori* and *Rickettsia prowazekii* contain fewer components than does the prototypic type IV system of *A. tumefaciens* (6, 7). Moreover, Selbach et al. recently were able to distinguish VirD4-dependent and VirD4 independent functions of the *H. pylori* type IV secretion system (34). These workers have shown that while VirD4 is necessary for secretion of CagA through the type IV secretion machinery, induction of IL-8 release from gastric epithelial cells requires other components of the type IV system but not VirD4 (34).

Mutants with mutations in 14 other genes were screened for invasion defects. Nine of these are mutants with mutations in genes that appear to be unique to the campylobacter plasmid (ORFs 12, 15, 16, 23, 24, 29, 30, 44, and 51), four had mutations in orthologs of *H. pylori* genes (Cjp8, Cjp11, Cjp26, and Cjp49), and the last had a mutation in Cjp32, which is homologous to Cj0041 and FliH of NCTC 11168 (see above). Of the 14 mutants, 4 showed significant reductions in invasion. The mutant with a mutation in Cjp15, encoding a predicted soluble protein of 25.8 kDa, invaded at 53.6% of the level of wild-type 81-176 ($P < 0.01$). The Cjp32 mutant invaded at 34.8% of the level of wild-type 81-176 ($P < 0.001$). The mutant with a mutation in gene Cjp49, encoding the ortholog of HP0996, invaded at 26% the level of wild-type 81-176 ($P < 0.0001$). Intriguingly, the Cjp49 gene is bracketed by a perfect pair of 36-bp direct repeats (see above), which may render this gene unstable via intragenomic recombination. The fourth mutant affected in invasion had a mutation in the Cjp29 gene, which encodes a glutamine-rich protein of 44.8 kDa, which showed no significant match by BLASTP analyses. The mutant with a

TABLE 2. Mutational analyses of pVir

^{*a*} The insertion point of each transposon insertion was determined by DNA sequence analyses using primers within the transposon (13, 14); a plus sign indicates that the *cat* gene was inserted in the same orientation a ^b Actual invasion of wild-type 81-176 was 2.43% \pm 0.58% of the inoculum. Values are given relative to 81-176, which was set at 100%.

^c P < 0.0001.

^d P < 0.001.

^e P < 0.01.

mutation in Cjp29 invaded at 15.2% of the level of wild-type 81-176 ($P < 0.0001$). There was no significant difference in the growth curves of the mutants that showed invasion defects compared to 81-176 in MH broth (data not shown). Additionally, RT-PCR was done on downstream genes, as described in Materials and Methods, to confirm that the Cm^r cassette was not exerting a polar effect on downstream genes. In all cases, the gene downstream of the mutated gene was expressed (data not shown).

pVir alone is insufficient for invasion. A mutant of 81-176 insertionally inactivated in Cjp11, encoding a putative topoisomerase, invaded at wild-type levels (Tables 2 and 3). Plasmid DNA (pVir/Cjp11::Cm) was isolated from this mutant and used to electroporate NCTC 11168 to Cm^r. Comparison of the restriction enzyme patterns of pVir/Cjp11::Cm isolated from an NCTC 11168 recipient and that from the 81-176 parent showed no differences (data not shown), indicating that the intact plasmid had been transferred into the recipient. Plasmid pVir/Cjp11::Cm appeared to replicate stably in NCTC 11168,

TABLE 3. Effect of pVir on invasion of INT407 cells by *C. jejuni* NCTC 11168

Strain	% Invasion ^a

^a Invasion is expressed as the percentage of the input inoculum internalized in 2 h.

and yields of plasmid DNA were comparable to yields from 81-176, suggesting similar copy numbers in the two strains. However, the presence of the plasmid in NCTC 11168 did not alter its low level of invasion into INT407 cells, as shown in Table 3. This indicates that although the plasmid is required for maximum invasion by 81-176, additional virulence factors not found in NCTC 11168 are also necessary. This is consistent with previous reports about the requirement of both the Penner 23,36 capsule and particular ganglioside mimicries in LOS for invasion of 81-176 into INT407 cells (4, 16). We have previously shown that a strain of 81-176 that has spontaneously lost pTet, the second large plasmid in this strain, is fully invasive (3), but the possibility remains that additional, novel chromosomal genes of 81-176 may contribute to invasion. The pVir plasmid appears remarkably stable in 81-176, and we have been unable, so far, to cure tagged versions (data not shown).

Conclusions. The pVir plasmid contributes to the ability of 81-176 to invade INT407 cells in vitro, a marker that correlates with virulence in the ferret diarrheal disease model (3, 4, 42). Plasmid mutants show reductions of 2- to 11-fold in invasion of INT407 cells. The greatest reduction in invasion was that seen in the *virB11* mutant, which invaded at 9% of the level of the wild type (3) . The new mutants described here display about two- to sevenfold reductions in invasion compared to the parent. This range in invasion defects is difficult to interpret but may be due, at least in part, to redundancy of some of these genes. Although apparently dispensable to invasion in wildtype 81-176, there is an additional type IV secretion system on the pTet plasmid (R. A. Batchelor and P. Guerry, unpublished

data) that may be able to complement partially some pVir mutants. Similarly, Cjp32 is similar to Cj0041, a gene shown in a random transposon mutagenesis study to be required for motility and invasion (10). This observation, coupled with the similarity of the Cjp32 protein to FliH, a component of the flagellar type III secretion system, is also intriguing. The pVir plasmid encodes a number of genes that are closely related to potential virulence genes of *H. pylori*, including both the type IV secretion system genes and plasticity zone genes. Although the mutants need additional characterization, preliminary analyses indicate that four additional pVir genes, Cjp15, Cjp29, Cjp32, and Cjp49, significantly affected *C. jejuni* invasion into intestinal epithelial cells. All four genes encode proteins that are predicted to be soluble and located in the cytoplasm. It is possible that one or more of these are effector proteins of this putative type IV secretion system, a hypothesis that is currently being tested. It is particularly interesting that Cjp49, which encodes an ortholog of *H. pylori* plasticity zone protein, appears to affect the virulence of *C. jejuni*, and we are also examining the possible role of the plasmid in induction of proinflammatory cytokines. Additionally, we are exploring the apparent contribution of the putative type IV secretion system to natural transformation (3). The data presented here provide a framework for further analyses of the role of these *C. jejuni* plasmid genes in the pathogenesis of this important human pathogen.

ACKNOWLEDGMENTS

This work was supported by work units 601152N.00004.001.A0072 from the Office of Naval Research and 61102A.S13.0.A0005 from the Naval Medical Research Center.

We thank Gary Majam and Isabelle Walker for technical assistance and Scarlett Goon for advice with the RT-PCR experiments.

REFERENCES

- 1. **Alm, R. A., L.-S. L. Ling, D. Moir, B. L. King, E. D. Brown, P. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jian, D. E. Taylor, G. F. Vovis, and T. J. Trust.** 1999. Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori.* Nature **397:**160–189.
- 2. **Anderson, L. B., A. V. Hertzel, and A. Das.** 1996. *Agrobacterium tumefaciens* VirB7 and VirB9 form a disulfide-linked protein complex. Proc. Natl. Acad. Sci. USA **93:**8889–8894.
- 3. **Bacon, D. J., R. A. Alm, D. H. Burr, L. Hu, D. J. Kopecko, C. P. Ewing, T. J. Trust, and P. Guerry.** 2000. Involvement of a plasmid in the virulence of *Campylobacter jejuni* 81-176. Infect. Immun. **68:**4384–4390.
- 4. **Bacon, D. J., C. M. Szymanski, D. H. Burr, R. P. Silver, R. A. Alm, and P. Guerry.** 2001. A phase variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. Mol. Microbiol. **40:**769–778.
- 5. **Black, R. E., M. M. Levine, M. L. Clements, T. P. Hughes, and M. J. Blaser.** 1988. Experimental *Campylobacter jejuni* infection in humans. J. Infect. Dis. **157:**472–479.
- 6. **Christie P. J., and J. P. Vogel.** 2000. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol. **8:**354–360.
- 7. **Covacci, A., J. L. Telford, G. Del Giudice, J. Parsonnet, and R. Rappuoli.** 1999. *Helicobacter pylori* virulence and genetic geography. Science **284:**1328– 1333.
- 8. **Das, A., and Y. H. Xie.** 2000. The *Agrobacterium* T-DNA transport proteins VirB8, VirB9 and VirB10 interact with one another. J. Bacteriol. **182:**758– 763.
- 9. **De Ungria, M. C, T. Kolesnikow, P. T. Cox, and A. Lee.** 1999. Molecular characterization and interstrain variability of pHPS1, a plasmid isolated from the Sydney strain (SS1) of *Helicobacter pylori.* Plasmid **41:**97–109.
- 10. **Golden, N. J., and D. W. K. Acheson.** 2002. Identification of motility and autoagglutination *Campylobacter jejuni* mutants by random transposon mutagenesis. Infect. Immun. **70:**1761–1771.
- 11. **Goodner, B., G. Hinkle, S. Gattung, N. Miller, M. Blanchard, B. Qurollo, B. S. Goldman, Y. Cao, M. Askenazi, C. Halling, L. Mullin, K. Houmiel, J.**

Gordon, M. Vaudin, O. Iartchouk, A. Epp, F. Liu, C. Wollam, M. Allinger, D. Doughty, C. Scott, C. Lappas, B. Markelz, C. Flanagan, C. Crowell, J. Gurson, C. Lomo, C. Sear, G. Strub, C. Cielo, and S. Slater. 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science **294:**2323–2328.

- 12. **Grant, C. C. R., M. E. Konkel, W. Cieplak, Jr., and L. S. Tompkins.** 1993. Role of flagella in adherence, colonization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures. Infect. Immun. **61:**1764–1771.
- 13. **Guerry, P., R. A. Alm, M. E. Power, and T. J. Trust.** 1992. Molecular and structural analysis of *Campylobacter* flagellin, p. 267–281. *In* I. Nachamkin, M. J. Blaser, and L. S. Tompkins (ed.), *Campylobacter jejuni*: current status and future trends. American Society for Microbiology, Washington, D.C.
- 14. **Guerry, P., R. A. Alm, M. E. Power, S. M. Logan, and T. J. Trust.** 1991. The role of two flagellin genes in *Campylobacter* motility. J. Bacteriol. **173:**4757– 4764.
- 15. **Guerry, P., C. P. Ewing, T. E. Hickey., M. M. Prendergast, and A. P. Moran.** 2000. Sialylation of lipooligosaccharide cores affects immunogenicity and serum resistance of *Campylobacter jejuni.* Infect. Immun. **68:**6656–6662.
- 16. **Guerry, P., C. M. Szymanski, M. M. Prendergast, T. E. Hickey, C. P. Ewing, D. L. Pattarini, and A. P. Moran.** 2002. Phase variation of *Campylobacter jejuni* 81-176 lipooligosaccharide affects ganglioside mimicry and invasiveness in vitro. Infect. Immun. **70:**787–793.
- 17. **Hickey, T. E., A. McVeigh, D. A. Scott, R. E. Michielutti, A. Bixby, S. A. Carroll, A. L. Bourgeois, and P. Guerry.** 2000. *Campylobacter jejuni* cytolethal distending toxin mediates release of interleukin-8 from intestinal epithelial cells. Infect. Immun. **68:**6535–6541.
- 18. **Hickey, T. E., S. Baqar, A. L. Bourgeois, C. P. Ewing, and P. Guerry.** 1999. *Campylobacter jejuni*-stimulated secretion of interleukin 8 by intestinal epithelial cells. Infect. Immun. **67:**88–93.
- 19. **Hofreuter, D., and R. Haas.** 2002. Characterization of two cryptic *Helicobacter pylori* plasmids: a putative source for horizontal gene transfer and gene shuffling. J. Bacteriol. **184:**2755–2766.
- 20. **Hu, L., and D. J. Kopecko.** 1999. *Campylobacter jejuni* 81-176 associates with microtubules and dynein during invasion of human intestinal cells. Infect. Immun. **67:**4171–4182.
- 21. **Israel, D. A., N. Salama, U. Krishna, U. M. Rieger, J. C. Atherton, S. Falkow, and R. M. Peek, Jr.** 2001. *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. Proc. Natl. Acad. Sci. USA **98:**14625– 14630.
- 22. **Korlath, J. A., M. T. Osterholm, L. A. Judy, J. C. Forfang, and R. A. Robinson.** 1985. A point-source outbreak of campylobacteriosis associated with consumption of raw milk. J. Infect. Dis. **152:**592–596.
- 23. **Lara-Tejero, M., and J. E. Galan.** 2001. CdtA, CdtB and CdtC form a tripartite complex thatis required for cytolethal distending toxin activity. Infect. Immun. **69:**4358–4365.
- 24. **Lee, W.-K., Y.-S. An, K.-H. Kim, S.-H. Kim, J.-Y. Song, B.-D. Ryu, Y.-J. Choi, Y.-H. Yoon, S.-C. Baik, K.-H. Rhee, and M.-J. Cho.** 1997. Construction of a *Helicobacter pylori-Escherichia coli* shuttle vector for gene transfer in *Helicobacter pylori.* Appl. Environ. Microbiol. **63:**4866–4871.
- 25. **Luo, N., and Q. Zhang.** 2001. Molecular characterization of a cryptic plasmid from *Campylobacter jejuni.* Plasmid **45:**127–133.
- 26. **Minnis, J. A., T. E. Taylor, J. E. Knesek, W. L. Peterson, and S. A. McIntire.** 1995. Characterization of a 3.5-kbp plasmid from *Helicobacter pylori.* Plasmid **34:**22–36.
- 27. **Oberhelman, R. A., and D. N. Taylor.** 2000. *Campylobacter* infections in developing countries, p. 139–153. *In* I. Nachamkin and M. J. Blaser (ed.), *Campylobacter.* American Society for Microbiology, Washington, D.C.
- 28. **Occhialini, A., A. Marais, R. Alm, F. Garcia, R. Sierra, and F. Megraud.** 2000. Distribution of open reading frames of plasticity region of strain J99 in *Helicobacter pylori* strains isolated from gastric carcinoma and gastritis patients in Costa Rica. Infect. Immun. **68:**6240–6249.
- 29. **Oelschlaeger, T. A., P. Guerry, and D. J. Kopecko.** 1993. Unusual microtubule-dependent endocytosis mechanisms triggered by *Campylobacter jejuni* and *Citrobacter freundii.* Proc. Natl. Acad. Sci. USA **90:**6884–6888.
- 30. **Parkhill, J., B. W. Wren, K. Mungall, J. M. Ketley, C. Churcher, D. Basham, T. Chillingworth, R. M. Davies, T. Feltwell, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Moule, M. J. Pallen, C. W. Penn, M. A. Quail, M. A. Rajandream, K. M. Rutherford, A. H. M. van Vliet, S. Whitehead, and B. G. Barrell.** 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature **403:**665–668.
- 31. **Pavlovskis, O. R., D. M. Rollins, R. L. Haberberger, Jr., A. E. Green, L. Habash, S. Stroko, and R. I. Walker.** 1991. Significance of flagella in colonization resistance of rabbits immunized with *Campylobacter* spp. Infect. Immun. **59:**2259–2264.
- 32. **Pickett, C. L., E. C. Pesci, D. L. Cottle, G. Russell, A. N. Erdem, and H. Zeytin.** 1996. Prevalence of cytolethal distending toxin production in *Campylobacter jejuni* and relatedness of *Campylobacter* spp. *cdtB* genes. Infect. Immun. **64:**2070–2078.
- 33. **Rivera-Amill, V., B. J. Kim, J. Seshu, and M. E. Konkel.** 2001. Secretion of the virulence-associated *Campylobacter* invasion antigens from *Campylobacter jejuni* requires a stimulatory signal. J. Infect. Dis. **183:**1607–1616.
- 34. **Selbach, M., S. Moese, T. F. Meyer, and S. Backert.** 2002. Functional analysis of the *Helicobacter pylori cag* pathogenicity island reveals both VirD4-CagA dependent and VirD4-CagA-independent mechanisms. Infect. Immun. **70:** 665–671.
- 35. **Spudich, G. M., D. Fernandez, X.-R. Zhou and P. J. Christie.** 1996. Intermolecular disulfide bonds stabilize VirB7 homodimers and VirB7/VirB9 heterodimers during biogenesis of the *Agrobacterium tumefaciens* T-complex transport apparatus. Proc. Natl. Acad. Sci. USA **93:**7512–7517.
- 36. **Tomb, J. F., O. White, A. R. Kerlavage, R. A. Clayton, G. G. Sutton, R. D. Fleischmann., K. A. Ketchum, H. P. Klenk, S. Gill, B. A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E. F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H. G. Khalak, A. Glodek, K. McKenney, L. M. Fitzgerald, N. Lee, M. D. Adams, E. K. Hickey, D. E. Berg, J. D. Gocayne, T. R. Utterback, J. D. Peterson, J. M. Kelley, M. D. Cotton, J. M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W. S. Hayes, M. Borodovsky, P. D. Karp, H. O. Smith, C. M. Fraser, and J. C. Venter.** 1997. The complete genome of the gastric pathogen *Helicobacter pylori.* Nature **388:**539–547.
- 37. **Wassenaar, T. M., N. M. C. Bleumink-Pluym, and B. A. M. van der Zeijst.** 1991. Inactivation of *Campylobacter jejuni* flagellin genes by homologous

Editor: V. J. DiRita

recombination demonstrates that *flaA* but not *flaB* is required for invasion. EMBO J. **10:**2055–2061.

- 38. **Waterman, S. R., J. A. Hackett, and P. A. Manning.** 1993. Characterization of the replicative region of the small cryptic plasmid of *Campylobacter hyointestinalis.* Gene **125:**11–17.
- 39. **Whitehouse, C. A., P. B. Balbo, E. C. Pesci, D. L. Cottle, P. M. Mirabito, and C. L. Pickett.** 1998. *Campylobacter jejuni* cytolethal distending toxin causes a G2-phase cell cycle block. Infect. Immun. **66:**1934–1940.
- 40. **Yamaguchi, K., F. Wu, and M. Inouye.** 1988. A single amino acid determinant of the membrane localization of lipoproteins in *Escherichia coli.* Cell **53:**423–432.
- 41. **Yao, R., R. A. Alm, T. J. Trust, and P. Guerry.** 1993. Construction of new *Campylobacter* cloning vectors and a new mutational *cat* cassette. Gene **130:**127–130.
- 42. **Yao, R., D. H. Burr, P. Doig, T. J. Trust, H. Niu, and P. Guerry.** 1994. Isolation of motile and non-motile insertional mutants of *Campylobacter jejuni:* role of motility in adherence and invasion of eukaryotic cells. Mol. Microbiol. **14:**883–893.