Campylobacter corcagiensis sp. nov., isolated from faeces of captive lion-tailed macaques (Macaca silenus)

Monika Koziel,[1] Pat O'Doherty,[2] Peter Vandamme,[3] Gerard D. Corcoran,[4] Roy D. Sleator[1] and Brigid Lucey[1]

Correspondence
Roy D. Sleator
roy.sleator@cit.ie

[1]Department of Biological Sciences, Cork Institute of Technology, Bishopstown, Cork, Ireland

[2]Gilabbey Veterinary Hospital, Vicars Road, Cork, Ireland

[3]Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium

[4]Department of Microbiology, Cork University Hospital, Wilton, Cork, Ireland

An investigation of the prevalence of *Campylobacter ureolyticus* in a variety of animals led to the identification of the strain CIT 045[T], in the faeces of captive lion-tailed macaques (*Macaca silenus*). Originally, believed to be *Campylobacter ureolyticus* based on the colony morphology and positive urease test, analysis of 16S rRNA and *hsp60* gene sequences of this isolate revealed that the strain differs significantly from other species of the genus *Campylobacter* described to date. Species-specific primers for 16S rRNA and *hsp60* genes were designed and used to identify two additional strains isolated from faeces samples from other macaques. Nucleotide sequence analysis of the 16S rRNA and *hsp60* genes revealed $\leqslant 95\%$ and $\leqslant 82\%$ sequence similarity to recognized species of the genus *Campylobacter* respectively. All three isolates formed a distinct group within the genus *Campylobacter* based on their 16S rRNA and *hsp60* sequences and matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) profiles. The unique species status was further supported by phenotypic characteristics of the isolates. All isolates were found to be oxidase-, catalase- and urease-positive, they grew well at 37 °C and 42 °C and produced H_2S on TSI (triple-sugar iron) and SIM (sulfide indole motility) media. The name *Campylobacter corcagiensis* sp. nov. is proposed for this novel species, with the strain CIT 045[T] as the type strain CIT 045[T] (=LMG 27932[T], CCUG 64942[T]).

The genus *Campylobacter* has expanded considerably since it was initially proposed by Sebald & Veron (1963), at the time of writing encompassing 24 species and eight subspecies (On, 2013). Species of the genus *Campylobacter* inhabit a wide variety of ecological niches and have been isolated from various sample sources and hosts; including humans, birds, livestock, pets and primates (Man, 2011). Furthermore, several of these species have been associated with diseases in humans and animals.

Interestingly, even though some species of the genus *Campylobacter*, such as *Campylobacter jejuni*, have been isolated from a variety of hosts, including wild birds, domestic animals and humans, multilocus sequence typing analysis has revealed a wide genetic strain diversity of *C. jejuni* among different hosts (Griekspoor *et al.*, 2013). Furthermore, some strains have been found to be host-specific, being confined to a particular wild bird species (rather than geographical location). Other genetic populations have been found to be more widely distributed between different species and have been found to be present in humans and livestock (Griekspoor *et al.*, 2013). This heterogeneity highlights the possibility of trans-species transmission and the zoonotic potential of many species of the genus *Campylobacter*.

As a part of a larger study investigating possible environmental sources of *Campylobacter ureolyticus*, faecal samples were collected from captive wild animals. Samples were collected using sterile swabs which were immediately placed in a sterile 25 ml container filled with Bolton Broth (Oxoid) supplemented with (l^{-1}) 2 g sodium formate, 3 g sodium fumarate, 10 mg amphotericin B, 10 mg nalidixic acid and

20 mg vancomycin (all of which were supplied by Sigma Aldrich). Broths were incubated aerobically at 37 °C, closed tightly (to create anaerobic conditions) for a minimum of 7 days. Following pre-enrichment, samples were plated onto two types of media: NAV medium (O'Doherty *et al.*, 2014), containing Anaerobe Basal agar (Oxoid) with additional agar and the supplements listed above and NAV medium with additional 10 mg trimethoprim l^{-1} (Sigma Aldrich). Plates were incubated anaerobically with AnaeroGen 2.5 l gas packs (Oxoid) at 37 °C for up to 21 days and examined at regular intervals for putative *C. ureolyticus* colonies (flat, spreading, translucent colonies, positive for urease). One sample, collected from a lion-tailed macaque (*Macaca silenus*), was positive for presumptive *C. ureolyticus* colonies. Genomic DNA of this isolate (CIT 045[T]) was extracted using a Qiagen QIAmp DNA Mini kit. However, colony PCR gave a negative result for *C. ureolyticus* using previously described primers (Bullman *et al.*, 2011) but a positive result with *Campylobacter*-genus-specific C412F and C1288R primers (Linton *et al.*, 1996).

Subsequently, the 16S rRNA and *hsp60* gene sequences of the isolate CIT 045[T] were amplified using universal primers for 16S rRNA - fD1 and rP1 (Weisburg *et al.*, 1991) and *hsp60* - H729 and H730 (Hill *et al.*, 2006) genes and sequenced using fD1 and rP1 primers (Weisburg *et al.*, 1991) for 16S rRNA and M13(−40)F and M14(48)R primers (Hill *et al.*, 2006) for *hsp60* amplicons respectively. Comparison of these sequences to the sequences of other species of the genus *Campylobacter* available in the GenBank database using the BLASTN algorithm, revealed that the novel isolate shared ⩽95 % and ⩽82 % sequence similarity to a species of the genus *Campylobacter* with a validly published name for the 16S rRNA and *hsp60* gene sequences respectively. Two additional isolates of the novel species CIT 056 and CIT 058 were obtained from faeces of two additional lion-tailed macaques using a combination of molecular screening (using primers designed to target the 16S rRNA and *hsp60* sequences of the novel strain) and the culture method described above. All PCRs (25 µl final volume) consisted of 1 U HotStarTaq Plus (Qiagen), 1 µM each primer (Eurofins, MWG Operon), 0.2 µM each dNTPs (Sigma Aldrich) and 100 ng DNA template.

The primers used were 045-16SF (5′-CTTAGCGTAG-GGGTAAAATCCGTAGAGGTT-3′), 045-16SR (5′-CTA-CTTGCGTAGGCAGTCTGTTTAGAGAA-3′), 045-hsp60F (5′-GGATAAAGAAGTAGCTGCTATAATTAG-3′), 045-hsp60F (5′-CCGCTTCAAGAACCGGAAGAAGG-3′). These primer sets amplified a region of 499 bp (16S rRNA) and 359 bp (*hsp60*) under the following PCR conditions: initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 60 s, extension at 72 °C for 60 s and a final extension step at 72 °C for 10 min. *In silico* analysis of the 16S rRNA and *hsp60* gene sequences which are available for other species of the genus *Campylobacter* confirmed that the regions selected for primers are specific to the novel strains and are unlikely to bind to sequences from other species of the

genus *Campylobacter*. This was also confirmed by PCR on control strains: *C. jejuni* subsp. *jejuni* DSM 4688[T], *C. coli* DSM 4689[T], *C. lari* subsp. *lari* DSM 11375[T], *C. fetus* subsp. *fetus* DSM 5361[T], *C. hyointestinalis* subsp. *hyointestinalis* DSM 19053[T], *C. ureolyticus* DSM 20703[T] and *C. concisus* DSM 9716[T].

All three isolates belonging to the novel taxon had identical *hsp60* and 16S rRNA sequences (100 % sequence identity over the entire length of the sequence). Moreover, the RAPD analysis using (GTG)₅ primer (Matsheka *et al.*, 2006) performed on those three isolates showed a high level of genetic similarity (Fig. S1, available in the online Supplementary Material), indicating that the three isolates are possibly clones of a single strain, present in different members of the macaque colony at this site. However, it is worth noting that these isolates have different phenotypic characteristics, such as enzyme activity, antibiotic susceptibility and MALDI-TOF profiles, which implies some strain diversification in different hosts.

To investigate the taxonomic position of the novel species, phylogenetic trees were reconstructed for 16S rRNA (Fig. 1) and *hsp60* (Fig. S2) genes by aligning the sequences with MEGA 5.10 software (Tamura *et al.*, 2011) using the CLUSTAL W algorithm and the neighbour-joining method for tree building, with bootstrap analysis set for 1000 replicates.

The DNA G + C content of the type strain was derived from the whole genome sequence analysis. The strain was sequenced using Illumina MiSeq with 250 bp paired-end reads. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number JFAP00000000. The genome was assembled *de novo* using the Velvet (v.1.2.10) assembly tool (Zerbino & Birney, 2008). The DNA G + C content was estimated as 31.9 mol%, which is within the 29–47 mol% range reported for species of the genus *Campylobacter* (Debruyne *et al.*, 2008).

Given that it has been previously suggested that at a 16S rRNA similarity level of 97 %, or below, it is unlikely that two organisms have DNA–DNA hybridization level of more than 60–70 % and thus belong to the same species (Stackebrandt & Goebel, 1994); DNA–DNA hybridization experiments were not performed in this study since the level of 16S rRNA sequence similarity of the novel organism was ⩽95 % to any species of the genus *Campylobacter* with a validly published name.

A number of phenotypic characteristic of the novel species were determined using a wide range of previously described methods (On & Holmes, 1991; On & Holmes, 1995; Ursing *et al.*, 1994). An overview of those results and the comparison with the closely related species of the genus *Campylobacter* (on the basis of 16S rRNA sequence similarity) is presented in Table 1. The complete list of phenotypic characteristics determined for the novel species is provided in Table S1. Resistance to metronidazole (4 mg l^{-1}), nalidixic acid (32 mg l^{-1}), ciprofloxacin (32 mg l^{-1})
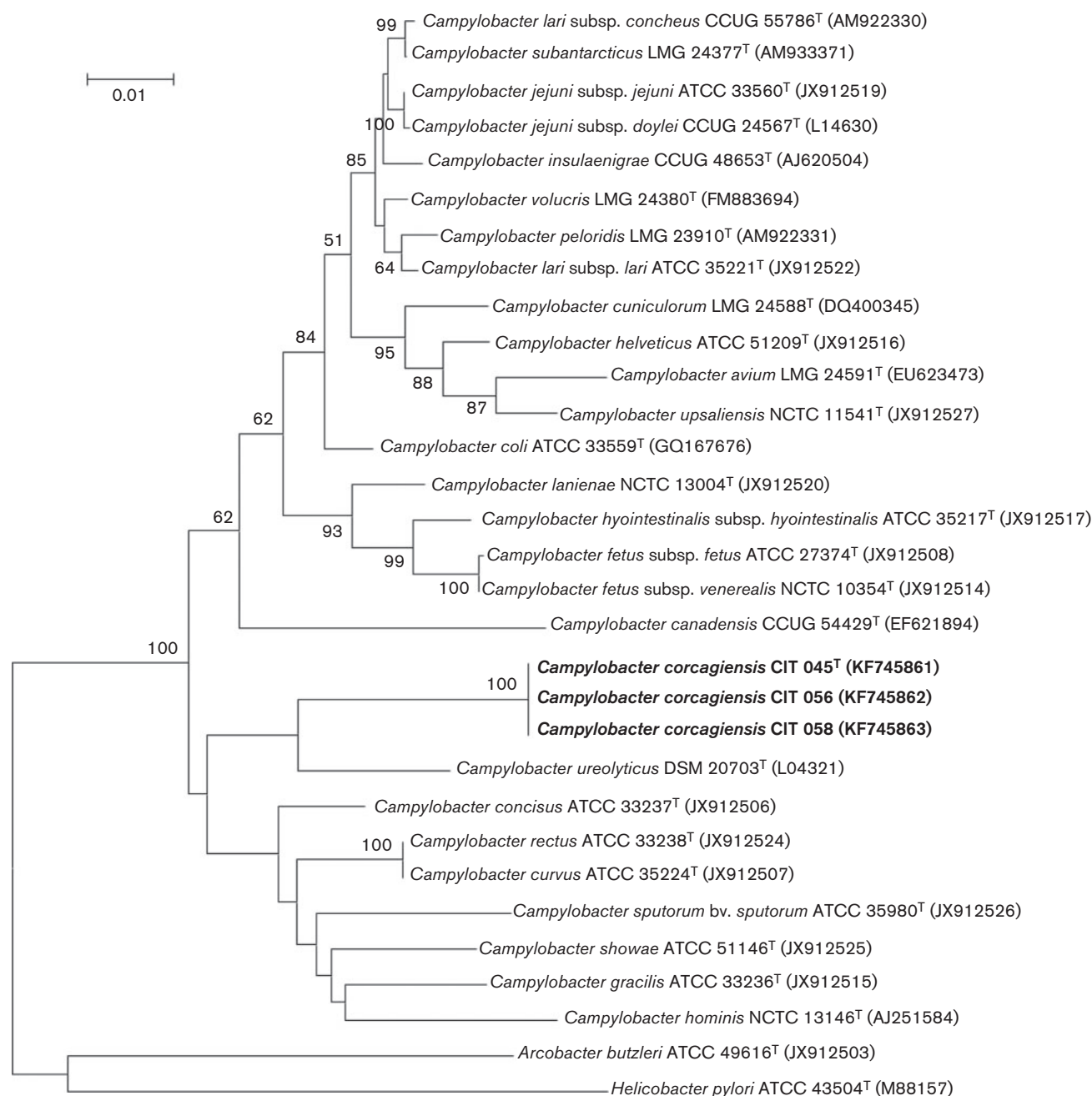
Fig. 1. Phylogenetic tree based on 16S rRNA gene sequences reconstructed by the neighbour-joining method. Bootstrap values (percentages) based on 1000 replicates are indicated at nodes. Bootstrap values lower than 50 % are not shown. Bar, 0.01 substitutions per nucleotide position.

and cefoperazone (64 mg l^{-1}) was also determined. The novel species differs from the majority of other species of the genus *Campylobacter* in possessing a urease enzyme and is phenotypically different from *C. ureolyticus* in its production of H$_2$S on TSI medium and its inability to hydrolyse gelatin (Vandamme *et al.*, 2010). Furthermore, unlike *C. ureolyticus* strains reported to date, this novel species shows resistance to nalidixic acid (32 mg l^{-1}).

The taxonomic divergence of the novel species was also supported by the distinct protein profiles generated using MALDI-TOF mass spectrometry which were compared with the profiles of type strains of other species of the genus *Campylobacter* available in the database and strains of its nearest phylogenetic neighbour *Campylobacter ureolyticus* (strains DSM 20703^T, CCUG 59468 and CCUG 59897). All strains were grown on Columbia Blood Agar (Oxoid)

Table 1. Phenotypic characteristics differentiating the novel strains from recognized, closely related species of the genus *Campylobacter*. The majority of phenotypic characteristics were determined with standardized methods, recommended for detection and characterization of species of the genus *Campylobacter* (On & Holmes, 1991, 1992; On & Holmes, 1995; On *et al.*, 1996; Ursing *et al.*, 1994)

Taxa: 1, *C. corcagiensis* sp. nov. (3 strains); 2, *C. concisus* (Zanoni *et al.*, 2009); 3, *C. curvus* (Lawson *et al.*, 2001); 4, *C. gracilis* (Lawson *et al.*, 2001); 5, *C. hominis* (Lawson *et al.*, 2001); 6, *C. rectus* (Debruyne *et al.*, 2010); 7, *C. showae* (Debruyne *et al.*, 2010); 8, *C. sputorum* (On *et al.*, 1998); 9, *C. ureolyticus* (Vandamme *et al.*, 2010). +, Positive; w, weakly positive; (+), 68–95 % of strains positive; v, 34–67 % strains positive; (−), 5–33 % strains positive; −, negative; NA, data not available; TSI, triple iron sugar medium; TTC, 2,3,5-triphenyltetrazolium chloride.

Characteristic	1	2	3	4	5	6	7	8	9
Catalase production	+	−	−	v	−	(−)	+	v	v
Urease production	+	−	−	−	−	−	−	v*	+
Alkaline phosphatase	+	v	v	−	−	−	−	−	−
Gelatin hydrolysis	−	NA	NA	NA	NA	NA	NA	NA	+
Hippurate hydrolysis	−	−	(−)	−	−	−	−	−	−
Indoxyl acetate hydrolysis	v	−	v	v	−	+	−	−	v
Nitrate reduction	(+)	(−)	+	(+)	−	+	+	+	−
Requirement for H_2	−	+	+	−	+	+	+	−	+
Production of H_2S (TSI)	+	−	(−)	−	−	−	v	+	−
Growth at/on/in the presence of:									
42 °C (microaerobic)	+	(+)	v	v	−	(−)	v	+	v
MacConkey agar	−	−	(+)	(+)	−	−	+	v	v
Nutrient agar	w	(−)	+	+	NA	(−)	v	+	+
NaCl (2 %)	+	(−)	v	v	+	v	+	+	+
Glycine (1 %)	+	(−)	+	+	+	+	v	+	+
Bile (1 %)	+	−	−	−	NA	−	−	v	v
TTC (0.04 %)	−	−	v	−	NA	−	−	−	−
Nalidixic acid (32 mg l^{-1})	+	(+)	+	v	v	(+)	−	(+)	−
Cefoperazone (64 mg l^{-1})	−	−	(+)	−	(−)	−	−	−	−
DNA G + C content (mol%)	31.9	37–41	45–46	44–46	32.5	45–46	44–46	29–32	28–30

*Strains belonging to *C. sputorum* biovar *paraureolyticus* are urease-positive, others are urease-negative.

supplemented with 5 % defibrinated horse blood (TCS Biosciences) and additionally with (l^{-1}) 2 g sodium formate and 3 g sodium fumarate for *C. ureolyticus* strains. The strains were incubated in anaerobic atmosphere using AnaeroGen 2.5 l gas packs at 37 °C for 72 h. All MALDI-TOF experiments were run in the Bruker MALDI Microflex Biotyper instrument. Proteins were extracted using the formic acid extraction method. Briefly, bacterial colonies were suspended in 300 µl deionized water and 900 µl absolute ethanol was added to the cell suspension and mixed by vortexing for 1 min. The cells were centrifuged for 2 min at 17 000 *g* and the supernatant was removed. The centrifugation step was repeated to remove any residual ethanol. The pellet was allowed to air dry in a class I safety cabinet and 50 µl of 70 % aqueous formic acid solution was added to the pellet. The solution was vortexed vigorously and incubated at room temperature for 2 min. Next, 50 µl of absolute acetonitrile solution was added to the suspension and mixed gently by pipetting it up and down. Protein extracts were centrifuged for 2 min at approximately 17 000 *g*. A 1 µl sample of the supernatant was spotted on the clean onto a mass spectra (MSP) (main spectra library) 96 target polished steel plate (Bruker

Daltonik). Samples were air-dried at room temperature and then immediately overlaid with 1 µl Bruker HCCA matrix solution (α-Cyano-4-hydroxycinnamic acid). The three novel strains were plated in 20 replicates, and culture collection strains of *C. ureolyticus* were plated in triplicate. Samples were air-dried again and mass spectra were generated with a Microflex mass spectrometer (Bruker Daltonics) using MALDI Biotyper Real-time Classification software, under the control of flexControl software (Version 3.4; Bruker Daltonics). For each measurement, 240 individual spectra (40 laser shots at six different spot positions) were collected. For identification, spectra of bacteria were matched with the spectra in the Bruker Daltonics database. A logarithmic score of 0 to 3.0 was assigned based on the peak matching as follows: a score of 0 to 1.699 indicated no reliable identification; scores of 1.700 to 1.999 indicated probable genus identification; scores of 2.000 to 2.299 indicated secure genus identification and probable species identification; and scores of 2.300 to 3.000 indicated highly probable species identification. No reliable identification was obtained for the novel strains using the database and the score values were ≤ 1.4. MSP were created – a single standard spectrum for each novel strain was generated from

20 raw spectra. The spectra for the isolates of the novel species were compared with spectra generated for *C. ureolyticus* strains DSM 20703[T], CCUG 59468 and CCUG 59897 and type strains of species of the genus *Campylobacter* available in the Bruker database and a MSP dendrogram was generated (Fig. S3.). The novel strains clustered together in the MSP dendrogram and their closest phylogenetic neighbour was shown to be *Campylobacter ureolyticus*. Full MALDI-TOF protein spectra of the novel species and phylogenetically related species are included in Fig. S4.

Using a polyphasic approach, we have shown that the three isolates from lion-tailed macaques represent a novel taxon within the genus *Campylobacter*. The novel species can be distinguished from other members of the genus by the analysis of the 16S rRNA (Fig. 1) and *hsp60* (Fig. S2) genes as well as biochemical characteristics (Table 1) and protein spectra using MALDI-TOF mass spectrometry (Figs. S3 and S4). On the basis of these findings, we propose classifying these isolates as strains of *Campylobacter corcagiensis* sp. nov. with strain CIT 045[T] (=LMG 27932[T], CCUG 64942[T]) as the type strain.

Description of the *Campylobacter corcagiensis* sp. nov.

Campylobacter corcagiensis (cor.ca.gien'sis N.L. masc. adj. *corcagiensis* pertaining to Corcagia, the Medieval Latin name for Cork, Ireland).

Cells are Gram-negative, motile, straight rods. All isolated strains are catalase, oxidase and urease-positive. All are alkaline phosphatase-positive. Some strains reduce nitrates and some hydrolyse indoxyl acetate. Hippurate and gelatin are not hydrolysed. All strains produce H_2S on TSI and SIM media. On blood agar, following 3 days incubation at 37 °C, the colonies are smooth, translucent in colour with convex centres and entire margins. Two types of colonies are observed: small, pinpoint colonies (1 mm in diameter) and flat spreading colonies (1–3 mm in diameter). No haemolysis on blood agar was observed. Growth is obtained on blood agar at 37 °C under anaerobic and microaerobic (weaker growth), but not aerobic conditions. Strains do not require H_2 to grow, but will grow in its presence. Strains grow anaerobically at 42 °C and 25 °C but not at 15 °C. Strains grow on blood agar medium supplemented with nalidixic acid, 2 % NaCl, 1 % glycine and 1 % bile. No growth is observed on media containing 0.04 % 2,3, 5-triphenyltetrazolium chloride and cefoperazone. Strain-dependent resistance to metronidazole is observed. Strain-dependent resistance to ciprofloxacin is also observed. Strains do not grow on MacConkey agar and growth on unsupplemented nutrient agar is weak. Pathogenicity has not been established to date.

The type strain is CIT 045[T] (=CCUG 64942[T], LMG 27932[T]), isolated from faeces of lion-tailed macaque (*Macaca silenus*) in February 2013. Strains CIT 056 (=LMG 27934, CCUG 64943) and CIT 058 (=LMG 27933, CCUG 64944) were isolated from other individuals of this species, from the same site.

References

Bullman, S., Corcoran, D., O'Leary, J., Lucey, B., Byrne, D. & Sleator, R. D. (2011). *Campylobacter ureolyticus*: an emerging gastrointestinal pathogen? *FEMS Immunol Med Microbiol* **61**, 228–230.

Debruyne, L., Gevers, D. & Vandamme, P. (2008). Taxonomy of the family Campylobacteraceae. In *Campylobacter*, pp. 3–26. Edited by I. Nachamkin, C. M. Szymanski & M. J. Blaser. Washington, DC: American Society for Microbiology.

Debruyne, L., Broman, T., Bergström, S., Olsen, B., On, S. L. & Vandamme, P. (2010). *Campylobacter volucris* sp. nov., isolated from black-headed gulls (*Larus ridibundus*). *Int J Syst Evol Microbiol* **60**, 1870–1875.

Griekspoor, P., Colles, F. M., McCarthy, N. D., Hansbro, P. M., Ashhurst-Smith, C., Olsen, B., Hasselquist, D., Maiden, M. C. & Waldenström, J. (2013). Marked host specificity and lack of phylogeographic population structure of *Campylobacter jejuni* in wild birds. *Mol Ecol* **22**, 1463–1472.

Hill, J. E., Paccagnella, A., Law, K., Melito, P. L., Woodward, D. L., Price, L., Leung, A. H., Ng, L. K., Hemmingsen, S. M. & Goh, S. H. (2006). Identification of *Campylobacter* spp. and discrimination from *Helicobacter* and *Arcobacter* spp. by direct sequencing of PCR-amplified *cpn60* sequences and comparison to cpnDB, a chaperonin reference sequence database. *J Med Microbiol* **55**, 393–399.

Lawson, A. J., On, S. L., Logan, J. M. & Stanley, J. (2001). *Campylobacter hominis* sp. nov., from the human gastrointestinal tract. *Int J Syst Evol Microbiol* **51**, 651–660.

Linton, D., Owen, R. J. & Stanley, J. (1996). Rapid identification by PCR of the genus *Campylobacter* and of five *Campylobacter* species enteropathogenic for man and animals. *Res Microbiol* **147**, 707–718.

Man, S. M. (2011). The clinical importance of emerging *Campylobacter* species. *Nat Rev Gastroenterol Hepatol* **8**, 669–685.

Matsheka, M. I., Lastovica, A. J., Zappe, H. & Elisha, B. G. (2006). The use of (GTG)$_5$ oligonucleotide as an RAPD primer to type *Campylobacter concisus*. *Lett Appl Microbiol* **42**, 600–605.

O'Doherty, A., Koziel, M., De Barra, L., Corcoran, D., Bullman, S., Lucey, B. & Sleator, R. D. (2014). Development of nalidixic acid amphotericin B vancomycin (NAV) medium for the isolation of *Campylobacter ureolyticus* from the stools of patients presenting with acute gastroenteritis. *Br J Biomed Sci* **71**, 6–12.

On, S. L. (2013). Isolation, identification and subtyping of *Campylobacter*: where to from here? *J Microbiol Methods* **95**, 3–7.

On, S. L. & Holmes, B. (1991). Effect of inoculum size on the phenotypic characterization of *Campylobacter* species. *J Clin Microbiol* **29**, 923–926.

On, S. L. & Holmes, B. (1992). Assessment of enzyme detection tests useful in identification of campylobacteria. *J Clin Microbiol* **30**, 746–749.

On, S. L. & Holmes, B. (1995). Classification and identification of campylobacters, helicobacters and allied taxa by numerical analysis of phenotypic characters. *Syst Appl Microbiol* **18**, 374–390.

On, S. L., Holmes, B. & Sackin, M. J. (1996). A probability matrix for the identification of campylobacters, helicobacters and allied taxa. *J Appl Bacteriol* **81**, 425–432.

On, S. L., Atabay, H. I., Corry, J. E., Harrington, C. S. & Vandamme, P. (1998). 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* **48**, 195–206.

Sebald, M. & Veron, M. (1963). [Base DNA content and classification of vibrios]. *Ann Inst Pasteur* (*Paris*) **105**, 897–910. (in French)

Stackebrandt, E. & Goebel, B. M. (1994). Taxonomic note: a place for DNA–DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44**, 846–849.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.

Ursing, J. B., Lior, H. & Owen, R. J. (1994). Proposal of minimal standards for describing new species of the family *Campylobacteraceae*. *Int J Syst Bacteriol* **44**, 842–845.

Vandamme, P., Debruyne, L., De Brandt, E. & Falsen, E. (2010). Reclassification of *Bacteroides ureolyticus* as *Campylobacter ureolyticus* comb. nov., and emended description of the genus *Campylobacter*. *Int J Syst Evol Microbiol* **60**, 2016–2022.

Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**, 697–703.

Zanoni, R. G., Debruyne, L., Rossi, M., Revez, J. & Vandamme, P. (2009). *Campylobacter cuniculorum* sp. nov., from rabbits. *Int J Syst Evol Microbiol* **59**, 1666–1671.

Zerbino, D. R. & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**, 821–829.