

The Complete Genome Sequence and Analysis of the Human Pathogen *Campylobacter lari*

William G. Miller,¹ Guilin Wang,¹ Tim T. Binnewies,² and Craig T. Parker¹

Abstract

Campylobacter lari is a member of the epsilon subdivision of the Proteobacteria and is part of the thermotolerant *Campylobacter* group, a clade that includes the human pathogen *C. jejuni*. Here we present the complete genome sequence of the human clinical isolate, *C. lari* RM2100. The genome of strain RM2100 is ~ 1.53 Mb and includes the 46 kb megaplasmid pCL2100. Also present within the strain RM2100 genome is a 36 kb putative prophage, termed CLIE1, which is similar to CJIE4, a putative prophage present within the *C. jejuni* RM1221 genome. Nearly all (90%) of the gene content in strain RM2100 is similar to genes present in the genomes of other characterized thermotolerant campylobacters. However, several genes involved in amino acid biosynthesis and energy metabolism, identified previously in other *Campylobacter* genomes, are absent from the *C. lari* RM2100 genome. Therefore, *C. lari* RM2100 is predicted to be multiply auxotrophic, unable to synthesize eight different amino acids, acetyl-coA, and pantothenate. Additionally, strain RM2100 does not contain a complete TCA cycle and is missing the CydAB terminal oxidase of the respiratory chain. Defects in the amino acid biosynthetic pathways in this organism could be potentially compensated by the large number of encoded peptidases. Nevertheless, the apparent absence of certain key enzymatic functions in strain RM2100 would be expected to have an impact on *C. lari* biology. It is also possible that the reduction in the *C. lari* metabolic machinery is related to its environmental range and host preference.

Introduction

CAMPYLOBACTER *LARI* is a member of the thermotolerant clade of the Gram-negative Epsilonproteobacteria. Other members of the phylogenetically related, thermotolerant *Campylobacter* group include *C. jejuni*, *C. coli*, *C. upsaliensis*, *C. helveticus*, and *C. insulaenigrae*. The thermotolerant campylobacters are characterized by the ability of strains within each species to grow at 42°C; optimal growth of these species is supported under microaerobic conditions where the [O₂] is ≤ 5%.

Most strains of *C. lari* (formerly *C. laridis*) were isolated originally from gulls (Skirrow and Benjamin, 1980; Benjamin *et al.*, 1983). However, *C. lari* strains have also been isolated from cattle (Giacoboni *et al.*, 1993; Aarestrup *et al.*, 1997), dogs (Engvall *et al.*, 2003), pigs (Harvey *et al.*, 1999; Lindblom *et al.*, 1990), poultry (Tresierra-Ayala *et al.*, 1994), and shellfish (Endtz *et al.*, 1997; Van Doorn *et al.*, 1998), as well as water sources (Obiri-Danso and Jones, 1999a; Obiri-Danso *et al.*, 2001). Although associated with birds and food animals, *C. lari* is isolated infrequently from animal or processed food sources.

¹Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, California.

²Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark.

However, *C. lari* is detected at moderate to high levels in shellfish; in a study by Endtz *et al.* (1997), nearly all of the *Campylobacter* strains isolated from shellfish (37/39) were identified as *C. lari*. The presence of relatively high levels of *C. lari* in both seawater and shorebirds (Glunder and Petermann, 1989; Obiri-Danso and Jones, 1999b; Obiri-Danso *et al.*, 2001; Moore *et al.*, 2002; Waldenstrom *et al.*, 2007) suggests that the preferred environmental niche for this organism is the marine shoreline, and would presumably explain the prevalence of *C. lari* in shellfish. Also consistent with a marine environmental range, *C. lari* is more halotolerant than other thermotolerant campylobacters (Skirrow and Benjamin, 1980; Smibert, 1984). Additionally, the phylogenetically related species *C. insulaenigrae* is isolated predominantly from marine mammals (Stoddard, 2005; Stoddard *et al.*, 2007); thus, *C. lari* and *C. insulaenigrae* may be members of a group of marine-adapted campylobacters.

Unlike *C. jejuni*, *C. lari* is isolated relatively infrequently from clinical samples and only one outbreak has been attributed to this organism (Broczyk *et al.*, 1987). Nevertheless, the clinical symptomatology of *C. lari*-related campylobacterioses is similar to that of *C. jejuni*, namely gastroenteritis with abdominal pain, fever, and diarrhea (Broczyk *et al.*, 1987; Lin *et al.*, 1998; Prasad *et al.*, 2001; Otasevic *et al.*, 2004). *Campylobacter lari* has been associated also with bacteremia in immunocompromised (Nachamkin *et al.*, 1984; Martinot *et al.*, 2001) or otherwise debilitated patients (Morris *et al.*, 1998) and in patients with gastroenteritis (Soderstrom *et al.*, 1991; Skirrow *et al.*, 1993).

Campylobacter lari strains are phenotypically and genotypically diverse and have been subdivided into four major phenotypic groups (Endtz *et al.*, 1997; On and Harrington, 2000; Duim *et al.*, 2004): the nalidixic acid-resistant thermophilic campylobacters (NARTC), the urease-producing thermophilic campylobacters (UPTC), the nalidixic acid-susceptible (NASC) group, and the urease-producing nalidixic acid-susceptible group. *Campylobacter lari* RM2100 (ATCC-BAA 1060; CDC strain D67, "case 6" [Tauxe *et al.*, 1985]) is a member of the NARTC *C. lari* subgroup and was isolated from an 8-month-old girl with watery diarrhea. Strain

RM2100 (multilocus sequence typing [MLST] ST-3 [Miller *et al.*, 2005a]) is also very closely related phylogenetically to the *C. lari* type strain CCUG 23947 (NARTC; MLST ST-4 [Miller *et al.*, 2005a]), which was isolated from a gull. Preliminary data on the draft RM2100 genome sequence was presented previously (Fouts *et al.*, 2005); however, many features present in this strain could not be addressed conclusively due to the incomplete nature of the genome sequence. *Campylobacter lari* RM2100 was sequenced initially because of both the clinical origin of this isolate and the classification of *C. lari* as an "emerging" *Campylobacter* species capable of causing disease in humans. Additionally, solution of the strain RM2100 genome would hopefully provide new insights into several aspects of *Campylobacter* biology, such as colonization-host adaptation, virulence, and synthesis of surface structures (e.g., capsule). This study presents the completed genomic sequence of the human clinical isolate RM2100. The genomic data revealed that this strain contains defects in multiple metabolic pathways that may pertain to *C. lari* biology.

Materials and Methods

Growth conditions and chemicals

Campylobacter lari RM2100 was cultured at 37°C on brain heart infusion agar (Becton Dickinson, Sparks, MD) amended with 5% (v/v) laked horse blood (Hema Resource & Supply, Aurora, OR). The incubation atmosphere was 5% H₂, 10% CO₂, and 85% N₂. Polymerase chain reaction (PCR) enzymes and reagents were purchased from New England Biolabs (Beverly, MA) or Epicentre (Madison, WI). All chemicals were purchased from Sigma-Aldrich Chemicals (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA). DNA sequencing chemicals and capillaries were purchased from Applied Biosystems (Foster City, CA).

Polymerase chain reactions

C. lari genomic DNA was prepared as described previously (Miller *et al.*, 2005b). Standard amplifications were performed on a Tetrad thermocycler (Bio-Rad, Hercules, CA) with the following settings: 30 seconds at 94°C; 30

seconds at 53°C; 2 minutes at 72°C (30 cycles). Each amplification mixture contained 50 ng genomic DNA, 1×PCR buffer (Epicentre), 1×PCR enhancer (Epicentre), 2.5 mM MgCl₂, 250 μM each dNTP, 50 pmol each primer, and 1 U polymerase (New England Biolabs). Amplicons were purified on a BioRobot 8000 Workstation (Qiagen, Santa Clarita, CA). Sequencing and PCR oligonucleotides were purchased from Qiagen.

DNA sequencing

Cycle sequencing reactions were performed on a 96-well Tetrad thermocycler (Bio-Rad) using the ABI PRISM BigDye terminator cycle sequencing kit (version 3.1) and standard protocols. Amplicon extension products were purified using DyeEx 96 well plates (Qiagen) according to the manufacturer's protocols. DNA sequencing was performed on an ABI PRISM 3130XL Genetic Analyzer (Applied Biosystems) using the POP-7 polymer and ABI PRISM Genetic Analyzer Data Collection and ABI PRISM Genetic Analyzer Sequencing Analysis software.

Closure of the *C. lari* RM2100 genome

ABI chromatograms from the *C. lari* RM2100 draft genome (Fouts *et al.*, 2005) were obtained from TIGR (The Institute for Genomic Research, Rockville, MD) and assembled using SeqMan (v 5.0, DNASTar, Madison, WI). All TIGR chromatograms in the assembly were then trimmed for quality at the 5' and 3' ends. Those reads that did not meet minimum quality requirements were discarded. Six contig gaps and multiple sequence gaps were present after assembly and trimming of the draft genome traces. Arrangement of the six contigs into a final single contig used combinatorial standard PCR to bridge the contig gaps. All gaps, both sequence and contig, present in the assembly were closed by PCR using primers designed with Primer Premier (v 5.0, Premier Biosoft, Palo Alto, CA). The three ~6 kb ribosomal RNA (rRNA) loci present in *C. lari* RM2100 were amplified independently using unique flanking primer sets and long PCR. rRNA region amplicons were sequenced, using rRNA primers common to all three loci, and assembled. The assembled rRNA loci were then integrated into the larger genome assembly

manually. The final assembly contained 29,574 reads. The final assembly also contained contiguous sequences (>2×coverage/nt) on both strands for an average coverage of 9.6×; ambiguous bases were resequenced at least twice.

Genome analysis

Putative coding sequences were predicted using ORFscan (Rational Genomics, South San Francisco, CA) with the parameters set to include all three potential start codons (ATG, GTG, and TTG) and a coding sequence (CDS) cut-off of 50 amino acids. Spurious CDSs (e.g., CDSs contained within larger CDSs) were deleted manually. Initial annotation was accomplished by comparing the predicted proteins to the NCBI nonredundant (nr) database using BLASTP. The list of putative CDSs was then used to create a preliminary GenBank-formatted (.gbk) file that was entered into Artemis (release 9.0; <http://www.sanger.ac.uk/Software/Artemis/> [Rutherford *et al.*, 2000]). Annotation within Artemis included the fusion of split CDSs into pseudogenes and the identification of small (usually ribosomal protein) genes overlooked in the initial ORFscan analysis. The start codon of each putative CDS was curated manually, either through visual inspection within Artemis of the ribosome binding site, or through BLAST comparison of each CDS to its epsilonproteobacterial homologs, where present; CDSs with dubious ribosome binding sites, lengths shorter than 200 bp, and a low BLASTP score ($E > 10^{-10}$) were removed from the final list of CDSs. tRNAs were annotated using tRNAscan-SE (v 1.23; <http://bioweb.pasteur.fr/seqanal/interfaces/trnscan.html> [Lowe and Eddy, 1997]). rRNA loci were identified through BLASTN comparison to *Campylobacter* rRNA loci.

Final annotation of the strain RM2100 genome was performed using 1) BLASTP comparison to proteins in the NCBI nr database, 2) CDS identification, cluster of orthologous group function prediction, cellular location prediction, and identification of Pfam domains and PROSITE motifs using the BASys Bacterial Annotation System (<http://wishart.biology.ualberta.ca/basys/cgi/submit.pl>), and 3) additional Pfam domain identification using the Sanger Centre Pfam search engine (release 22.0;

pfam.sanger.ac.uk/). Four levels of gene assignment were used for each CDS: 1) CDSs predicted to encode proteins with high similarity across multiple taxa to proteins of known function (e.g., PurA, HisB, GyrA) were assigned that function, 2) CDSs predicted to encode proteins with only a general function, but similar to members of various Pfam families or superfamilies, or containing various PROSITE motifs, were annotated to reflect the tentative general function (e.g., acetyltransferase, oxidoreductase), 3) CDSs predicted to encode proteins with no defined function or motifs but similar to other nondefined proteins ($E < 10^{-10}$) across multiple bacterial taxa were designated as "conserved hypothetical," and 4) CDSs predicted to encode proteins with no defined function and either unique to this genome or with a low similarity ($E > 10^{-10}$) to proteins from other taxa were designated as "hypothetical." In some instances, two or more adjacent CDSs were determined by BLAST analysis to represent a single CDS, fragmented by frameshift or nonsense mutations. These fragmented CDSs were fused into a single open reading frame and annotated as a putative pseudogene. The complete nucleotide sequence and annotation of *C. lari* RM2100 has been deposited at GenBank under the accession numbers CP000932 (chromosome) and CP000933 (megaplasmid pCL2100).

Results and Discussion

General features

The genome of *Campylobacter lari* RM2100 contains 1,525,460 bp, and thus represents the smallest characterized epsilonproteobacterial genome to date. The G+C content of the RM2100 genome is 29.7% and is consistent with other members of the thermotolerant *Campylobacter* group (30.3–34.5% G+C). A summary of the features of the strain RM2100 genome is presented in Table 1 and a diagrammatic representation of the strain RM2100 genome is presented in Suppl. Fig. S1 (to view all supplementary materials, go to www.liebertpub.com/fpd). The *C. lari* RM2100 genome is predicted to contain 1500 CDSs with an additional 13 fragmented CDSs, containing frameshift mutations or other point mutations, designated as putative pseudogenes (Table 1). The 1500

TABLE 1. FEATURES OF THE *CAMPYLOBACTER LARI* RM2100 GENOME

General features	No. or % of total
Chromosome	
Size (bp)	1,525,460
G + C content	29.70%
CDS numbers ^a	1500
Assigned function	736 (49%)
Pseudogenes	13
General function	504 (34%)
Conserved	260 (17%)
hypothetical/hypothetical	
Prophage/genetic islands	1/0
Ribosomal RNA operons	3
IS elements/CRISPRs	0/0
Poly G:C tracts (hypervariable)	
6 bp	7 (0)
7 bp	0
8 bp	0
9 bp	8 (1)
10 bp	4 (3)
11 bp	3 (1)
Plasmid pCL2100	
Size (bp)	46,201
G + C content	26.86%
CDS numbers	42
Gene classes	
Chemotaxis proteins	23
Che/Mot proteins	9
Methyl-accepting chemotaxis proteins	12
Restriction/modification systems	
Type I	1 ^b
Type II/IIS	2
Type III	0
Transcriptional regulators	
Regulatory proteins	18
Non-ECF family σ factors	3
Two-component systems	
Response regulator	8
Sensor histidine kinase	6

^aTotal does not include pseudogenes.

^b*hsdM* contains frameshift; therefore, the *hsd* locus in strain RM2100 is presumably nonfunctional. CDS, coding sequence; IS, insertion sequence; CRISPRs, clustered regularly interspaced short palindromic repeats; ECF, extracytoplasmic function.

proteins predicted to be encoded by strain RM2100 were compared by BLASTP analysis to proteins contained in the NCBI nr database (September 24, 2007 release). Based on this analysis, approximately half of the proteins (736; 49%) were assigned a specific function, with the remainder determined to be proteins of either general (504; 34%) or unknown (260; 17%) function (Table 1). Thirty-one proteins (2.1%) were not similar to any of the proteins contained within the nr database, and an additional 19

proteins (1.3%) had alignment match lengths < 75% of the lengths of both the query and subject sequences. Genes encoding this latter group of 19 proteins are potentially gene fragments. A complete list of the *C. lari* RM2100 CDSs and their annotation is presented in Suppl. Table S1. Categorization of these CDSs by function is presented in Suppl. Table S2. The five best homologs, where applicable, for each protein encoded by strain RM2100 are presented in Suppl. Table S3.

Similar to other campylobacters, *C. lari* RM2100 contains homopolymeric G:C tracts, some of which are hypervariable (Table 1). Strain RM2100 contains 15 homopolymeric tracts of either 9, 10, or 11 base pairs. Of these, five were determined to be hypervariable, based on the identification during genomic assembly of overlapping tracts of different lengths. Within *Campylobacter*, homopolymeric G:C tracts are found often within genes involved in the biosynthesis of surface structures (Fouts *et al.*, 2005; Gilbert *et al.*, 2005; Parkhill *et al.*, 2000), e.g., the lipooligosaccharide (LOS), capsular polysaccharide (CPS), and flagellar modification loci. Likewise in the strain RM2100 genome, 11 of the 15 homopolymeric G:C tracts identified are located within these three loci: two of the tracts (both hypervariable) are located within the CPS locus (*cl0305* and *cl0307*), two of the tracts are located within the LOS locus (*cl1197* and *cl1199* [hypervariable]), and seven of the tracts are located in the putative flagellar modification locus (*cl1260*, *cl1266*, *cl1269*, *cl1273*, *cl1274*, *cl1276*, and *cl1281*). Determination of hypervariability in strain RM2100 was based on a level of coverage within the assembly sufficient to identify homopolymeric tracts of different lengths. It is possible that some of the “non-hypervariable” tracts were located within regions of coverage too low to adequately assess hypervariability. It is also possible that some tracts are not hypervariable under the conditions in which the *C. lari* cells were grown to harvest the genomic DNA. Additional investigations will be necessary to present a more accurate picture of homopolymeric tract hypervariability within strain RM2100.

Comparative genomics of C. lari RM2100

Campylobacter lari is a member of the phylogenetically related thermotolerant *Campylo-*

bacter group, which includes the species *C. jejuni*, *C. coli*, and *C. upsaliensis*. The genomic sequences of multiple strains within this group have already been determined, at least to draft level (Parkhill *et al.*, 2000; Fouts *et al.*, 2005; Hofreuter *et al.*, 2006; Pearson *et al.*, 2007; Fouts *et al.*, manuscript in preparation). Therefore, it is not surprising that a substantial majority (90%) of the non-unique *C. lari* RM2100 proteins have their best matches in proteins encoded by other thermotolerant campylobacters (Table 2 and Suppl. Table S3). The majority of *C. lari* proteins (53.5%) have their best matches in proteins encoded by either of the two *C. jejuni* subspecies with a further quarter similar to proteins encoded by *C. coli*. In fact, only 3.9% of the non-unique strain RM2100 proteins have their best matches in proteins encoded by non-*Campylobacter* taxa and only 2.6% in proteins encoded by non-epsilonproteobacterial taxa.

TABLE 2. SIMILARITY OF PREDICTED *CAMPYLOBACTER LARI* PROTEINS TO PROTEINS FROM OTHER TAXA

Taxon	Best matches		Within best 5 matches %
	n	%	
<i>C. jejuni</i> subsp. <i>jejuni</i>	677	46.6%	33.7%
<i>C. jejuni</i> subsp. <i>doylei</i>	99	6.8%	15.5%
<i>C. coli</i>	371	25.6%	19.1%
<i>C. upsaliensis</i>	161	11.1%	11.6%
<i>C. fetus</i>	35	2.4%	3.1%
<i>C. curvus</i>	25	1.7%	3.0%
<i>C. concisus</i>	17	1.2%	2.6%
<i>C. hominis</i>	7	0.5%	1.2%
<i>Helicobacter</i>	13	0.9%	1.0%
Other ϵ -proteobacteria	6	0.4%	1.8%
γ -proteobacteria	12	0.8%	3.0%
δ -proteobacteria	5	0.3%	0.5%
β -proteobacteria	0	0.0%	0.3%
α -proteobacteria	2	0.1%	0.3%
Bacteroidetes/Chlorobi	5	0.3%	1.1%
Firmicutes	11	0.8%	1.6%
Other bacteria	4	0.3%	0.6%

Predicted *C. lari* RM2100 proteins were compared to proteins in the NCBI nonredundant (nr) database by BLASTP. Matches with an Expect (E) value of $>1 \times 10^{-5}$, an identity of < 25%, and an alignment length across either the query or match sequences of < 75% were excluded. The best matches or the top five matches (where applicable) were identified at each locus (supplementary Table S3) and classified by taxon. Percentage values for the defined taxa do not include the 50 loci (3.3% of the 1500 identified loci) for which no homologs were identified or for which the alignment lengths were < 75%. For the top five matches, multiple strains of the same species (e.g., *C. jejuni* 81-176 and *C. jejuni* NCTC 11168) at a given locus were counted once.

Divergent loci within *C. lari* RM2100

As described above, nearly all of the *C. lari* RM2100 CDSs are well-conserved among *Campylobacter*, especially among thermotolerant campylobacters. However, some strain RM2100 CDSs encode proteins most similar to those of non-thermotolerant *Campylobacter* species or non-epsilonproteobacterial taxa, or encode proteins with no matches within the nr database, and are termed here divergent CDSs (Fig. 1). Many of these divergent CDSs are clustered and include genes encoding a MoxR/RavA AAA + ATPase/Von Willebrand Factor Type A protein pair (*cl0032* and *cl0033*); *nrdD* (anaerobic ribonucleotide triphosphate reductase; *cl0047*) and *nrdG* (NrdD-activation protein; *cl0048*); a region (*cl0256-cl0261*) bounded by genes encoding a

UspA-like universal stress protein and a LysR-family transcriptional regulator; a region (*cl0834-cl0854*) encoding a putative peptidyl-arginine deiminase, two AAA-family ATPases and a SIR2-family NAD-dependent protein deacetylase; and a region encoding two components (*ModA*; *cl1241* and *ModC*; *cl1243*) of a molybdenum ABC transporter (*ModB* is also divergent but located elsewhere on the chromosome [*cl0449*]). In some instances (e.g., *ModABC* and *NrdDG*) these divergent proteins are encoded by other *Campylobacter* species, but these proteins have low or no similarity to those encoded by strain RM2100. For example, with the exception of the thermotolerant campylobacters and *Helicobacter*, *nrdD* and *nrdG* have been identified in all characterized epsilonproteobacterial genomes (Miller *et al.*, 2007);

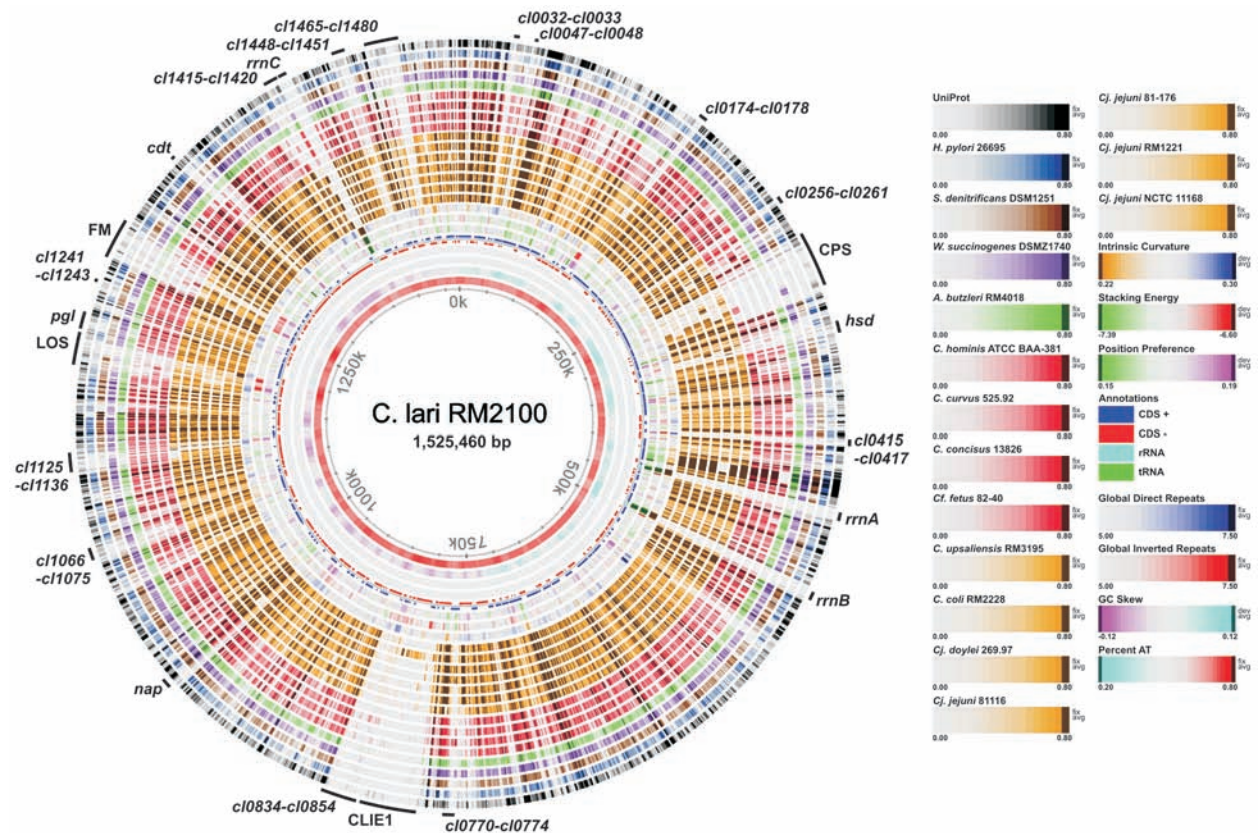


FIG. 1. Genomic BLAST atlas of *C. lari* RM2100. *Campylobacter lari* RM2100 is the reference genome and is compared to a set of 15 other epsilonproteobacterial genomes, including different thermotolerant *Campylobacter* (gold rings 9–15 from center), non-thermotolerant *Campylobacter* (red rings 16–19 from center), *A. butzleri* (green ring 20 from center), and *Helicobacteraceae* (purple, brown, and blue rings 21–23 from center) strains, as well as the UniProt database (outermost ring in black). Physical genomic parameters and annotations are illustrated in rings 1–8 from center. Outermost labels represented selected genes and loci. A web-based zoomable BLAST atlas is available at <http://www.cbs.dtu.dk/services/GenomeAtlas/suppl/zoomatlas/>. CPS, capsular polysaccharide; LOS, lipooligosaccharide; FM, flagellar modification.

however, strain RM2100 NrdD and NrdG are more similar to proteins encoded by *Clostridium* and *Desulfitobacterium*, and thus perhaps represent an independent acquisition by *C. lari*.

Biosynthetic pathways and respiration

The salient feature of the *C. lari* RM2100 genome is the predicted absence of several key metabolic proteins, including multiple enzymes involved in amino acid or cofactor biosynthesis and energy metabolism or respiration (Table 3). Genes not present in the strain RM2100 genome include those involved in the biosynthesis of acetyl-coenzyme A (*acs*), arginine (*argBCDFO*), glutamate (*gltBD*), leucine (*leuABCD*), methionine (*metABEF*), pantothenate (*panBCD*), proline (*proAB*), and tryptophan (*trpACDEF*). Additionally, strain RM2100 is not predicted to encode citrate synthase (*GltA*), aconitase (*Acn*), isocitrate dehydrogenase (*Icd*), succinyl-CoA synthetase (*SucCD*), succinate dehydrogenase

(*SdhABC*), and the *CydAB* terminal oxidase of the respiratory chain. Thus, *C. lari* RM2100 is multiply auxotrophic and does not contain a TCA cycle.

The absence of the glutamate synthase *GltBD* is apparently compensated for in strain RM2100 through the NADP-dependent amination of α -ketoglutarate via the alternate glutamate synthase *GdhA*, an enzyme not encoded by the other thermotolerant campylobacters. In addition to *C. lari*, the campylobacters *C. concisus*, *C. curvus*, and *C. hominis* are predicted also to have *gltBD*⁻, *gdhA*⁺ genotypes (Fouts *et al.*, manuscript in preparation). Since *GdhA* requires a cytoplasmic pool of NH₄⁺ and *gltBD* mutants demonstrate increased osmosensitivity and reduced growth in low ammonium media (Csonka *et al.*, 1994), the *gltBD*⁻, *gdhA*⁺ genotype of strain RM2100 would restrict presumably the environmental niches available to *C. lari*. Additionally, although several arginine biosynthetic proteins are not encoded by strain RM2100, the RM2100 genome is predicted to encode *ArgG* and *ArgH* which convert citrulline to arginine. Thus, this organism could potentially synthesize arginine, if provided with an exogenous source of citrulline. Also, it is noteworthy that, unlike *ArgG*, RM2100 *ArgH* is most similar to the *ArgH* proteins found within the δ -proteobacterial taxon (e.g., *Pelobacter* and *Geobacter*), rather than to other *Campylobacter* *ArgH* proteins. However, even in the presence of citrulline, strain RM2100 would still be auxotrophic presumably for alanine, asparagine, leucine, lysine, methionine, proline, and tryptophan. Therefore, as these amino acids cannot be synthesized *de novo* by *C. lari*, they must be obtained exogenously as either free amino acids or via protein degradation. Thus, it is not surprising that *C. lari* RM2100 is predicted to contain the largest repertoire of peptidases among the characterized *Campylobacter* genomes, including the *PepE*, *PepQ*, and *PepT* peptidases that are not encoded by *C. jejuni*.

The absence of two biosynthetic genes, *argD* and *fabI*, from the *C. lari* RM2100 genome has important ramifications for the biology of this organism, relating especially to the cellular envelope. *ArgD*, as described above, is a component of the arginine biosynthetic pathway. However, in *Campylobacter*, *ArgD*

TABLE 3. *C. JEJUNI* GENES ABSENT FROM THE *C. LARI* RM2100 GENOME

Gene(s)	Pathway
Amino acid/cofactor biosynthesis	
<i>argCOBD, argF</i>	Arginine biosynthesis
<i>gltB, gltD</i>	Glutamate biosynthesis
<i>leuABCD</i>	Leucine biosynthesis
<i>metAB, metE, metF</i>	Methionine biosynthesis
<i>panBCD</i>	Pantothenate biosynthesis
<i>proA, proB</i>	Proline biosynthesis
<i>trpAFDE, trpC</i>	Tryptophan biosynthesis
Energy metabolism	
<i>acnB</i>	TCA cycle
<i>gltA</i>	TCA cycle
<i>icd</i>	TCA cycle
<i>sdhABC</i>	TCA cycle
<i>sucCD</i>	TCA cycle
<i>cydAB</i>	Respiratory chain (terminal oxidase)
Miscellaneous	
gene(s)	Function
<i>cfa</i>	Cyclopropane-fatty-acyl-phospholipid synthase
<i>fabI</i>	Enoyl-(acyl carrier protein) reductase
<i>ceuBCDE</i>	Enterochelin uptake ABC transporter
<i>chuABCD</i>	Haemin uptake ABC transporter
<i>acs</i>	Acetyl-coenzyme A synthetase
<i>luxS</i>	Autoinducer-2 production

also catalyzes the conversion of *N*-succinyl-2-amino-6-ketopimelate into *N*-succinyl-*L,L*-2,6-diaminopimelate during lysine biosynthesis. *N*-succinyl-*L,L*-2,6-diaminopimelate is converted further to *meso*-diaminopimelate, a common component of peptidoglycan. Thus, the absence of *argD* would affect not only arginine biosynthesis but lysine and peptidoglycan synthesis as well. ArgD is not predicted to be encoded either by *C. concisus*; however, this organism is predicted to encode the Ddh diaminopimelate dehydrogenase, identified also in *Corynebacterium* and *Bacillus*, which converts tetrahydrodipicolinate to *meso*-diaminopimelate in a single step (Fouts *et al.*, manuscript in preparation), thus obviating the need for ArgD. No obvious ArgD or Ddh homologs are predicted to be encoded by strain RM2100. Therefore, additional investigations will be necessary to determine if a complete lysine biosynthetic pathway is present in *C. lari* and/or whether the *C. lari* peptidoglycan contains *meso*-diaminopimelate. The absence of the enoyl-ACP reductase FabI is also intriguing. FabI catalyzes the terminal step in the elongation cycle of bacterial type II fatty acid biosynthesis and has been shown to be necessary for bacterial viability (Bergler *et al.*, 1996). Some organisms contain the enoyl reductases FabK or FabL in place of or in addition to FabI (Heath and Rock, 2000; Heath *et al.*, 2000); however, no FabK- or FabL-like enoyl reductase is predicted to be encoded by strain RM2100. Enoyl-ACP reductases have a characteristic Tyr-Xaa₆-Lys motif that is also present in the strain RM2100 proteins CL0338 and CL0934. However, additional research will be necessary to determine if these two proteins represent a novel type of enoyl-ACP reductase.

It is unclear what role the absence of both the TCA cycle and the CydAB terminal oxidase plays in *C. lari* biology. *Campylobacter lari* RM2100 does encode pyruvate carboxylase in addition to fumarate dehydrogenase, fumarase, and malate dehydrogenase. Thus, these enzymes may function in part to interconvert metabolic precursors, such as pyruvate and succinate, and provide a limited number of electrons to the respiratory chain. The absence of the citrate synthase gene *gltA* in *C. lari* was identified previously during the development of a novel MLST method (Miller *et al.*, 2005a), since

gltA is a member of the *C. jejuni* MLST gene set. Additionally, the absence of *gltA* was noted during development of a novel MLST method for *C. insulaenigrae* (Stoddard *et al.*, 2007), suggesting the possibility that the TCA cycle may be absent also from this related species.

Halotolerance

Campylobacter lari has an elevated NaCl tolerance, when compared to many other *Campylobacter* species, such as *C. jejuni*. Most *C. lari* strains grow on media amended with 1.5% NaCl (Benjamin *et al.*, 1983); however, growth of *C. lari* is not supported at NaCl concentrations $\geq 3\%$. Therefore, *C. lari* would be considered weakly halotolerant. Analysis of the strain RM2100 genome indicated that this organism cannot synthesize any compatible solutes, e.g., betaine, ectoine, proline, or trehalose, which bacteria use often as protection against hyperosmotic stress (Kempf and Bremer, 1998). However, the *C. lari* RM2100 genome is predicted to encode a BCCT (betaine/choline/carnitine transporter) family protein (CL1410). BCCT family proteins are involved in the osmoregulated uptake of compatible solutes (Peter *et al.*, 1998; Steger *et al.*, 2004). Thus, the observed halotolerance of *C. lari* RM2100 is most likely due to the ability of this strain to import compatible solutes from the environment. Among the characterized *Campylobacter* genomes, BCCT family proteins also have been identified only in the genomes of *C. concisus* 13826 and *C. hominis* ATCC BAA-381. The presence of a BCCT family protein in *C. concisus* is consistent with the growth of this organism on media amended with 1.5% NaCl (Holt, 1994); growth of *C. hominis* on 1.5% NaCl has not been tested (Lawson *et al.*, 2001). Besides *C. lari* and *C. concisus*, only *C. sputorum* and *C. mucosalis* have been shown to grow routinely on media amended with 1.5% NaCl (Holt, 1994; On *et al.*, 1998); it is likely that the *C. sputorum* and *C. mucosalis* genomes also encode BCCT family proteins.

Surface structures

The surface-associated LOS, CPS, and flagella of bacterial pathogens play roles in the pathogenicity of the bacteria and the immune re-

sponse of the host. The LOS is a major component inducing proinflammatory responses, and both the LOS and CPS are immunodominant antigens and interact with host cell factors. Likewise, flagella are antigenic and in *C. jejuni* play a role in internalization and translocation across epithelial cells (Grant *et al.*, 1993). The LOS biosynthesis locus of *C. lari* RM2100 (*cl1187-cl1208*) possesses an organization similar to *C. jejuni* (Gilbert *et al.*, 2005; Parker *et al.*, 2005). At both ends of the locus are genes involved in the addition of heptose (*waaF* [*cl1187*] and *waaC* [*cl1208*]). In between is a region containing genes whose products have functions related to LOS biosynthesis, including several glycosyltransferases (*cl1201-cl1206*) and genes involved in synthesis of sugar intermediates (*cl1189-cl1198*) (Suppl. Table S1). Of particular note, many of the genes involved in the synthesis of sugar intermediates (*cl1189-cl1198*) are similar to genes of the flagellar modification locus of *C. jejuni*. Until mutational analysis of these genes is performed, the role of these genes in either LOS biosynthesis or flagellar modification is not certain. Indeed, it has been observed that polysaccharide biosynthesis genes can function in different pathways (Karlyshev *et al.*, 2005). The organization of the *C. lari* RM2100 CPS biosynthesis locus (*cl0285-cl0314*) is also similar to *C. jejuni* (Karlyshev *et al.*, 2005). At both ends of the locus are genes involved in the export of CPS (*kpsMTEDF* [*cl0285-cl0289*] and *kpsCS* [*cl0313-cl0314*]). Like the LOS locus, the region between these *kps* genes contains genes whose products have functions related to CPS biosynthesis, including glycosyltransferases and genes involved in synthesis of sugar intermediates. There is also evidence of phase variation of the capsule synthesis or capsular structure, due to the presence of two genes (*cl0305* and *cl0307*) with variable length homopolymeric G:C tracts. Finally, similar to *C. jejuni*, the *C. lari* strain RM2100 is a motile bacterium that synthesizes a polar flagellum and the flagellar genes are widely dispersed in small gene clusters throughout the genome.

Prophage

Campylobacter lari RM2100 contains an inserted element, termed CLIE1 (*Campylobacter*

lari integrated element 1; [Fouts *et al.*, 2005]). CLIE1 is 36,280 bp in length (bp 789,761–826,040), contains 47 CDSs (*cl0787-cl0833*), is integrated into a leucyl-tRNA (tRNA-Leu-1; Suppl. Table S1), and is bounded by 29 bp terminal direct repeats. The presence of phage capsid-encoding genes within CLIE1 suggests that this element is an integrated prophage and similar to CJIE4, a putative prophage in the genome of *C. jejuni* RM1221 (Fouts *et al.*, 2005). Thirty-two of the 47 proteins predicted to be encoded by CLIE1 are between 32% and 85% identical to proteins encoded by CJIE4 (Suppl. Table S2). Moreover, a large region of synteny exists between the two integrated elements: *cl0808-cl0825* and *cje1450-cje1469*. Directly upstream of this syntenic region are four genes (*cl0803-cl0807*) with similarity to four genes (*cjejd_02001129-cjejd02001126*) contained within the *C. jejuni* subsp. *doylei* 269.97 integrated element DIE5 (*doylei* integrated element 5; Fouts *et al.*, manuscript in preparation). Phage-related genes are also present within DIE5, suggesting that CLIE1, CJIE4, and perhaps DIE5 are members of a related *Campylobacter* bacteriophage family. The remainder of non-phage-related genes within CLIE1 encodes proteins of no discernable function; no toxins or other virulence factors are predicted to be encoded by CLIE1. Therefore, it is uncertain what role, if any, CLIE1 plays in *C. lari* biology.

Antibiotic resistance

The pattern of antibiotic resistance in *C. lari* RM2100 is similar to that of *C. jejuni* NCTC 11168 (data not shown) with two exceptions: strain RM2100 demonstrated a reduced resistance to the cephalosporin antibiotics cefmetazole, cefoperazone, cefotaxime, cefotetan, cefoxitin, ceftazidime, and ceftibuten, and was resistant to the aminocoumarin (e.g., novobiocin), quinolone (e.g., nalidixic acid and cinoxacin), and fluoroquinolone (e.g., ciprofloxacin and enrofloxacin) classes of DNA gyrase inhibitors. Among *C. lari* strains, decreased resistance to second and third generation cephalosporins has not been reported; therefore, it cannot be determined whether the reduced resistance of strain RM2100 to these antibiotics reflects the species as a whole. Cephalosporin resistance in

Gram-negative bacteria is due to altered binding of the antibiotic to penicillin-binding proteins, altered outer membrane permeability or the action of β -lactamases (Georgopapadakou, 1993). It is unknown which of these three mechanisms, alone or in combination, has been altered in strain RM2100. Analysis of the strain RM2100 genome indicates that this organism has an altered penicillin-binding protein profile relative to strain NCTC 11168; however, additional investigations will be necessary to determine the nature of the increased sensitivity to cephalosporins in this strain. Resistance to DNA gyrase inhibitors is common among *C. lari* strains (Thwaites and Frost, 1999; Leatherbarrow *et al.*, 2007). Quinolone resistance in *Campylobacter* is often the result of amino acid changes (specifically at Thr-86, Asp-90, or Ala-70) within the DNA gyrase subunit GyrA (Taylor and Tracz, 2005); mutations at Thr-86 have been shown to be associated with high level resistance to fluoroquinolones (Wang *et al.*, 1993). Consistent with these findings, the predicted GyrA protein of strain RM2100 contains a Thr-86-Val substitution, resulting presumably in the resistance of this strain to a broad spectrum of DNA gyrase inhibitors.

Virulence determinants

Campylobacter lari RM2100 is a human clinical isolate. Moreover, *C. lari* strains have been isolated, albeit infrequently, from human diarrheic stools (Tauxe *et al.*, 1985; Broczyk *et al.*, 1987; Lastovica and Skirrow, 2000). Therefore, it is likely that virulence determinants would be present in the strain RM2100 genome. One of the major virulence determinants discussed in relation to *Campylobacter* is the cytolethal distending toxin CDT (encoded by *cdtABC*; Pickett and Lee, 2005). CDT is an exotoxin that blocks eukaryotic cells irreversibly in the G₁ or G₂ phase (Pickett and Lee, 2005). Within *Campylobacter*, *cdt* genes have been identified in *C. jejuni* subsp. *jejuni*, *C. coli*, *C. upsaliensis*, and *C. fetus* (Fouts *et al.*, 2005; Asakura *et al.*, 2007). *Campylobacter lari* RM2100 contains also all three *cdt* genes; however, *cdtB* is truncated at the 5' end with a suboptimal ribosome binding site. Therefore, it is possible that *cdtB* in this strain is

a pseudogene and strain RM2100 possesses no CDT activity. Other *C. lari cdtB* genes submitted to the NCBI nr database show similar 5' truncations, suggesting that *cdtB* 5' truncation may be a general feature of *C. lari* strains. Truncations and deletions within the *cdt* locus were also identified in *C. jejuni* subsp. *doylei* (Parker *et al.*, 2007). It is noteworthy that, like *C. jejuni* subsp. *doylei*, *C. lari* strains have also been isolated from patients with bacteremia (Martinot *et al.*, 2001); therefore, it is possible that a link exists between the absence of CDT and bacteremia. In addition to *cdtABC*, other *Campylobacter* virulence-associated genes, such as *ciaB* and *moiN*, are present in the strain RM2100 genome; however, the JlpA lipoprotein is not encoded by strain RM2100. All three putative fibronectin-binding proteins identified in *C. jejuni* NCTC 11168 (CadF, Cj1279c, and Cj1349c) are encoded also by strain RM2100 (CL0381, CL0576, and CL1486, respectively).

In addition to virulence determinants described previously in *Campylobacter*, *C. lari* RM2100 is predicted to encode a filamentous hemagglutinin (FHA)-like hemagglutinin (CL1478). Also, one potential virulence determinant, novel to the thermotolerant campylobacters, in strain RM2100, is the Cu/Zn superoxide dismutase SodC. All of the characterized *Campylobacteraceae* genomes are predicted to encode the cytoplasmic Fe superoxide dismutase SodB (Miller *et al.*, 2007) that serves to detoxify superoxides arising from aerobic respiration. However, in addition to SodB, the genomes of *C. fetus*, *C. concisus*, *C. curvus*, and *C. lari* RM2100 are predicted to encode the periplasmic SodC. The apparent impermeability of the inner membrane to superoxides (Sadosky *et al.*, 1994) suggests that the periplasmic location of SodC confers protection against superoxides of exogenous origin and not against cytoplasmic pools of superoxides (Steinman, 1993). Sources of exogenous superoxides include macrophages, which use a respiratory burst of superoxides as a bacterial cytotoxic mechanism. Indeed, SodC has been demonstrated to contribute to survival within host macrophages (Gee *et al.*, 2005; Keith and Valvano, 2007). SodC has been implicated also in biofilm formation (Kim *et al.*, 2006) and protection against heme toxicity (Negari *et al.*, 2008).

Plasmids

Campylobacter lari RM2100 contains a 46,201 bp megaplasmid, termed pCL2100 (pCL46 [Fouts *et al.*, 2005]; Fig. 2). The majority of proteins (24/42) predicted to be encoded by pCL2100 are similar to those encoded by the *C. jejuni* strain 81-176 conjugative plasmid pTet (Batchelor *et al.*, 2004; Table 4). Nevertheless, differences are present between the two megaplasmids in both gene content and organization, notably the absence of antibiotic-resistance genes in pCL2100. Many of the 42 proteins predicted to be encoded by pCL2100 are similar also to Tra/Vir type IV secretion system (T4SS)

proteins encoded by Gram-negative conjugative plasmids (Lawley *et al.*, 2004). These include a putative pilin protein (the VirB2 homolog Cla008), proteins comprising the conjugative pore (the VirB5, VirB6, VirB8, VirB9, and VirB10 homologs Cla010, Cla014, Cla016, Cla017, and Cla018, respectively), and other proteins predicted to be components of the conjugative apparatus. The pTet and pCL2100 T4SS gene clusters have a similar organization, with complete synteny between the *cla016-cla026* and *cmgB8-cpp49* regions. However, differences between the two T4SS gene clusters were identified, such as the position of the *ssb* single-stranded DNA binding protein-encoding gene

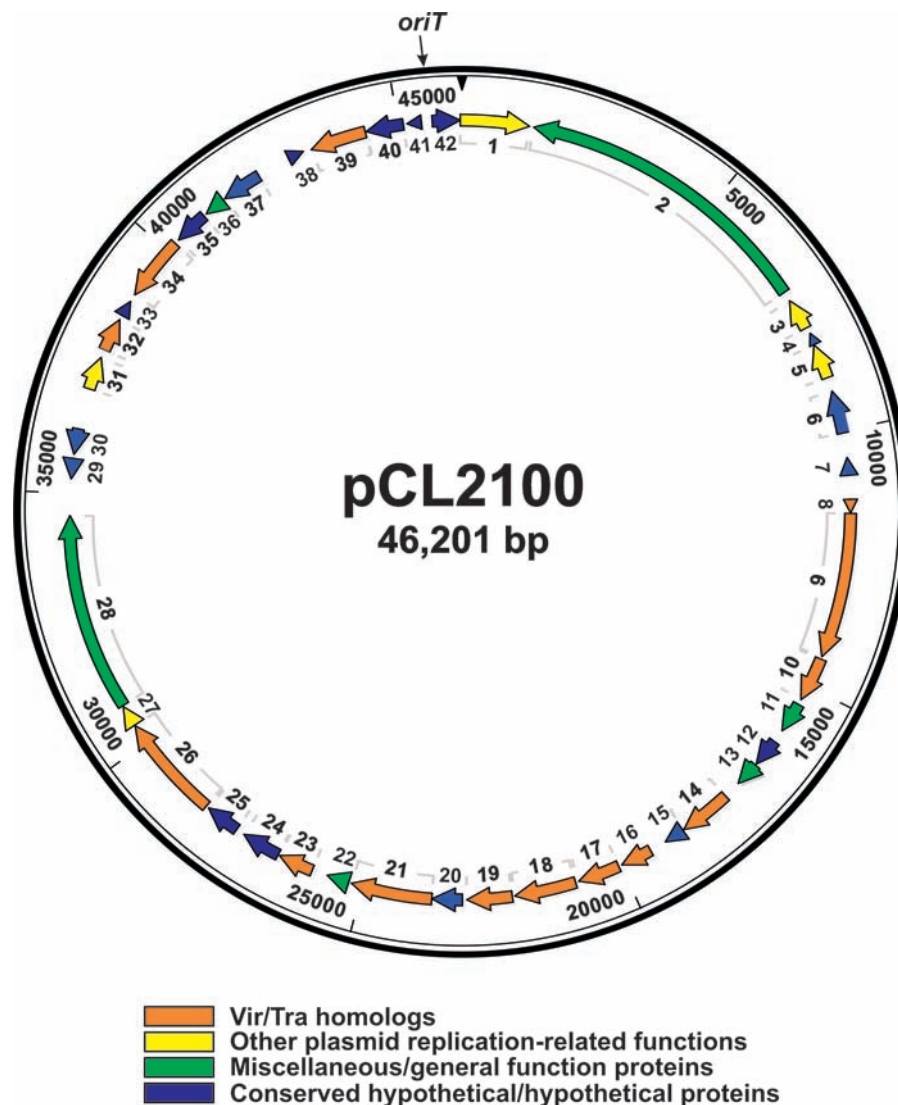


FIG. 2. Diagrammatic representation of the *C. lari* RM2100 megaplasmid pCL2100. All genes are drawn to scale. *oriT*, putative origin of transfer.

and the presence of two putative antirepressor-encoding genes (*cla011* and *cla013*) between the *cmgB5* and *cmgB6* homologs, genes adjacent on pTet (Batchelor *et al.*, 2004). pCL2100 contains also a putative origin of DNA transfer (*oriT*; Fig. 2) between the *cyp19* and *cyp18* homologs *cla041* and *cla042*, similar to the pTet *oriT* region (Batchelor *et al.*, 2004). This region includes an RP4-like *nic* nickase cleavage site (5' TATCCTGC 3') (Lawley *et al.*, 2004). Finally, the megaplasmid pCL2100 is predicted to encode a putative invasin/adhesin (Cla028). This protein

is similar to the *Salmonella* SinH and *E. coli* EaeH invasin/adhesins. However, both the levels of similarity and regions of overlap between Cla028 and the SinH/EaeH-like proteins are small. No experimental evidence exists that suggests that Cla028 is involved in *C. lari* virulence. It is likely that Cla028 has an alternative function; one possibility, given its proximity to the T4SS gene cluster, is that Cla028 might play a role in plasmid conjugation, perhaps as a IncF1 family, TraN-like adhesin (Lawley *et al.*, 2004).

TABLE 4. PLASMID pCL2100 CODING SEQUENCES

CDS	Function ^a	pTet/pVir homolog [aa identity]
<i>cla001</i>	Nickase/relaxase	Cpp17 (pTet) [61%]
<i>cla002</i>	DNA methylase/helicase	Cpp14 (pTet) [80%]
<i>cla003</i>	Site-specific recombinase/resolvase	Cpp27 (pTet) [57%]
<i>cla004</i>	Conserved hypothetical protein	
<i>cla005</i>	Plasmid partitioning protein	Cjp26 (pVir) [56%]
<i>cla006</i>	Conserved hypothetical protein	
<i>cla007</i>	Hypothetical protein	
<i>cla008</i>	Attachment mediating protein TraC/VirB2	
<i>cla009</i>	Secretion/conjugal transfer ATPase VirB3/4	CmgB3/4 (pTet) [66%]
<i>cla010</i>	Conjugal transfer protein VirB5	CmgB5 (pTet) [62%]
<i>cla011</i>	Putative antirepressor, Rha family	Cpp32 (pTet) [52%]
<i>cla012</i>	Conserved hypothetical protein	Cpp33 (pTet) [71%]
<i>cla013</i>	Putative antirepressor, BRO family	
<i>cla014</i>	Conjugal transfer protein TrbL/VirB6	CmgB6 (pTet) [44%]
<i>cla015</i>	Conserved hypothetical protein	
<i>cla016</i>	Conjugal transfer protein TraJ/VirB8	CmgB8 (pTet) [74%]
<i>cla017</i>	Conjugal transfer protein TraK/VirB9	CmgB9 (pTet) [68%]
<i>cla018</i>	Conjugal transfer protein TrbI/VirB10	CmgB10 (pTet) [58%]
<i>cla019</i>	Conjugal transfer protein VirB11	CmgB11 (pTet) [72%]
<i>cla020</i>	Conserved hypothetical protein	
<i>cla021</i>	Conjugal transfer protein TraG/VirD4	CmgD4 (pTet) [68%]
<i>cla022</i>	Conserved hypothetical protein, putative CagT	Cpp44 (pTet) [48%]
<i>cla023</i>	TrbM mating pair formation protein homolog	Cpp45 (pTet) [50%]
<i>cla024</i>	Conserved hypothetical protein	Cpp46 (pTet) [34%]
<i>cla025</i>	Conserved hypothetical protein	Cpp47 (pTet) [55%]
<i>cla026</i>	DNA topoisomerase III TraE	Cpp49 (pTet) [65%]
<i>cla027</i>	Single-stranded DNA binding protein	
<i>cla028</i>	Putative adhesin/invasin	
<i>cla029</i>	Hypothetical protein	
<i>cla030</i>	Hypothetical protein	
<i>cla031</i>	Site-specific recombinase/resolvase	
<i>cla032</i>	Putative DNA primase TraC	
<i>cla033</i>	Conserved hypothetical protein	
<i>cla034</i>	DNA primase TraC	Cpp22 (pTet) [64%]
<i>cla035</i>	Conserved hypothetical protein	Cpp21 (pTet) [84%]
<i>cla036</i>	PIN domain protein	
<i>cla037</i>	Conserved hypothetical protein	
<i>cla038</i>	Conserved hypothetical protein	
<i>cla039</i>	DNA primase TraC	Cpp22 (pTet) [70%]
<i>cla040</i>	Conserved hypothetical protein	Cpp20 (pTet) [38%]
<i>cla041</i>	Conserved hypothetical protein	Cpp19 (pTet) [38%]
<i>cla042</i>	Conserved hypothetical protein	Cpp18 (pTet) [47%]

^aPredicted CDS functions are based on BLASTP comparisons to proteins in the NCBI nonredundant (nr) database. CDS, coding sequence; aa, amino acid.

Conclusions

Campylobacter lari RM2100 is a member of the thermotolerant campylobacters and, as such, is very similar in gene content to other members of the group, for example *C. jejuni* and *C. coli*. Consistent with other *Campylobacter* species, *C. lari* is predicted to be unable to utilize sugars as carbon sources and contains an incomplete glycolytic pathway. Additionally, the *C. lari* RM2100 genome contains multiple homopolymeric G:C tracts that are located predominantly within surface structure-related genes, a feature shared especially with other thermotolerant *Campylobacter*. Indeed, 90% of the non-unique *C. lari* strain RM2100 proteins are similar to those encoded by at least one of the other thermotolerant campylobacters. Therefore, with a few exceptions, the singular aspect of the *C. lari* RM2100 genome is not its gene content, *per se*, but is related rather to those genes predicted to be absent in this strain.

The majority of these missing genes encode proteins involved primarily in amino acid or cofactor biosynthesis and energy metabolism. *Campylobacter lari* RM2100 is predicted to be unable to synthesize acetyl-coA, alanine, arginine, asparagine, leucine, lysine, methionine, pantothenate, proline, and tryptophan *de novo*. In addition, this strain is not predicted to contain a complete TCA cycle, due to the absence of five components, including citrate synthase. The absence of these genes, and others, from the *C. lari* RM2100 genome, would necessarily affect the biology of this organism.

Relative to other thermotolerant campylobacters, *C. lari* is a phenotypically and genotypically diverse species. Strain RM2100 is a member of one of the four major *C. lari* subgroups, the nalidixic acid-resistant thermophilic campylobacters. No information on the gene content of the other three major subgroups is available, other than that the UPTC subgroup would obviously encode proteins involved in urease biosynthesis. In light of the above findings, it should prove interesting to determine if members of the UPTC or NASC *C. lari* groups possess similar defects in their metabolic pathways. The genome of strain RM2100 could serve as a base genome for such comparisons. Analysis of *C. insulaenigrae* might also prove useful

in studies concerning *C. lari*; a CLUSTAL alignment of concatenated MLST allele sequences (Suppl. Fig. S2) indicated that *C. insulaenigrae*, among the present thermotolerant taxa, is the closest relative of *C. lari* (86.2% average nucleotide identity; Suppl. Fig. S3). Thus, it is noteworthy that the citrate synthase gene *gltA* is also absent presumably from *C. insulaenigrae*, and suggests that many of the features identified in this study are common to the taxonomic group that includes *C. lari* and *C. insulaenigrae*. Since *C. lari* is isolated often from seawater, shorebirds, and shellfish and *C. insulaenigrae* is isolated predominantly from marine mammals, it is intriguing to speculate that features common to both species, identified perhaps in this study, might permit adaptation of this taxonomic group to marine environments.

Acknowledgments

This work was funded by the United States Department of Agriculture, Agricultural Research Service, CRIS project 5325-42000-230-045.

References

- Aarestrup FM, Nielsen EM, Madsen M, and Engberg J. Antimicrobial susceptibility patterns of thermophilic *Campylobacter* spp. from humans, pigs, cattle, and broilers in Denmark. *Antimicrob. Agents Chemother.* 1997;41:2244–2250.
- Asakura M, Samosornsuk W, Taguchi M, Kobayashi K, Misawa N, Kusumoto M, Nishimura K, Matsuhisa A, and Yamasaki S. Comparative analysis of cytolethal distending toxin (*cdt*) genes among *Campylobacter jejuni*, *C. coli* and *C. fetus* strains. *Microb. Pathog.* 2007;42:174–183.
- Batchelor RA, Pearson BM, Friis LM, Guerry P, and Wells JM. Nucleotide sequences and comparison of two large conjugative plasmids from different *Campylobacter* species. *Microbiology* 2004;150:3507–3517.
- Benjamin J, Leaper S, Owen RJ, and Skirrow MB. Description of *Campylobacter laridis*, a new species comprising the nalidixic acid resistant thermophilic *Campylobacter* (NARTC) group. *Curr. Microbiol.* 1983; 8:221–238.
- Bergler H, Fuchsbichler S, Hogenauer G, and Turnowsky F. The enoyl-[acyl-carrier-protein] reductase (FabI) of *Escherichia coli*, which catalyzes a key regulatory step in fatty acid biosynthesis, accepts NADH and NADPH as cofactors and is inhibited by palmitoyl-CoA. *Eur. J. Biochem.* 1996;242:689–694.
- Broczyk A, Thompson S, Smith D, and Lior H. Waterborne outbreak of *Campylobacter laridis*-associated gastroenteritis. *Lancet* 1987;1:164–165.

- Csonka LN, Ikeda TP, Fletcher SA, and Kustu S. The accumulation of glutamate is necessary for optimal growth of *Salmonella typhimurium* in media of high osmolality but not induction of the *proU* operon. *J. Bacteriol.* 1994;176:6324–6333.
- Duim B, Wagenaar JA, Dijkstra JR, Goris J, Endtz HP, and Vandamme PA. Identification of distinct *Campylobacter lari* genogroups by amplified fragment length polymorphism and protein electrophoretic profiles. *Appl. Environ. Microbiol.* 2004;70:18–24.
- Endtz HP, Vliegthart JS, Vandamme P, Weverink HW, van den Braak NP, Verbrugh HA, and van Belkum A. Genotypic diversity of *Campylobacter lari* isolated from mussels and oysters in the Netherlands. *Int. J. Food Microbiol.* 1997;34:79–88.
- Engvall EO, Brandstrom B, Andersson L, Baverud V, Trowald-Wigh G, and Englund L. Isolation and identification of thermophilic *Campylobacter* species in faecal samples from Swedish dogs. *Scand. J. Infect. Dis.* 2003;35:713–718.
- Fouts DE, Mongodin EF, Mandrell RE, Miller WG, Rasko DA, Ravel J, Brinkac LM, Deboy RT, Parker CT, Daugherty SC, Dodson RJ, Durkin AS, Madupu R, Sullivan SA, Shetty JU, Ayodeji MA, Shvartsbeyn A, Schatz MC, Badger JH, Fraser CM, and Nelson KE. Major structural differences and novel potential virulence mechanisms from the genomes of multiple campylobacter species. *PLoS Biol.* 2005;3:e15.
- Gee JM, Valderas MW, Kovach ME, Grippe VK, Robertson GT, Ng WL, Richardson JM, Winkler ME, and Roop RM 2nd. The *Brucella abortus* Cu,Zn superoxide dismutase is required for optimal resistance to oxidative killing by murine macrophages and wild-type virulence in experimentally infected mice. *Infect. Immun.* 2005;73:2873–2880.
- Georgopapadakou NH. Penicillin-binding proteins and bacterial resistance to beta-lactams. *Antimicrob Agents Chemother.* 1993;37:2045–2053.
- Giacoboni GI, Itoh K, Hirayama K, Takahashi E, and Mitsuoka T. Comparison of fecal *Campylobacter* in calves and cattle of different ages and areas in Japan. *J. Vet. Med. Sci.* 1993;55:555–559.
- Gilbert M, Godschalk P, Parker CT, Endtz H, and Warkarchuk W. Genetic basis for the variation in the lipooligosaccharide outer core of *Campylobacter jejuni* and possible association of glycosyltransferase genes with post-infectious neuropathies. In: *Campylobacter: Molecular and Cellular Biology*. Ketley, J. M. and M. E. Konkel (eds). Norfolk, UK: Horizon Bioscience, 2005, pp. 219–248.
- Glunder G and Petermann S. The occurrence and characterization of *Campylobacter* spp. in silver gulls (*Larus argentatus*), three-toed gulls (*Rissa tridactyla*) and house sparrows (*Passer domesticus*) [In German]. *Zentralbl. Veterinarmed. B.* 1989;36:123–130.
- Grant CC, Konkel ME, Cieplak W Jr, and Tompkins LS. Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures. *Infect. Immun.* 1993; 61:1764–1771.
- Harvey RB, Young CR, Ziprin RL, Hume ME, Genovese KJ, Anderson RC, Droleskey RE, Stanker LH, and Nisbet DJ. Prevalence of *Campylobacter* spp isolated from the intestinal tract of pigs raised in an integrated swine production system. *J. Am. Vet. Med. Assoc.* 1999; 215:1601–1604.
- Heath RJ and Rock CO. A triclosan-resistant bacterial enzyme. *Nature* 2000;406:145–146.
- Heath RJ, Su N, Murphy CK, and Rock CO. The enoyl-[acyl-carrier-protein] reductases FabI and FabL from *Bacillus subtilis*. *J. Biol. Chem.* 2000;275:40128–40133.
- Hofreuter D, Tsai J, Watson RO, Novik V, Altman B, Benitez M, Clark C, Perbost C, Jarvie T, Du L, and Galan JE. Unique features of a highly pathogenic *Campylobacter jejuni* strain. *Infect. Immun.* 2006;74:4694–4707.
- Holt JG. *Bergey's Manual of Determinative Bacteriology*. Baltimore, MD: Lippincott Williams & Wilkins, 1994.
- Karlyshev AV, Champion OL, Churcher C, Brisson JR, Jarrell HC, Gilbert M, Brochu D, St Michael F, Li J, Warkarchuk WW, Goodhead I, Sanders M, Stevens K, White B, Parkhill J, Wren BW, and Szymanski CM. Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol. Microbiol.* 2005;55:90–103.
- Keith KE and Valvano MA. Characterization of SodC, a periplasmic superoxide dismutase from *Burkholderia cenocepacia*. *Infect. Immun.* 2007;75:2451–2460.
- Kempf B and Bremer E. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* 1998;170: 319–330.
- Kim YH, Lee Y, Kim S, Yeom J, Yeom S, Seok Kim B, Oh S, Park S, Jeon CO, and Park W. The role of periplasmic antioxidant enzymes (superoxide dismutase and thiol peroxidase) of the Shiga toxin-producing *Escherichia coli* O157:H7 in the formation of biofilms. *Proteomics* 2006;6:6181–6193.
- Lastovica AJ and Skirrow MB. Clinical significance of *Campylobacter* and related species other than *Campylobacter jejuni* and *C. coli*. In: *Campylobacter*. Nachamkin I and Blaser MJ (eds). Washington, DC: ASM Press, 2000, pp. 89–120.
- Lawley T, Wilkins BM, and Frost LS. Bacterial conjugation in Gram-negative bacteria. In: *Plasmid Biology*. Funnell BE and Phillips GJ (eds). Washington, DC: ASM Press, 2004, pp. 203–226.
- Lawson AJ, On SL, Logan JM, and Stanley J. *Campylobacter hominis* sp. nov., from the human gastrointestinal tract. *Int. J. Syst. Evol. Microbiol.* 2001;51:651–660.
- Leatherbarrow AJ, Griffiths R, Hart CA, Kemp R, Williams NJ, Diggle PJ, Wright EJ, Sutherst J, Houghton P, and French NP. *Campylobacter lari*: genotype and antibiotic resistance of isolates from cattle, wildlife and water in an area of mixed dairy farmland in the United Kingdom. *Environ. Microbiol.* 2007;9:1772–1779.
- Lin CW, Yin PL, and Cheng KS. Incidence and clinical manifestations of *Campylobacter* enteritis in central Taiwan. *Zhonghua Yi Xue Za Zhi (Taipei)* 1998;61:339–345.
- Lindblom GB, Johnny M, Khalil K, Mazhar K, Ruiz-Palacios GM, and Kaijser B. Enterotoxigenicity and frequency of

- Campylobacter jejuni*, *C. coli* and *C. lariidis* in human and animal stool isolates from different countries. FEMS Microbiol. Lett. 1990;54:163–167.
- Lowe TM and Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25:955–964.
- Martinot M, Jaulhac B, Moog R, De Martino S, Kehrl P, Monteil H, and Piemont Y. *Campylobacter lari* bacteremia. Clin. Microbiol. Infect. 2001;7:96–97.
- Miller WG, On SL, Wang G, Fontanoz S, Lastovica AJ, and Mandrell RE. Extended multilocus sequence typing system for *Campylobacter coli*, *C. lari*, *C. upsaliensis*, and *C. helveticus*. J. Clin. Microbiol. 2005a;43:2315–2329.
- Miller WG, Parker CT, Rubenfield M, Mendz GL, Wosten MM, Ussery DW, Stolz JF, Binnewies TT, Hallin PF, Wang G, Malek JA, Rogosin A, Stanker LH, and Mandrell RE. The complete genome sequence and analysis of the Epsilonproteobacterium *Arcobacter butzleri*. PLoS ONE. 2007;2:e1358.
- Miller WG, Pearson BM, Wells JM, Parker CT, Kapitonov VV, and Mandrell RE. Diversity within the *Campylobacter jejuni* type I restriction-modification loci. Microbiology 2005b;151:337–351.
- Moore JE, Gilpin D, Crothers E, Canney A, Kaneko A, and Matsuda M. Occurrence of *Campylobacter* spp. and *Cryptosporidium* spp. in seagulls (*Larus* spp.). Vector Borne Zoonotic Dis. 2002;2:111–114.
- Morris CN, Scully B, and Garvey GJ. *Campylobacter lari* associated with permanent pacemaker infection and bacteremia. Clin. Infect. Dis. 1998;27:220–221.
- Nachamkin I, Stowell C, Skalina D, Jones AM, Hoop RM 2nd, and Smibert RM. *Campylobacter lariidis* causing bacteremia in an immunosuppressed patient. Ann Intern Med. 1984;101:55–57.
- Negari S, Sulpher J, Pacello F, Ingrey K, Battistoni A, and Lee BC. A role for *Haemophilus ducreyi* Cu,ZnSOD in resistance to heme toxicity. Biometals 2008;21:249–258.
- Obiri-Danso K and Jones K. Distribution and seasonality of microbial indicators and thermophilic campylobacters in two freshwater bathing sites on the River Lune in northwest England. J. Appl. Microbiol. 1999a;87:822–832.
- Obiri-Danso K and Jones K. The effect of a new sewage treatment plant on faecal indicator numbers, campylobacters and bathing water compliance in Morecambe Bay. J. Appl. Microbiol. 1999b;86:603–614.
- Obiri-Danso K, Paul N, and Jones K. The effects of UVB and temperature on the survival of natural populations and pure cultures of *Campylobacter jejuni*, *Camp. coli*, *Camp. lari* and urease-positive thermophilic campylobacters (UPTC) in surface waters. J. Appl. Microbiol. 2001;90:256–267.
- On SL, Atabay HI, Corry JE, Harrington CS, and Vandamme P. Emended description of *Campylobacter sputorum* and revision of its infrasubspecific (biovar) divisions, including *C. sputorum* biovar *paraureolyticus*, a urease-producing variant from cattle and humans. Int. J. Syst. Bacteriol. 1998;48 Pt 1:195–206.
- On SL and Harrington CS. Identification of taxonomic and epidemiological relationships among *Campylobacter* species by numerical analysis of AFLP profiles. FEMS Microbiol. Lett. 2000;193:161–169.
- Otasevic M, Lazarevic-Jovanovic B, Tasic-Dimov D, Dordevic N, and Miljkovic-Selimovic B. The role of certain *Campylobacter* types in the etiology of enterocolitis [In Serbian]. Vojnosanit Pregl. 2004;61:21–27.
- Parker CT, Horn ST, Gilbert M, Miller WG, Woodward DL, and Mandrell RE. Comparison of *Campylobacter jejuni* lipooligosaccharide biosynthesis loci from a variety of sources. J. Clin. Microbiol. 2005;43:2771–2781.
- Parker CT, Miller WG, Horn ST, and Lastovica AJ. Common genomic features of *Campylobacter jejuni* subsp. *doylei* strains distinguish them from *C. jejuni* subsp. *jejuni*. BMC Microbiol. 2007;7:50.
- Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream MA, Rutherford KM, van Vliet AH, Whitehead S, and Barrell BG. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature 2000;403:665–668.
- Pearson BM, Gaskin DJ, Segers RP, Wells JM, Nuijten PJ, and van Vliet AH. The complete genome sequence of *Campylobacter jejuni* strain 81116 (NCTC11828). J. Bacteriol. 2007;189:8402–8403.
- Peter H, Weil B, Burkovski A, Kramer R, and Morbach S. *Corynebacterium glutamicum* is equipped with four secondary carriers for compatible solutes: identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP. J. Bacteriol. 1998;180:6005–6012.
- Pickett CL and Lee RB. Cytolethal distending toxin. In: *Campylobacter: Molecular and Cellular biology*. Ketley JM and Konkel ME (eds). Norfolk, UK: Horizon Bioscience, 2005, pp. 385–396.
- Prasad KN, Dixit AK, and Ayyagari A. *Campylobacter* species associated with diarrhoea in patients from a tertiary care centre of north India. Indian J. Med. Res. 2001;114:12–17.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, and Barrell B. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–945.
- Sadosky AB, Wilson JW, Steinman HM, and Shuman HA. The iron superoxide dismutase of *Legionella pneumophila* is essential for viability. J. Bacteriol. 1994;176:3790–3799.
- Skirrow MB and Benjamin J. A human strain of *Campylobacter fetus* subsp. *intestinalis* grown at 42 degrees C. J. Clin. Pathol. 1980;33:603–604.
- Skirrow MB, Jones DM, Sutcliffe E, and Benjamin J. *Campylobacter* bacteraemia in England and Wales, 1981–91. Epidemiol. Infect. 1993;110:567–573.
- Smibert RM. Genus *Campylobacter*. In: *Bergey's manual of systematic bacteriology*. Krieg NR and Holt HG (eds). Baltimore, MD: The Williams & Wilkins Co., 1984, pp. 111–118.
- Soderstrom C, Schalen C, and Walder M. Septicaemia caused by unusual *Campylobacter* species (*C. lariidis* and *C. mucosalis*). Scand. J. Infect. Dis. 1991;23:369–371.

- Steger R, Weinand M, Kramer R, and Morbach S. LcoP, an osmoregulated betaine/ectoine uptake system from *Corynebacterium glutamicum*. FEBS Lett. 2004;573:155–3160.
- Steinman HM. Function of periplasmic copper-zinc superoxide dismutase in *Caulobacter crescentus*. J. Bacteriol. 1993;175:1198–1202.
- Stoddard RA. *Salmonella* and *Campylobacter* spp. in northern elephant seals, California. Emerg. Infect. Dis. 2005;11:1967–1969.
- Stoddard RA, Miller WG, Foley JE, Lawrence J, Gulland FM, Conrad PA, and Byrne BA. *Campylobacter insulaenigrae* Isolates from northern elephant seals (*Mirounga angustirostris*) in California. Appl. Environ. Microbiol. 2007;73:1729–1735.
- Tauxe RV, Patton CM, Edmonds P, Barrett TJ, Brenner DJ, and Blake PA. Illness associated with *Campylobacter lari*, a newly recognized *Campylobacter* species. J. Clin. Microbiol. 1985;21:222–225.
- Taylor DE and Tracz DM. Mechanisms of antimicrobial resistance in *Campylobacter*. In: *Campylobacter: Molecular and Cellular biology*. Ketley JM and Konkel ME (eds). Norfolk, UK: Horizon Bioscience, 2005, pp. 193–204.
- Thwaites RT and Frost JA. Drug resistance in *Campylobacter jejuni*, *C coli*, and *C lari* isolated from humans in north west England and Wales, 1997. J. Clin. Pathol. 1999;52:812–814.
- Tresierra-Ayala A, Bendayan ME, Bernuy A, Pereyra G, and Fernandez H. Chicken as potential contamination source of *Campylobacter lari* in Iquitos, Peru. Rev. Inst. Med. Trop. Sao Paulo 1994;36:497–499.
- Van Doorn LJ, Verschuuren-Van Haperen A, Van Belkum A, Endtz HP, Vliegthart JS, Vandamme P, and Quint WG. Rapid identification of diverse *Campylobacter lari* strains isolated from mussels and oysters using a reverse hybridization line probe assay. J. Appl. Microbiol. 1998;84:545–550.
- Waldenstrom J, On SL, Ottvall R, Hasselquist D, and Olsen B. Species diversity of campylobacteria in a wild bird community in Sweden. J. Appl. Microbiol. 2007;102:424–432.
- Wang Y, Huang WM, and Taylor DE. Cloning and nucleotide sequence of the *Campylobacter jejuni gyrA* gene and characterization of quinolone resistance mutations. Antimicrob. Agents Chemother. 1993;37: 457–463.

Address correspondence to:

William G. Miller, Ph.D.

Produce Safety and

Microbiology Research Unit

U.S. Department of Agriculture

Agricultural Research Service

800 Buchanan St.

Albany, CA 94710

E-mail: William.Miller@ars.usda.gov