CAMPYLOBACTER VULPIS SP. NOV. ISOLATED FROM WILD RED FOXES

Antonio Parisi[a*], Matteo Chiara[b], Monica Caffara[c], Domenico Mion[c], William G. Miller[d], Marta Caruso[a], Caterina Manzari[e], Daniela Florio[c], Loredana Capozzi[a], Anna Maria D'Erchia[e,f], Viviana Manzulli[a], and Renato Giulio Zanoni[c]

[a] Experimental Zooprophylactic Institute of Apulia and Basilicata, Via Manfredonia 20, 71121 Foggia, Italy.

[b] Department of Biosciences, University of Milan, Via Celoria 26, 20133 Milano, Italy

[c] Department of Veterinary Medical Sciences, *Alma Mater Studiorum* Università di Bologna, Via Tolara di Sopra 50, 40064 Ozzano Emilia, Bologna, Italy.

[d] Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA, USA

[e] Institute of Biomembranes and Bioenergetics, Italian National Research Council, Via Amendola 165A, 70126 Bari, Italy

[f] Department of Biosciences, Biotechnology and Biopharmaceutics, University of Bari "A. Moro", Via Orabona, 4, 70126 Bari, Italy

*Corresponding author Antonio Parisi; Experimental Zooprophylactic Institute of Apulia and Basilicata, Via Manfredonia 20, 71121 Foggia, Italy Phone: +39 0804057858 ; Fax: +39 0804057753; E-mail: antonio.parisi@izspb.it

ABSTRACT

During a sampling of wild red foxes (*Vulpes vulpes*) for the detection of Epsilonproteobacteria, 14 strains were isolated from the caecal contents of 14 epidemiologically-unrelated animals. A genus-specific PCR indicated that the isolates belonged to the genus *Campylobacter*. Based on the results of a species-specific PCR, the isolates were initially identified as *C. upsaliensis*. However, multi-locus sequence typing (MLST) revealed that the isolates were significantly different from the *C. upsaliensis* present in the MLST database. A polyphasic study, including conventional biochemical and tolerance characteristics, morphology by transmission electron microscopy (TEM), MALDI-TOF analysis, and genetic comparisons based on partial 16S rDNA and *atpA* gene sequences, was undertaken. Finally, the complete genome sequence of the type strain 251/13^T and the draft genome sequences of the other isolates were determined. Average nucleotide identity, average amino acid identity and *in silico* DNA-DNA hybridization analyses confirmed that the isolates represent a novel taxon for which the name *Campylobacter vulpis* sp. nov. is proposed, with isolate 251/13^T (= CCUG 70587^T = LMG 30110^T) as the type strain. In order to allow a rapid discrimination of *C. vulpis* from the closely-related *C. upsaliensis*, a specific PCR test was designed, based on *atpA* gene sequences.

Keywords: *Campylobacter vulpis* sp. nov., polyphasic taxonomic study, red foxes, *Vulpes vulpes*.

INTRODUCTION

The genus *Campylobacter* currently contains 37 validly-described taxa. Many of these taxa have been recovered from a variety of terrestrial and marine mammals, reptiles, and both domesticated and wild birds [21]. *Campylobacter* strains are routinely isolated from the fecal samples and rectal swabs of domestic dogs, both healthy and diarrheic [1, 19]. One of the primary *Campylobacter* colonizers of domestic dogs is *Campylobacter upsaliensis* [1, 21]. The related species, *Campylobacter helveticus* [47], is also recovered from domestic dogs, although at a lower frequency [4, 14]. In addition to *C. upsaliensis* and *C. helveticus*, *C. jejuni, C. coli, C. lari* and *C. hyointestinalis* also colonize domestic dogs [19], along with at least nine other *Campylobacter* species (i.e., *C. showae, C. concisus, C. fetus, C. gracilis, C. mucosalis, C. rectus, C. sputorum, C. ureolyticus* and *C. volucris*) at a lower reported frequency [11, 19], suggesting that domestic dogs are host to a wide variety of campylobacters. Several of these emerging campylobacters have also been isolated from human clinical samples [1, 21, 29]. Thus, their presence in domestic dogs is a potential human health concern due to the close contact of dogs and humans. Most of the reported canine-associated epidemiology of *Campylobacter* pertains to pet and sheltered animals. Therefore, data on the prevalence of *Campylobacter* within wild canid populations is limited, although *C. upsaliensis* has been recovered from coyotes [33].

In 2013, a survey on the prevalence of *Campylobacter* spp. was carried out on wild red foxes (*Vulpes vulpes*) shot during population control programs in different municipalities of the Province of Bologna. Fourteen isolates of Gram-negative, spiral-shaped cells resembling members of the genus *Campylobacter* were recovered from the caecal contents of an equal number of epidemiologically-unrelated wild red foxes. In this study, these isolates were characterized through biochemical tests and by genomic analyses and were determined to comprise a distinct and novel taxon within the genus *Campylobacter*.

MATERIALS AND METHODS

Isolation and culturing

From January to July 2013 the caecal contents collected from a total of 45 wild red foxes (26 females and 19 males, 12 subadults <1 year and 33 adults) were examined for *Campylobacter* spp. and *Helicobacter* spp. Foxes were shot during population control programs in the province of Bologna. For this survey, only animals dead ≤12 hours and kept at refrigeration temperature until necropsy were selected. Bacterial strain isolations were performed on Nutrient sheep-blood agar [Nutrient Broth No 2 (Oxoid) amended with 1.5% (w/v) Bacto Agar (Difco) and 5% (v/v) sheep blood], using the filter technique of Steele and McDermott that was modified as previously described [52], and on selective media [Campylobacter Selective Agar (Skirrow) and Campylobacter CAT Agar (Oxoid)]. Plates were incubated at 37 ± 1°C for 72 h in a microaerobic atmosphere, with hydrogen, obtained by the gas replacement method, using an anaerobic gas mixture (10% H_2, 10% CO_2, 80% N_2) as described [5]. Pure cultures were obtained after dilution and repeated sub-culturing onto plates of Nutrient sheep-blood agar never older than 7 days.

Preliminary taxonomic identification and MLST typing

Genomic DNA was purified using the REDExtract-N-Amp Tissue PCR Kit (Sigma-Aldrich, St. Louis, MO, USA) and amplified using a *Campylobacter* genus-specific PCR [28] and a species-specific PCR for *C. upsaliensis* and *C. helveticus* as previously described [27]. Based on a preliminary identification of *C. upsaliensis*, the isolates were typed by multi-locus sequence typing (MLST) according to published protocols [31], with the PCR amplifications performed as described elsewhere (https://pubmlst.org/organisms/campylobacter-non-jejunicoli/c-upsaliensis-primers). DNA sequences of the seven genes used for the MLST

analysis were assembled using BioNumerics 7.6 software (Applied Maths, Belgium). Novel alleles and sequence types (STs) were assigned following submission of the DNA sequences and profiles to the non *jejuni/coli Campylobacter* PubMLST database (http://pubmlst.org/campylobacter/).

Concatenated sequences representing all *C. upsaliensis* sequence types within the *Campylobacter* PubMLST database were downloaded from PubMLST and combined with sequences representing the profiles from this study. A concatenated sequence based on the *C. upsaliensis* MLST gene get was also extracted from the type strain genome of *C. helveticus.* These sequences were imported into BioNumerics and aligned using the Fast algorithm. Within BioNumerics, a Neighbor-joining dendrogram was constructed from the aligned profile sequences. A minimum spanning tree (MST) was constructed based on the sequence distances between the concatenated profile sequences and using the default priority rules, 'Permutation resampling' resampling strategy and 'Highscore summary' methods. MST nodes were color-coded within BioNumerics according to taxon.

Biochemical and microscopic characterization

The 14 isolates were tested for the following phenotypic traits: oxidase, catalase, urease and alkaline phosphatase activity; hydrolysis of hippurate and indoxyl acetate; reduction of nitrate, selenite and 2,3,5-triphenyltetrazolium chloride (TTC); H_2S production on triple sugar iron agar; α-haemolysis; growth at 25 °C, 30 °C, 37 °C and 42 °C; growth under aerobic, microaerobic and anaerobic conditions; growth on CCDA and on media amended with 1% (w/v) glycine, 2% (w/v) NaCl or 1% (w/v) bile. Resistance to nalidixic acid and cephalothin was also determined. All tests were performed using standard methods as previously described [35-39] and were conducted at least twice.

Morphological characteristics of three of the novel strains were determined using transmission electron microscopy. Cells grown on Nutrient Broth No.2 (Oxoid) with 1.5% of Bacto Agar (Difco) and 10% of sheep blood for 48 h were gently suspended in 0.1 M phosphate-buffered saline (PBS) at a concentration of about 108 cells per ml. Samples were negatively stained with 2% (w/v) phosphotungstic acid (pH 6.5) for 30 s. The specimens were then examined using a Philips EM208S TEM.

MALDI-TOF MS

A Microflex LT/SH MALDI-TOF MS (Bruker Daltonics, Germany) with a 60-Hz nitrogen laser was used to analyze spectra over a mass range of 2,000-20,000 Da. All the isolates and *C. upsaliensis* ATCC 43954^T were processed according to the manufacturer's instructions. A sterile wooden tip was used to pick an isolated bacterial colony freshly grown on Mueller-Hinton agar (Oxoid) supplemented with 5% sheep blood for 48 h at 37 °C under microaerobic conditions. A thin film in five replicates was then smeared onto a MALDI 96-target polished steel plate (Direct Transfer procedure). Microbial films were overlaid directly with 1.0 µl of a 10 mg/ml α-cyano-4-hydroxycinnamic acid (MALDI-TOF HCCA) matrix solution. The sample-matrix mixture was dried at room temperature and subsequently inserted into the system for data acquisition. The spectra were constructed based on processed raw spectra after smoothing, baseline subtraction, normalization, and peak picking. For each spot, 50 sub-spectra for each of 40 randomized positions within the spot were collected and presented as one main spectrum. A sum spectrum was acquired by summing the laser shots. Quality controls were internally calibrated using the *Escherichia coli* DH5α supplied by Bruker Daltonics, following the same procedure. The data were processed automatically by the instrument software, and the spectra were compared with reference libraries for bacterial identification matching. Spectra were analyzed using MBT Compass 4.1.70.1 database

version 7.0.0.0 (Bruker Daltonics, Germany). Manufacturer-recommended score cut-offs were used to determine genus level (1.7000 to 1.999) or species level (≥2.000) of the organism. A score of <1.7 was considered unreliable for genus identification. Further analyses were performed using the ClinProTools 3.0 software package (Bruker Daltonics, Germany). For all the characteristic proteins, the results of the following statistical tests were taken into consideration: Anderson-Darling, Wilcoxon, t-test, Variation Coefficient and ROC curves.

rDNA, atpA *and core gene phylogenetic analysis*

The 16S rDNA and *atpA* genes were amplified by the universal internal primers p27f and p1492r [7] and atpAFC3 and atpARC4 [33], respectively, and then sequenced with an ABI 3730 DNA analyzer at StarSEQ GmbH (Mainz, Germany). For each gene set, the sequences from the 14 strains were combined with gene sequences extracted from the *Campylobacter* type strain genomes (defined hereafter as the genomes of all validly-described *Campylobacter* species type strains, the genome of *C. portucalensis* sp. nov. strain LMG 31504^T [46], and the genomes of *C. aviculae* sp. nov. strain LMG 31272^T, *C. estrildidarum* sp. nov. strain LMG 31271^T and *C. taeniopygiae* sp. nov. strain LMG 30935^T [8]) and the *Helicobacter pylori* type strain genome (as an outgroup) and aligned using Clustal X [25]. Neighbor-Joining trees were generated from these alignments in MEGA6 [48], using the Kimura 2-parameter method.

Additionally, the sequences of 20 core genes (*aroC*, *atpA*, *dnaN*, *eno*, *fabH*, *frr*, *glnA*, *groEL*, *hemB*, *ileS*, *lpxA*, *miaB*, *mrp*, *nrdB*, *pnp*, *prfA*, *queA*, *speA*, *spoT* and *tkt*) were extracted from the genomes of the 14 strains, the *Campylobacter* type strain genomes, and the *Helicobacter pylori* type strain genome (as an outgroup). The genes were aligned individually using MUSCLE in Geneious Prime (v. 2020.0.5) with default parameters. The 20 alignments were

then concatenated alphabetically and a Neighbor-Joining dendrogram was constructed in Geneious Prime using the Tamura-Nei genetic distance model.

Whole genome sequencing and annotation

The complete genome sequence of the type strain 251/13^T was determined using PacBio sequencing; draft genome sequences of the remaining strains were obtained using Illumina sequencing. Illumina sequencing was also performed on the type strain to improve PacBio base calling. For all 14 strains, genomic DNAs were prepared from loopfuls of cells using the Promega Wizard Genomic DNA Purification Kit. The same genomic DNA preparation was used to construct the Illumina and PacBio libraries for the type strain. Illumina libraries were prepared using the Illumina Nextera DNA Flex Library Prep Kit with 20ng of genomic DNA. Pooled libraries were sequenced on a MiSeq instrument at 8.0 pM, with dual index reads, paired end, using the MiSeq Reagent Kit v2 (300-cycles). For the type strain, a 20 kb PacBio library was prepared using 15 µg of genomic DNA, the PacBio SMRTbell Template Prep Kit 1.0, and manufacturer's protocols. SMRT sequencing was performed on an RSII sequencer. PacBio reads were assembled using the Hierarchical Genome Assembly Process (HGAP ver. 3.0) in the SMRT Analysis software (ver. 2.3.0). Sequencing metrics for all 14 strains are presented in Supplementary Table 1. PacBio sequencing of strain 251/13^T resulted in two contigs: one representing the chromosome and another representing a 36,943 bp plasmid, termed pVULP. Both contigs were circularized manually within Geneious Prime (ver. 2019.1.3; Biomatters Ltd., Auckland, New Zealand), and base calling was further improved by mapping the MiSeq reads onto the circularized PacBio contigs within Geneious Prime. Using the Geneious "Find Variations/SNPs" module, with a default minimum variation of 0.3, genomic variations in each contig were identified and corrected to the MiSeq consensus sequence.

171 Annotation of the type strain genome was performed manually. Putative coding sequences,
172 ribosomal loci and tRNA/transfer-messenger RNA (tmRNA) genes were identified using
173 GeneMark, RNAmmer and ARAGORN [3, 24, 26] respectively. The start point of each
174 coding sequence was curated manually within Artemis (v.16) [44]. Gene function was
175 initially assigned following a BLASTP comparison of the 251/13^T proteome against the
176 proteomes derived from the other annotated *Campylobacter* type strains. Annotation was
177 further refined by identification of Pfam motifs [13] and a BLASTP analysis that utilized the
178 larger NCBI non-redundant (nr) protein database. Annotation of the thirteen draft genomes
179 was performed by means of the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [49]
180 upon submission of the genomes to GenBank.
181
182 *Whole-genome analyses*
183 Analyses based on average nucleotide and amino acid identities and *in silico* DNA-DNA
184 hybridization were used to determine if the new isolates belonged to any of the previously-
185 characterized species of *Campylobacter*. The average nucleotide identity based on BLAST
186 (ANIb) analysis was performed using JSpecies (v. 1.2.1) [42]. The data set analyzed included
187 the complete 251/13^T genome and the *Campylobacter* type strain genomes. The proteomes of
188 the same taxa used in the ANI analysis were compiled into a single FASTA-formatted 72,437
189 protein amino acid sequence file. An all-against-all BLASTP [2] analysis was performed
190 using this composite proteome file. The BLASTP output was subsequently analyzed with
191 BlastPTrimmer14 [34] using 35% sequence similarity and 75% alignment length as input
192 parameters, which identified the proteins conserved among the 43 taxa and generated pair-
193 wise average amino-acid identity (AAI) values using the core protein set. In silico DNA-
194 DNA hybridization (isDDH) values were determined using the Genome-to-Genome Distance
195 Calculator 2.1 (GGDC; https://ggdc.dsmz.de/ggdc.php#; accessed Sept 2020) [30]. ANI, AAI

and isDDH analyses were also performed on a 28-strain set that included an additional 25 *C. vulpis*, *C. upsaliensis* and *C. helveticus* strains. In addition to the three complete *C. vulpis*, *C. upsaliensis* and *C. helveticus* type strain genomes and the thirteen *C. vulpis* draft genomes sequenced here, ten *C. upsaliensis* genomes and two *C. helveticus* genomes were downloaded from NCBI. ANIb, AAI and isDDH analyses were performed as described above, using a FASTA-formatted 48,744 protein amino acid sequence file for the AAI analysis. Whole-genome based average nucleotide identity (gANI) analysis was performed using the JGI Microbial Species Identifier (MiSI; https://ani.jgi.doe.gov/html/home.php?; accessed Mar 2021 [51]) and FASTA-formatted coding sequence files. The second isDDH analysis used a mixture of draft and complete genomes; therefore, Formula 2 (recommended for draft genomes as it is independent of genome length) was used throughout both isDDH analyses to maintain consistency.

Clusters of Orthologous Genes (COGS) and Rarefaction Analyses

Clusters of reciprocal best BLASTP matches [putative Clusters of Orthologous Genes (COGs)] were established using all-against-all BLASTP searches, which employed the BLOSUM80 matrix and accepted only the best reciprocal matches with E-values $\leq 1\times10^{-5}$, where "second-best" matches produced bit scores < 90% of those associated with the best matches, in order to avoid confounding effects deriving from recent duplications and paralogous genes. The estimated size of the core and accessory genome, based on the COGs identified by the method described above, were recorded for all possible combinations of the 2 to 14 genomes included in this study. The core genome of the genus *Campylobacter* was determined by applying the same method to a collection of 38 *Campylobacter* type strain proteomes, with the inclusion of 251/13^{T} as the type strain of the putative new species. Plots were prepared showing mean and standard deviation of these combinations. The scripts used

for the identification of COGs and execution of rarefaction analysis for the estimation of the size of the core and accessory genomes of the studied isolates are available at https://github.com/cvulpispaper/compute_aai_and_cogs.

Identification of "plasmid-like" sequences

The complete collection of plasmid sequences included in the PLSDB database [16] was retrieved from https://ccb-microbe.cs.uni-saarland.de/plsdb/plasmids/download/?zip (accessed Sept 2020). Sequence similarity searches against the PLSDB were performed using BLASTN, as available from the BLAST+ suite [10].

Species-specific PCR

In order to define a diagnostic method for the rapid detection and identification of the putative new species, specific PCR primers targeting the *atpA* gene were designed: atp_FW18 5'-TGC CGC TTT ACA ATA TCT CGC T-3' and atp_Rev405 5'-CCC CAC ACG CGA AAC AGA CAA G-3'. Amplification parameters were as follows: 30 cycles of 30 s at 94 °C, 30 s at 58 °C and 60 s at 72 °C, preceded by a denaturation step at 95 °C for 5 min and followed by an extended elongation step at 72 °C for 7 min. To assess the sensitivity and specificity of this PCR method, all 14 strains and a selection of strains representing different species of *Campylobacter* were analyzed using the specified conditions.

Data availability

The GenBank accession numbers for the 16S rRNA gene and *atpA* sequences of strains 36/13, 43/13, 52/13, 73/13, 99/13, 108/13, 131/13, 146/13, 147/13, 154/13, 166/13, 227/13, 250/13 and 251/13^{T} are KU855032-KU855045 (16S) and KU855018-KU855031 (*atpA*). The complete sequence of the type strain 251/13^{T} has been deposited in GenBank [CP041617

(chromosome) and CP041618 (megaplasmid pVULP)]. The draft genome sequences of the other strains have been deposited in GenBank with the following accession numbers: VJYO00000000 (36/13), VJYW00000000 (43/13), VJYU00000000 (52/13), VJYY00000000 (73/13), VJYM00000000 (99/13), VJYT00000000 (108/13), VJYV00000000 (131/13), VJYN00000000 (146/13), VJYS00000000 (147/13), VJYX00000000 (154/13), VJYQ00000000 (166/13), VJYP00000000 (227/13), VJYR00000000 (250/13).

RESULTS AND DISCUSSION

Morphology and phenotypic characterization

After 48 h of incubation at 37 ± 1°C, colonies appear 2-3 mm in diameter, α-haemolytic, grey, translucent, flat with an irregular edge, and show a tendency to spread along the direction of the streak and to swarm and coalesce. Cells are Gram-negative, sigmoid to allantoid in shape, 0.3-0.4 µm in width and 1.2-3.0 µm in length with one bipolar unsheathed flagellum (Fig. 1), and motile with characteristic darting movements when observed by dark field microscopy. Cells appear coccoid after 5-6 days of incubation or when exposed to air.

All 14 isolates from red foxes were identified initially as *C. upsaliensis* according to the results of a *Campylobacter* genus-specific PCR and a species-specific PCR for *C. upsaliensis* and *C. helveticus*. Thus, it was not unexpected that results of the standard biochemical and growth tests showed a strong similarity between the composite phenotypic profile observed for the 14 fox isolates and the phenotypic profile reported previously for *C. upsaliensis* (Table 1). Nevertheless, the 14 fox isolates could be unambiguously distinguished from *C. upsaliensis* by their inability to grow at 30 °C under microaerobic conditions (Table 1). Additionally, alkaline phosphatase activity is variable within our strain set but negative in *C. upsaliensis*, and TTC reduction is negative in our strain set but variable in *C. upsaliensis*.

Although the profile from our strain set is also highly similar to the profile reported for *C. helveticus*, they are clearly distinguished from *C. helveticus* by their ability to reduce selenite (Table 1).

MALDI-TOF analysis

We determined if our 14 fox isolates could also be distinguished from *C. upsaliensis* using MALDI-TOF MS. MALDI-TOF MS has been used in previous descriptions of novel *Campylobacter* taxa; *Campylobacter fetus* subsp. *testudinum* could not be distinguished from *C. fetus* subsp. *fetus* or *C. fetus* subsp. *venerealis* by standard phenotypic tests but were clearly distinct from those subspecies based on MALDI-TOF MS profiles [15], and MALDI-TOF MS analysis was used as additional supporting evidence in the description of *Campylobacter armoricus* [6].

MALDI-TOF MS spectra were generated for the 14 strains. The result "*No Organism Identification Possible*" was obtained by comparing these experimental mass spectra against the data contained in the Compass database used in this study. The average mass spectra of all acquired experimental data for *C. upsaliensis* NCTC 11541^{T} and the 14 isolates overlap across the entire mass range, although some differences are noticeable in the protein profiles. In the spectral regions between *m/z* 8200-8600 and *m/z* 10180-10260, there are characteristic peaks for both species (Supplementary Fig. 1). Specific peaks at *m/z* values of 8376, 8424 and 10201 for *C. upsaliensis* NCTC 11541^{T} and *m/z* values of 8510 and 10227 for two fox strains (251/13^{T}; 73/13), as monoprotonated molecules, were identified (Supplementary Fig. 1). The same proteins are revealed in the mass spectra as doubly-charged ions, as demonstrated by the replication of the same profiles in the spectral range between *m/z* 4180-4270 and 5095-5120 (Supplementary Fig. 1). The results of the Variation Coefficient tests, in terms of standard deviation of intensity values, and the ROC curves, in terms of Area Under

296 the Curve (AUC), reveal a statistical significance for the peaks under study. Finally, main
297 spectra library (MSP) phylogenetic analysis based on the average mass spectra of all acquired
298 experimental data for the fox isolates formed a clade distinct from the other *Campylobacter*
299 taxa present in the MBT Compass software database (Supplementary Fig. 2); this clade is
300 also clearly distinct from the clade containing *C. upsaliensis*. Thus, the proteomic data
301 indicate that the isolates under study, hereafter named *C. vulpis*, exhibit a specific protein
302 pattern that could also be used to discriminate them from the closely-related *C. upsaliensis*.
303

304 C. vulpis *multi-locus sequence typing*

305 MLST was performed on the 14 *C. vulpis* strains. For each of the seven MLST genes and for
306 all 14 isolates, the MLST alleles identified here are notably different, when compared to the
307 *C. upsaliensis* alleles present in the *Campylobacter* MLST database, and all the identified
308 alleles and STs are novel. Nine sequence types were identified among the 14 *C. vulpis* strains,
309 indicating genotypic diversity within the strain set. A minimum spanning tree was created
310 from the MLST profile sequences of *C. vulpis*, the type strain of *C. helveticus*, and all *C.*
311 *upsaliensis* sequence types currently present within PubMLST. Within this tree, the *C. vulpis*
312 isolates group together and are clearly separated from *C. upsaliensis* (Fig. 2).
313

314 *16S rDNA,* atpA *and core gene phylogenetic analysis*

315 Alignment of the 16S rDNA sequences from *C. vulpis* and the *C. upsaliensis* type strain
316 indicated identity (0% sequence distance) between the *C. vulpis* strains, with near identity
317 (99.8% similarity) to the 16S rDNA sequence of the *C. upsaliensis* type strain. This is
318 reflected in the dendrogram constructed from the aligned sequences (Fig. 3). However,
319 similarity values between the *C. vulpis* 16S rDNA loci and those of five additional *C.*

upsaliensis strains range from 98.5% to 99.8% (data not shown). High similarity between the 16S rDNA loci of related taxa has been reported previously in *Campylobacter*; the 16S rDNA sequences of three novel species recovered from Zebra finches are 99% similar [8]. Nevertheless, the neighbor-joining tree demonstrates that our strains and the *C. upsaliensis* type strain form a well-supported (92% bootstrap) clade that is separate from the rest of the *Campylobacter* species included in the dataset (Fig. 3).

To further investigate the taxonomic position of *C. vulpis,* we used *atpA* sequencing, which has been shown to have a higher resolving power than 16S-based analyses in *Campylobacter* [33].

In this analysis, the *C. vulpis* strains form a clade that is clearly distinct from the other *Campylobacter* taxa, including *C. upsaliensis* (Fig. 4). A phylogenetic tree was also constructed using the concatenated sequences of 20 core genes (Supplementary Fig. 3). Consistent with the results of the *atpA* analysis, *C. vulpis* forms a clade that is distinct from the other *Campylobacter* taxa.

Whole genome sequencing and analysis

The genome of the *C. vulpis* type strain 251/13 was sequenced to completion using a combination of PacBio and Illumina MiSeq sequencing. The genome of 251/13^T is 1,645 kbp with a G+C content of 34.7% (Supplementary Table 1). The 251/13^T genome putatively encodes 1564 genes, 52 pseudogenes and 3 ribosomal loci, although 2 of these ribosomal loci are split, as observed also in the related species *C. helveticus* [32]. The 251/13^T genome contains three genomic islands (43, 33, and 19 kbp) and no CRISPR/Cas loci. A 36,943 bp plasmid, termed pVULP, was also identified in the 251/13^T genome; this plasmid is presumably conjugative, due to the presence of genes encoding a P-type type IV conjugative transfer system. The other thirteen genomes were sequenced to draft level by Illumina MiSeq

sequencing (Supplementary Table 1). The approximate sizes of these genomes range from 1,574 to 1,724 kbp with G+C contents ranging from 34.4 to 34.7%. These %G+C values are similar to those reported for *C. upsaliensis* (i.e., 34-35%) and are within the range reported for other members of the genus *Campylobacter* (29–47 %) [12]. BLASTN analysis of the *C. vulpis* draft genomes, using pVULP as a query sequence, indicated the potential presence of plasmids in nine of the thirteen strains (Supplementary Table 1); these results were confirmed by BLAST sequence similarity searches against the PLSDB database, a specialized database containing the complete sequences of more than 18,000 plasmids. However, it is possible that these plasmid sequences in the draft genomes are located in genomic islands rather than on extrachromosomal elements; in addition to pVULP, P-type type IV conjugative transfer system encoding genes are also located within the 19 and 33 kbp genomic islands on the 251/13^{T} chromosome. Therefore, determination of the plasmid content of the draft genomes may require genome closure.

Rarefaction analysis based on the *C. vulpis* genomes sequenced in this study (Fig. 5), indicates that the core genome of *C. vulpis* is 1176 genes and accounts for more than 70% of the average *C. vulpis* gene content. Furthermore, the core genome curve illustrated in Figure 5 suggests that the size of the core genome is not likely to be affected substantially by the sequencing of additional genomes. However, the accessory genome presently contains 1260 genes. This is clearly a minimum value (Fig. 5), suggesting that currently available data offer only a partial representation of the pan-genome of *C. vulpis*.

Whole genome comparisons

Average nucleotide identity based on BLAST (ANIb) analysis indicated that the *C. vulpis* type strain showed the highest levels of sequence identity with the *C. upsaliensis* type strain, with an ANIb value of 89.9 (Supplementary Table 2A). Similar values were observed for the

other 13 strains, when compared to the *C. upsaliensis* type strain, with an average ANIb value of 90.1% (Table 2, Supplementary Table 2B). The average ANIb value between *C. vulpis* strains was 99.2% (Supplementary Table 2B). Importantly, ANIb values determined for every isolate considered in this study are well below the 94-95% cut-off that has been recently recommended for the discrimination of bacterial species [18, 23, 42]. Moreover, the average pairwise *C. vulpis/C. upsaliensis*^{T} ANIb value is below the ANIb values observed for all currently-recognized *Campylobacter* subspecies pairs: *C. jejuni* subspp. *jejuni* and *doylei* (96.1%); *C. lari* subspp. *lari* and *concheus* (93.2%); *C. fetus* subspp. *fetus/venerealis* and *testudinum* (91.8%); *C. hyointestinalis* subspp. *hyointestinalis* and *lawsonii* (94.2%); and *C. pinnipediorum* subspp. *caledonicus* and *pinnipediorum* (94.3%) (Supplementary Table 2A). This average pairwise ANIb value is also below that observed between *C. lari* subsp. *concheus* and *C. ornithocola* (91.0%; Supplementary Table 2A).

To further characterize the average nucleotide identities within *C. vulpis, C. upsaliensis* and *C. helveticus*, we used whole-genome based average nucleotide identity (gANI) analysis in addition to ANIb on an expanded strain set that included all 14 *C. vulpis* strains, 11 *C. upsaliensis* strains and 3 *C. helveticus* strains. Each method calculates the ANI in a different fashion: ANIb cuts the query genome into 1020 nt fragments which are then used to search against the reference genome using BLASTN [18], whereas gANI utilizes FASTA-formatted coding sequence files for each genome, which are then compared using BLASTN [23]. Although, in general, the gANI value was slightly higher than the ANIb value for each strain pair, the two ANI methods yielded very similar results (Supplementary Table 2B). Interestingly, ANI analysis within *C. upsaliensis* identified two strains, RM3195 and NCTC 12264, that are more divergent than the other nine strains, with intergroup ANI values of ~96.6% (Supplementary Table 2B). Nevertheless, despite this variation within *C. upsaliensis*, both ANI methods yielded *C. vulpis* vs. *C. upsaliensis* values [90.1 ± 0.2 (ANIb); 90.4 ± 0.2

(gANI)] well below the 94-95% ANI species boundary (Supplementary Table 2B). Palmer et al. [40] concluded that the various described ANI methods do not completely correlate within the region of 90-100% similarity, a region that includes the proposed ANI species boundary. This lack of correlation would be more critical if the calculated ANI values were in close proximity to the species boundary, i.e. ~94-95%. However, the *C. vulpis* vs. *C. upsaliensis* ANI values determined here, which are ~4% below the species boundary using either ANI method, strongly suggest that the *C. vulpis* and *C. upsaliensis* strains comprise two distinct *Campylobacter* species, a conclusion that is likely ANI method independent.

Average amino acid identity (AAI) analysis, based on the determination of pairwise similarities among the *Campylobacter* core protein set, indicated that the *C. vulpis* core proteins are most similar to those of the *C. upsaliensis* type strain, with an average AAI of 95% (Table 2). *Campylobacter vulpis* core proteins are also similar to those of *C. helveticus* (AAI = 91%; Table 2). AAI values between *C. vulpis* and the remaining *Campylobacter* taxa were ≤ 81% (Table 2, Supplementary Table 2A). AAI analysis using the expanded strain set described above yielded similar results. The average pairwise AAI value between the 14 *C. vulpis* strains and the 11 *C. upsaliensis* strains was 94.6% (Supplementary Table 2C), which is below the intraspecific *C. vulpis* and *C. upsaliensis* values of 99.6% and 98.8%, respectively (Supplementary Table 2C). Here also, AAI values between *C. vulpis* and *C. upsaliensis* are below those observed within established *Campylobacter* subspecies pairs (Supplementary Table 2A). Additionally, the AAI values are similar to those observed between some *Campylobacter* species [e.g., *C. rectus* and *C. showae* (95%), and *C. subantarcticus* and *C. lari* (95%); Supplementary Table 2A].

The estimated in silico DNA-DNA hybridization (isDDH) value between the *C. upsaliensis* type strain genome and the *C. vulpis* type strain genome is 40.1% (CI [37.7 – 42.7%]) (Supplementary Table 3). Analysis of the expanded strain set yielded a similar average value

(40.2%; Supplementary Table 2C). isDDH analysis using the 28-strain set is complicated by the mixture of draft and complete genomes. Formulas 2 and 3 on the GGDC service divide the high-scoring segment pairs (HSPs) by HSP length or genome length, respectively. Thus, Formula 2 is recommended by the GGDC service for draft genomes of uncertain genome length, and we utilized Formula 2 throughout this study to maintain consistency. Notably, the variation within *C. upsaliensis* described above is also observed here, with isDDH values of ~70% when strains RM3195 and NCTC 12264 were compared to the other nine *C. upsaliensis* strains (Supplementary Table 2C). However, as with the ANI results, this variation within *C. upsaliensis* did not substantially affect the pairwise *C. vulpis/C. upsaliensis* isDDH values, although these data may indicate that further research on *C. upsaliensis* taxonomy may be warranted. Although these isDDH data support the description of *C. vulpis* as a novel *Campylobacter* species, the presence of incomplete genomes within the data set suggest that isDDH analysis should not be used as the sole taxonomic determinant.

C. vulpis-*specific PCR*

To specifically amplify *C. vulpis* strains, a novel primer pair was designed that targets that *atpA* gene. Amplification of all 14 *C. vulpis* isolates produced a 409 bp PCR product, whereas the other *Campylobacter* species were PCR-negative. Additionally, the 39 *C. upsaliensis* strains characterized by Rossi et al. [43] from domestic dogs and cats were also PCR-negative. DNA sequencing of the PCR products from *C. vulpis* confirmed the specificity of the amplicons. This suggests that the *C. vulpis*-specific PCR is robust and can be used to rapidly discriminate between *C. vulpis* and *C. upsaliensis* strains.

CONCLUSIONS

Fourteen *Campylobacter* isolates were recovered from the cecal contents of wild red foxes in Northern Italy in 2013. These strains were identified initially as *C. upsaliensis* by *Campylobacter* species-specific PCR; however, MLST placed these isolates in a clade separate from *C. upsaliensis.* MLST also identified nine sequence types within the 14 isolates, indicating genomic variation within the strain set. Although phenotypically similar to both *C. upsaliensis* and *C. helveticus,* these 14 fox isolates could be distinguished from *C. upsaliensis* and *C. helveticus* by microaerobic growth at 30 °C and reduction of selenite, respectively. Discrimination of the fox isolates and *C. upsaliensis* could also be observed using MALDI-TOF MS analysis. These results suggested that the 14 fox isolates, although related to *C. upsaliensis* and *C. helveticus,* were representatives of a novel species distinct from other validly-described species of the genus *Campylobacter*. The name *Campylobacter vulpis* sp. nov. was proposed, with isolate 251/13^{T} (=CCUG 70587^{T}; =LMG 30110^{T}) as the type strain. 16S rDNA, *atp*A, core gene phylogenetic analyses and whole-genome sequence comparisons, including ANI, AAI and isDDH analyses, were also performed to further address the taxonomic placement of *C. vulpis*. *atpA* and core gene dendrograms placed *C. vulpis* strains into clades related to but well separated from *C. upsaliensis.* Additionally, pairwise *C. vulpis/C. upsaliensis* ANI and AAI values were below those observed for existing *Campylobacter* subspecies pairs but consistent with other *Campylobacter* species pairs. Taken together, these data support the designation of *C. vulpis* sp. nov. as a novel *Campylobacter* species. A robust *C. vulpis*–specific PCR was developed which will permit the rapid identification of *C. vulpis* strains recovered from animal, environmental, or possibly human clinical samples. Description of *Campylobacter vulpis* sp. nov. is presented in Table 3.

ACKNOWLEDGEMENTS

We are in debt to A. Buchicchio and A. Bianco for their technical assistance. We also thank Emma Yee for technical assistance in sequencing the *C. vulpis* genomes. This publication made use of the *Campylobacter* Multi Locus Sequence Typing website (https://pubmlst.org/campylobacter/) sited at the University of Oxford [20]. The development of this site has been funded by the Wellcome Trust.

FUNDING INFORMATION

This work was funded by University of Bologna (Italy) Grant RFO14 and by the United States Department of Agriculture, Agricultural Research Service, CRIS project 2030-42000-230-051.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Acke, E. (2018) Campylobacteriosis in dogs and cats: a review. N. Z. Vet. J. 66, 221-228.

[2] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

[3] Besemer, J., Borodovsky, M. (2005) GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res. 33, W451-454.

[4] Bojanic, K., Midwinter, A.C., Marshall, J.C., Rogers, L.E., Biggs, P.J., Acke, E. (2016) Isolation of *Campylobacter* spp. from client-owned dogs and cats, and retail raw meat pet food in the Manawatu, New Zealand. Zoonoses Public Health.

[5] Bolton, F.J., Wareing, D.R.A., Skirrow, M.B., Hutchinson, D.N., Identification and biotyping of *Campylobacters*, in: R.G. Board, D. Jones, F.A. Skinner (Eds.)

Identification methods in applied and environmental microbiology, Blackwell Scientific Publications, Oxford, England, 1992, pp. 151-162.

[6] Boukerb, A.M., Penny, C., Serghine, J., Walczak, C., Cauchie, H.M., Miller, W.G., Losch, S., Ragimbeau, C., Mossong, J., Megraud, F., Lehours, P., Benejat, L., Gourmelon, M. (2019) *Campylobacter armoricus* sp. nov., a novel member of the *Campylobacter lari* group isolated from surface water and stools from humans with enteric infection. Int. J. Syst. Evol. Microbiol. 69, 3969-3979.

[7] Brosius, J., Palmer, M.L., Kennedy, P.J., Noller, H.F. (1978) Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. 75, 4801-4805.

[8] Bryant, E., Shen, Z., Mannion, A., Patterson, M., Buczek, J. (2020) *Campylobacter taeniopygiae* sp. nov., *Campylobacter aviculae* sp. nov., and *Campylobacter estrildidarum* sp. nov., novel species isolated from laboratory maintained Zebra Finches. Avian Dis. 64.

[9] Caceres, A., Munoz, I., Iraola, G., Diaz-Viraque, F., Collado, L. (2017) *Campylobacter ornithocola* sp. nov., a novel member of the *Campylobacter lari* group isolated from wild bird faecal samples. Int. J. Syst. Evol. Microbiol. 67, 1643-1649.

[10] Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. (2009) BLAST+: architecture and applications. BMC Bioinformatics. 10, 421.

[11] Chaban, B., Ngeleka, M., Hill, J.E. (2010) Detection and quantification of 14 *Campylobacter* species in pet dogs reveals an increase in species richness in feces of diarrheic animals. BMC Microbiol. 10, 73.

[12] Debruyne, L., Gevers, D., Vandamme, P., Taxonomy of the family *Campylobacteraceae*, in: I. Nachamkin, C.M. Szymanski, M.J. Blaser (Eds.) *Campylobacter*, ASM Press, Washington, D.C., 2008, pp. 3-25.

[13] El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., Finn, R.D. (2019) The Pfam protein families database in 2019. Nucleic Acids Res. 47, D427-D432.

[14] Engvall, E.O., Brandstrom, B., Andersson, L., Baverud, V., Trowald-Wigh, G., Englund, L. (2003) Isolation and identification of thermophilic *Campylobacter* species in faecal samples from Swedish dogs. Scand. J. Infect. Dis. 35, 713-718.

[15] Fitzgerald, C., Tu, Z.C., Patrick, M., Stiles, T., Lawson, A.J., Santovenia, M., Gilbert, M.J., van Bergen, M., Joyce, K., Pruckler, J., Stroika, S., Duim, B., Miller, W.G., Loparev, V.L., Sinnige, J.C., Fields, P.I., Tauxe, R.V., Blaser, M.J., Wagenaar, J.A. (2014) *Campylobacter fetus* subsp. *testudinum* subsp. nov., isolated from humans and reptiles. Int. J. Syst. Evol. Microbiol. 64, 2944-2948.

[16] Galata, V., Fehlmann, T., Backes, C., Keller, A. (2019) PLSDB: a resource of complete bacterial plasmids. Nucleic Acids Res. 47, D195-D202.

[17] Gilbert, M.J., Miller, W.G., Leger, J.S., Chapman, M.H., Timmerman, A.J., Duim, B., Foster, G., Wagenaar, J.A. (2017) *Campylobacter pinnipediorum* sp. nov., isolated from pinnipeds, comprising *Campylobacter pinnipediorum* subsp. *pinnipediorum* subsp. nov. and *Campylobacter pinnipediorum* subsp. *caledonicus* subsp. nov. Int. J. Syst. Evol. Microbiol. 67, 1961-1968.

[18] Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., Tiedje, J.M. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57, 81-91.

[19] Iannino, F., Salucci, S., Di Donato, G., Badagliacca, P., Vincifori, G., Di Giannatale, E. (2019) *Campylobacter* and antimicrobial resistance in dogs and humans: "One health" in practice. Vet. Ital. 55, 203-220.

[20] Jolley, K.A., Bray, J.E., Maiden, M.C.J. (2018) Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res. 3, 124.

[21] Kaakoush, N.O., Castano-Rodriguez, N., Mitchell, H.M., Man, S.M. (2015) Global epidemiology of *Campylobacter* infection. Clin. Microbiol. Rev. 28, 687-720.

[22] Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111-120.

[23] Konstantinidis, K.T., Tiedje, J.M. (2005) Genomic insights that advance the species definition for prokaryotes. Proc. Natl. Acad. Sci. U. S. A. 102, 2567-2572.

[24] Lagesen, K., Hallin, P., Rodland, E.A., Staerfeldt, H.H., Rognes, T., Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35, 3100-3108.

[25] Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. Bioinformatics. 23, 2947-2948.

[26] Laslett, D., Canback, B. (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. 32, 11-16.

[27] Lawson, A.J., Linton, D., Stanley, J., Owen, R.J. (1997) Polymerase chain reaction detection and speciation of *Campylobacter upsaliensis* and *C. helveticus* in human faeces and comparison with culture techniques. J. Appl. Microbiol. 83, 375-380.

568 [28] Linton, D., Owen, R.J., Stanley, J. (1996) Rapid identification by PCR of the genus
569 *Campylobacter* and of five *Campylobacter* species enteropathogenic for man and
570 animals. Res. Microbiol. 147, 707-718.
571 [29] Man, S.M. (2011) The clinical importance of emerging *Campylobacter* species. Nat.
572 Rev. Gastroenterol. Hepatol.
573 [30] Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P., Goker, M. (2013) Genome sequence-
574 based species delimitation with confidence intervals and improved distance functions.
575 BMC Bioinformatics. 14, 60.
576 [31] Miller, W.G., On, S.L., Wang, G., Fontanoz, S., Lastovica, A.J., Mandrell, R.E. (2005)
577 Extended multilocus sequence typing system for *Campylobacter coli*, *C. lari*, *C.*
578 *upsaliensis*, and *C. helveticus*. J. Clin. Microbiol. 43, 2315-2329.
579 [32] Miller, W.G., Yee, E., Bono, J.L. (2017) Complete genome sequence of the
580 *Campylobacter helveticus* type strain ATCC 51209. Genome Announc. 5: e00398-17.
581 [33] Miller, W.G., Yee, E., Jolley, K.A., Chapman, M.H. (2014) Use of an improved *atpA*
582 amplification and sequencing method to identify members of the
583 *Campylobacteraceae* and *Helicobacteraceae*. Lett. Appl. Microbiol. 58, 582-590.
584 [34] Miller, W.G., Yee, E., Lopes, B.S., Chapman, M.H., Huynh, S., Bono, J.L., Parker, C.T.,
585 Strachan, N.J., Forbes, K.J. (2017) Comparative genomic analysis identifies a
586 *Campylobacter* clade deficient in selenium metabolism. Genome Biol. Evol. 9, 1843-
587 1858.
588 [35] On, S.L., Holmes, B. (1991) Effect of inoculum size on the phenotypic characterization
589 of *Campylobacter* species. J. Clin. Microbiol. 29, 923-926.
590 [36] On, S.L., Holmes, B. (1991) Reproducibility of tolerance tests that are useful in the
591 identification of campylobacteria. J. Clin. Microbiol. 29, 1785-1788.

592 [37] On, S.L., Holmes, B. (1992) Assessment of enzyme detection tests useful in
593 identification of campylobacteria. J. Clin. Microbiol. 30, 746-749.
594 [38] On, S.L., Holmes, B., Sackin, M.J. (1996) A probability matrix for the identification of
595 campylobacters, helicobacters and allied taxa. J. Appl. Bacteriol. 81, 425-432.
596 [39] On, S.L.W., Miller, W.G., Houf, K., Fox, J.G., Vandamme, P. (2017) Minimal standards
597 for describing new species belonging to the families *Campylobacteraceae* and
598 *Helicobacteraceae*: *Campylobacter, Arcobacter*, *Helicobacter* and *Wolinella* spp. Int.
599 J. Syst. Evol. Microbiol. 67, 5296-5311.
600 [40] Palmer, M., Steenkamp, E.T., Blom, J., Hedlund, B.P., Venter, S.N. (2020) All ANIs are
601 not created equal: implications for prokaryotic species boundaries and integration of
602 ANIs into polyphasic taxonomy. Int. J. Syst. Evol. Microbiol. 70, 2937-2948.
603 [41] Piccirillo, A., Niero, G., Calleros, L., Perez, R., Naya, H., Iraola, G. (2016)
604 *Campylobacter geochelonis* sp. nov. isolated from the western Hermann's tortoise
605 (*Testudo hermanni hermanni*). Int. J. Syst. Evol. Microbiol. 66, 3468-3476.
606 [42] Richter, M., Rossello-Mora, R. (2009) Shifting the genomic gold standard for the
607 prokaryotic species definition. Proc. Natl. Acad. Sci. U. S. A. 106, 19126-19131.
608 [43] Rossi, M., Hanninen, M.L., Revez, J., Hannula, M., Zanoni, R.G. (2008) Occurrence and
609 species level diagnostics of *Campylobacter* spp., enteric *Helicobacter* spp. and
610 *Anaerobiospirillum* spp. in healthy and diarrheic dogs and cats. Vet. Microbiol. 129,
611 304-314.
612 [44] Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell,
613 B. (2000) Artemis: sequence visualization and annotation. Bioinformatics. 16, 944-
614 945.
615 [45] Saitou, N., Nei, M. (1987) The neighbor-joining method: a new method for
616 reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.

[46] Silva, M.F., Pereira, G., Carneiro, C., Hemphill, A., Mateus, L., Lopes-da-Costa, L., Silva, E. (2020) *Campylobacter portucalensis* sp. nov., a new species of *Campylobacter* isolated from the preputial mucosa of bulls. PLoS One. 15: e0227500.

[47] Stanley, J., Burnens, A.P., Linton, D., On, S.L., Costas, M., Owen, R.J. (1992) *Campylobacter helveticus* sp. nov., a new thermophilic species from domestic animals: characterization, and cloning of a species-specific DNA probe. J. Gen. Microbiol. 138, 2293-2303.

[48] Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.

[49] Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M., Ostell, J. (2016) NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44, 6614-6624.

[50] Van, T.T.H., Elshagmani, E., Gor, M.C., Scott, P.C., Moore, R.J. (2016) *Campylobacter hepaticus* sp. nov., isolated from chickens with spotty liver disease. Int. J. Syst. Evol. Microbiol. 66, 4518-4524.

[51] Varghese, N.J., Mukherjee, S., Ivanova, N., Konstantinidis, K.T., Mavrommatis, K., Kyrpides, N.C., Pati, A. (2015) Microbial species delineation using whole genome sequences. Nucleic Acids Res. 43, 6761-6771.

[52] Zanoni, R.G., Rossi, M., Giacomucci, D., Sanguinetti, V., Manfreda, G. (2007) Occurrence and antibiotic susceptibility of *Helicobacter pullorum* from broiler chickens and commercial laying hens in Italy. Int. J. Food Microbiol. 116, 168-173.

FIGURE AND SUPPLEMENTARY FIGURE LEGENDS

Figure 1. Electron micrograph of *C. vulpis* strain 251/13^T cells from a 48 h culture. (a) Cells with bipolar flagella. (b) Detail of unsheathed flagellum.

Figure 2. MLST-based minimum spanning tree illustrating the genotypic relationships of the *C. vulpis* sp. nov. strains from this study and strains representing all *C. upsaliensis* sequence types present within the PubMLST database (https://pubmlst.org/campylobacter/; accessed Sept 2020). The *C. helveticus* MLST locus set is different from that of *C. upsaliensis*; thus, to ensure consistency, allele sequences from the locus set used in the *C. upsaliensis* MLST method were extracted from the *C. helveticus* type strain genome sequence. Nodes derived from the type strains are labeled with a 'T'. *C. vulpis* nodes are labeled with their respective strains.

Figure 3. 16S phylogenetic tree representing *Campylobacter vulpis* sp. nov. and the *Campylobacter* type strains. The evolutionary history was inferred using the Neighbor-Joining method [45]. The percentage (>75%) of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method [22] and are in the units of the number of base substitutions per site.

Figure 4. *atpA* phylogenetic tree representing *Campylobacter vulpis* sp. nov. and the *Campylobacter* type strains. The evolutionary history was inferred using the Neighbor-Joining method [45]. The percentage (>75%) of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The

667 tree is drawn to scale, with branch lengths in the same units as those of the evolutionary
668 distances used to infer the phylogenetic tree. The evolutionary distances were computed
669 using the Kimura 2-parameter method [22] and are in the units of the number of base
670 substitutions per site.
671
672 **Figure 5.** Core and accessory genome of *C. vulpis* sp. nov. Core genome, blue. Accessory
673 genome, orange. The number of isolates considered is indicated on the x axis, while the
674 number of genes is represented on the y axis. Error bars are used to indicate the standard
675 deviation.
676

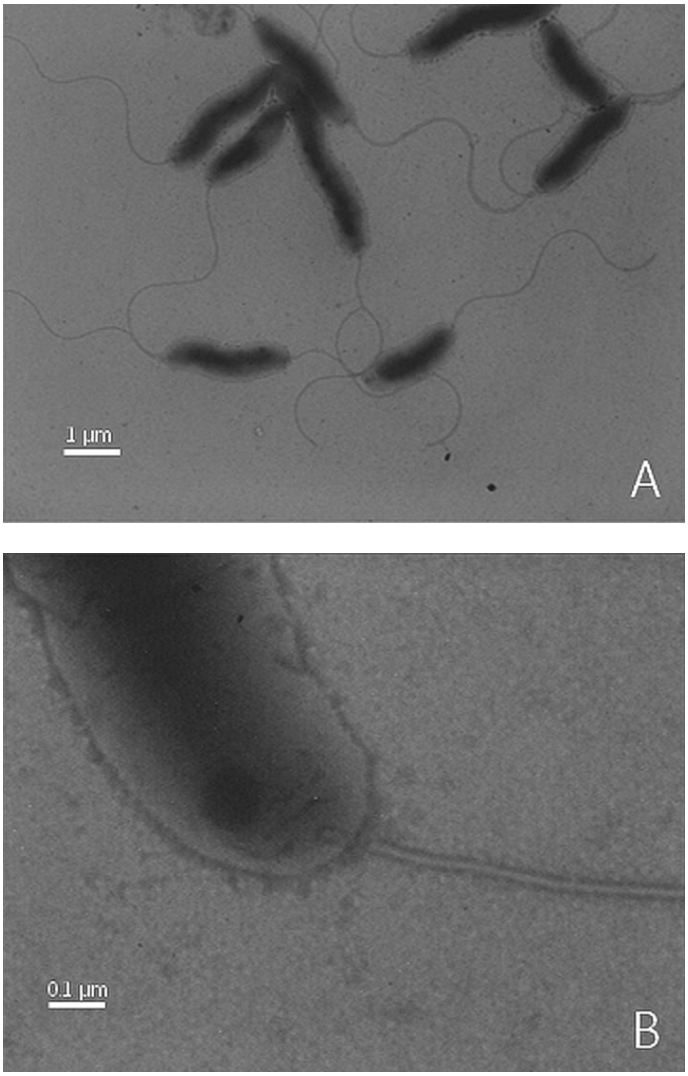

Figure 1. Electron micrograph of *C. vulpis* strain 251/13^{T} cells from a 48 h culture. (a) Cells with bipolar flagella. (b) Detail of unsheathed flagellum.

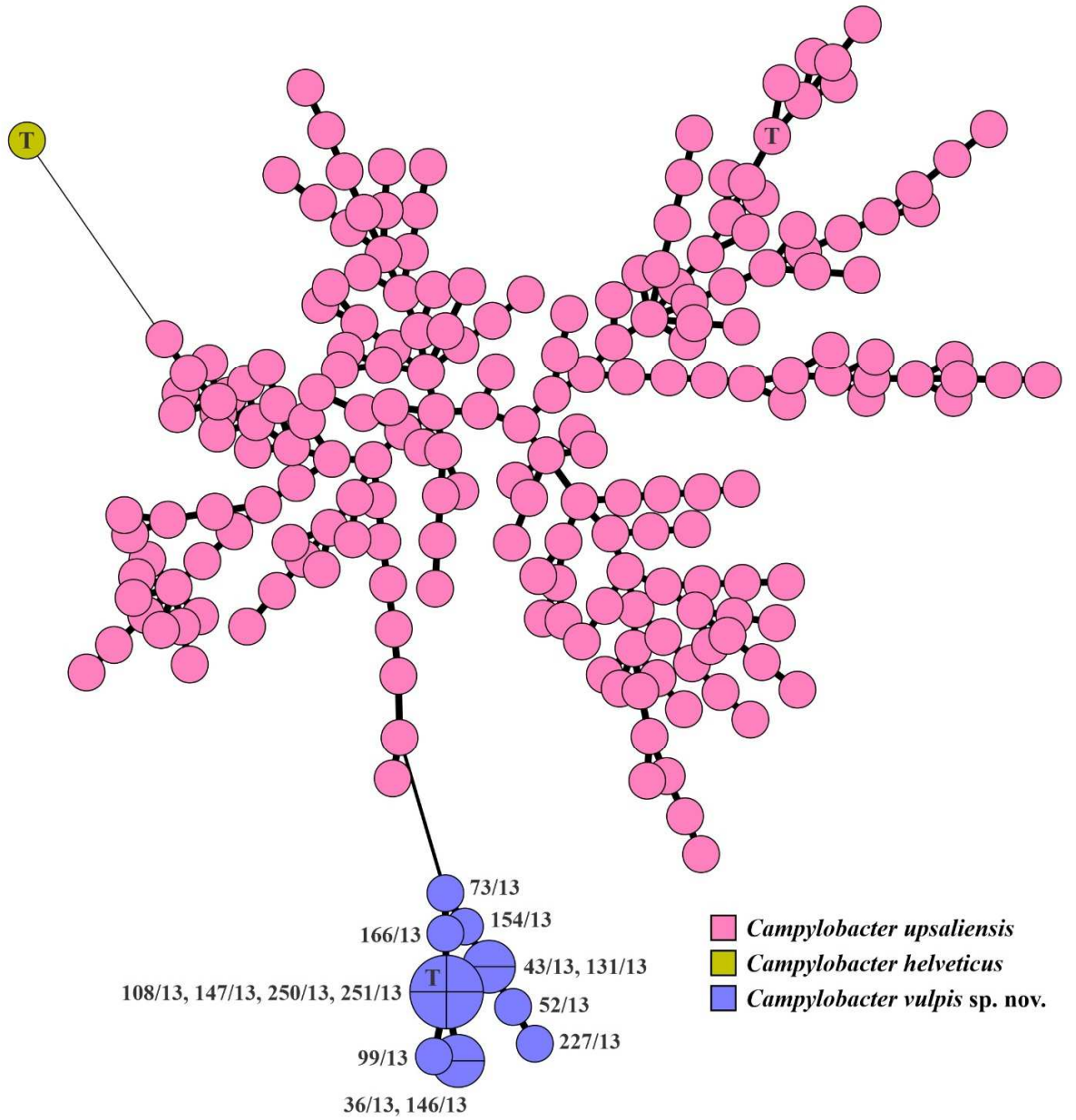

Figure 2. MLST-based minimum spanning tree illustrating the genotypic relationships of the *C. vulpis* sp. nov. strains from this study and strains representing all *C. upsaliensis* sequence types present within the PubMLST database (https://pubmlst.org/campylobacter/; accessed 9/14/2020). The *C. helveticus* MLST locus set is different from that of *C. upsaliensis*; thus, to ensure consistency, allele sequences from the locus set used in the *C. upsaliensis* MLST method were extracted from the *C. helveticus* type strain genome sequence. Nodes derived from the type strains are labeled with a 'T'. *C. vulpis* nodes are labeled with their respective strains.

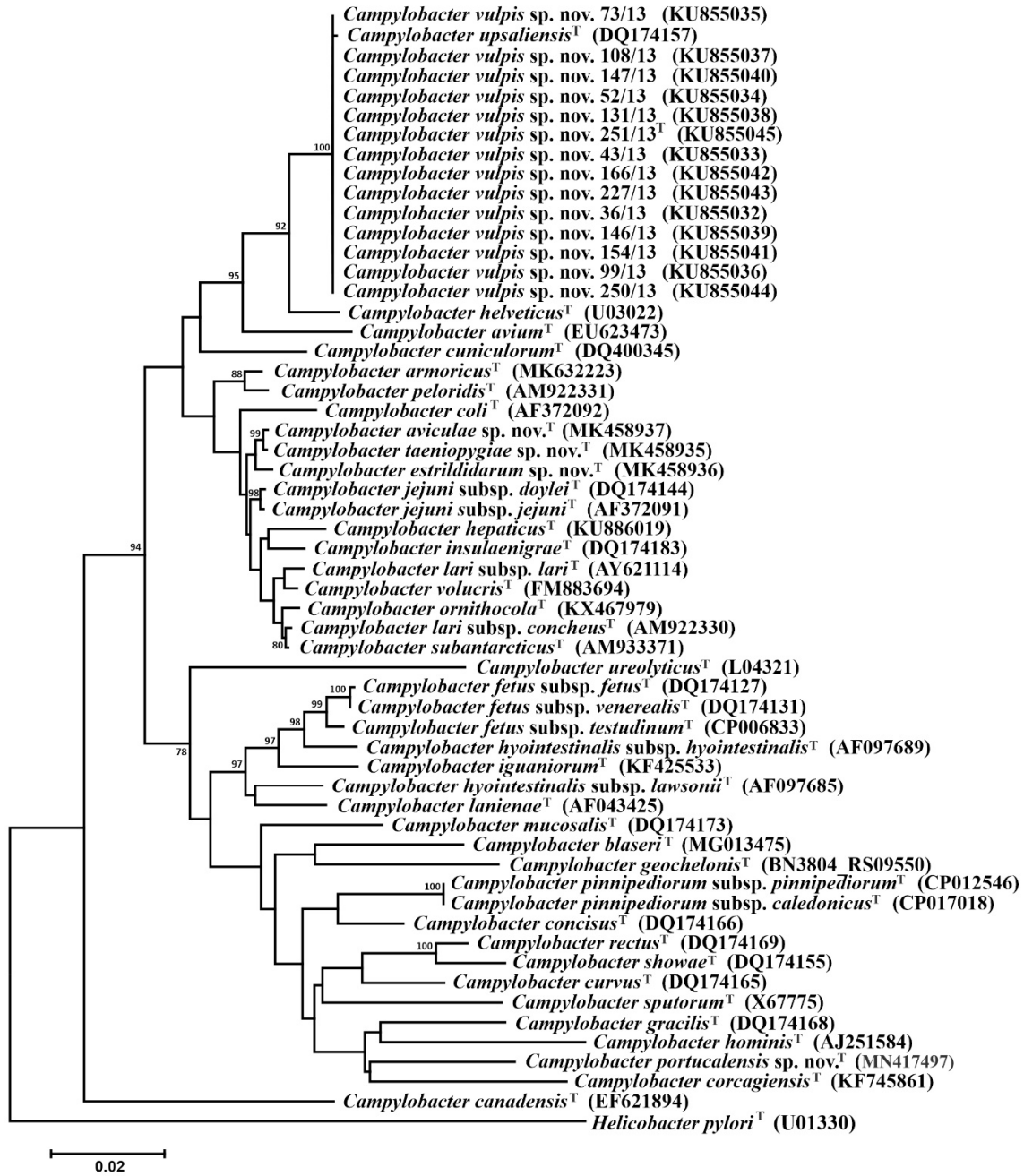

Figure 3: 16S phylogenetic tree representing *Campylobacter vulpis* sp. nov. and the *Campylobacter* type strains. The evolutionary history was inferred using the Neighbor-Joining method [45]. The percentage (>75%) of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method [22] and are in the units of the number of base substitutions per site.

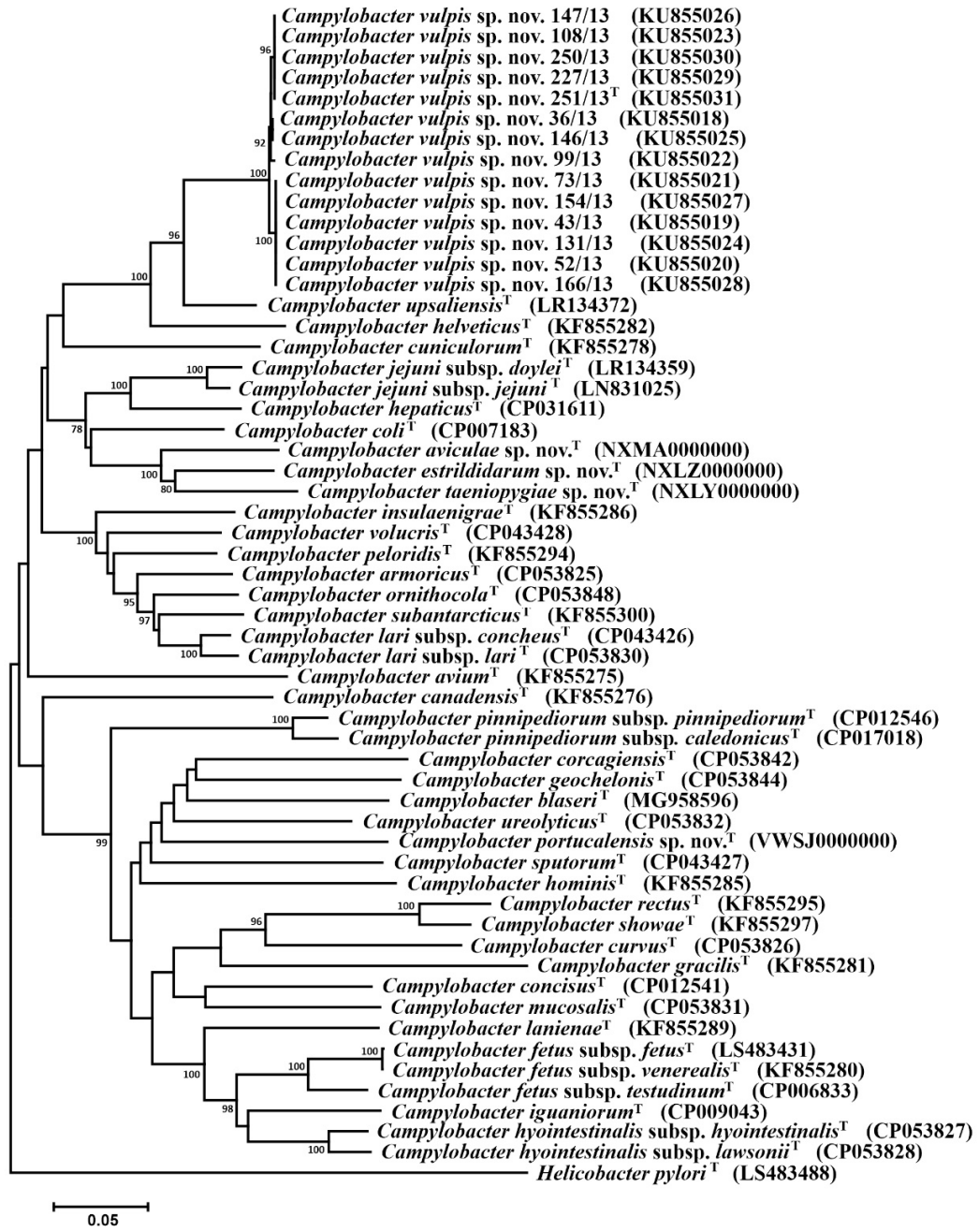

Figure 4: *atpA* phylogenetic tree representing *Campylobacter vulpis* sp. nov. and the *Campylobacter* type strains. The evolutionary history was inferred using the Neighbor-Joining method [45]. The percentage (>75%) of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method [22] and are in the units of the number of base substitutions per site.

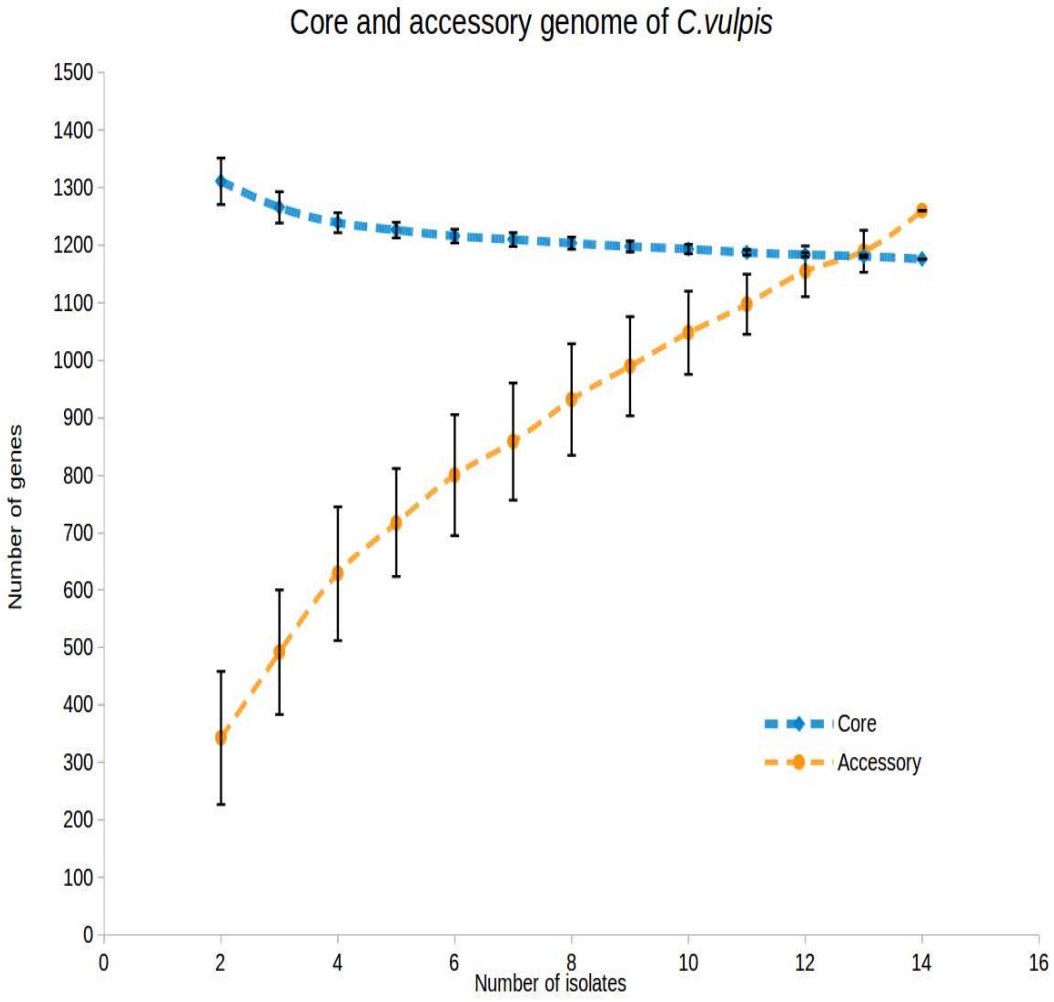

Figure 5. Core and accessory genome of *C. vulpis* sp. nov. Core genome, blue. Accessory genome, orange. The number of isolates considered is indicated on the x axis, while the number of genes is represented on the y axis. Error bars are used to indicate the standard deviation.

Table 1. Phenotypic characteristics of *Campylobacter* species

Characteristic	*vulpis* sp. nov. (n = 14)	*upsaliensis*	*helveticus*	*armoricus*	*aviculae*	*avium*	*blaseri*	*canadensis*	*coli*	*concisus*	*corcagiensis*	*cuniculorum*	*curvus*	*estrildidarum*	*fetus fetus*	*fetus testudinum*	*fetus venerealis*	*geochelonis*	*gracilis*	*hepaticus*	*hominis*	*hyointestinalis hyointestinalis*
Oxidase	+	+	+	+	+	+	+	+	+	v	+	+	+	+	+	+	+	+	-	+	+	+
Catalase	-	-	-	+	-	w	+	v	+	-	+	+	-	(-)	+	+	+	+	v	+	-	+
Urease	-	-	-	+	-	-	+	v	-	-	+	-	-	-	-	-	-	-	-	-	-	-
Alkaline phosphatase	v	-	-	-	-	-	+	-	-	(+)	+	-	v	-	-	ND	-	-	-	ND	-	-
Hydrolysis of:																						
Hippurate	-	-	-	-	+	+	-	-	-	-	-	-	(-)	(+)	-	-	-	-	-	(+)	-	-
Indoxyl acetate	+	+	+	-	-	+	+	-	+	-	v	+	v	-	-	-	-	-	(+)	+	-	-
Reduction of:																						
Nitrate	+	+	+	-	v	+	+	v	+	(-)	(+)	+	+	v	+	+	(+)	+	(+)	v	v	+
Selenite	+	+	-	v	ND	-	ND	ND	+	(-)	ND	-	-	ND	(+)	ND	(-)	-	-	ND	-	+
TTC	-	v	-	v	ND	-	ND	ND	+	-	-	v	v	ND	-	+	-	-	-	ND	-	(-)
H_2S production (TSI)	-	-	-	ND	ND	-	+	v	-	-	+	-	(-)	ND	-	-	-	-	-	-	-	+
α-haemolysis	+	+	+	-	ND	-	-	-	(-)	(-)	-	+	(-)	ND	-	ND	v	-	-	-	-	v
Growth at/in/on:																						
25 °C (microaerobic)	-	-	-	+	-	-	+	-	-	-	+	-	-	-	+	+	+	+	-	-	-	-
30 °C (microaerobic)	-	+	v	+	ND	-	+	-	-	(+)	+	-	+	ND	+	+	+	+	(+)	-	-	+
37 °C (microaerobic)	+	+	+	+	+	+	+	+	+	+	+	+	v	+	+	+	+	+	-	+	+	+
42 °C (microaerobic)	+	+	+	+	+	+	+	+	+	(+)	+	(+)	v	+	(+)	v	-	-	v	+	(-)	+
37 °C (anaerobic)	-	-	-	+	ND	-	+	+	-	+	+	-	+	ND	(-)	+	v	+	+	-	+	-
37 °C (aerobic)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CCDA	+	+	+	ND	ND	-	-	+	+	(-)	ND	(+)	(+)	ND	+	ND	+	+	v	ND	ND	+
Glycine (1% w/v)	+	+	v	+	(-)	-	w	v	(+)	(-)	+	-	+	(+)	+	+	(-)	+	+	+	+	+
NaCl (2% w/v)	-	-	(-)	-	ND	-	ND	-	-	(-)	+	-	v	ND	-	ND	-	+	v	-	ND	-
Bile (1% w/v)	+	+	+	ND	ND	v	ND	ND	(+)	-	+	ND	-	ND	+	ND	+	ND	-	+	ND	+
Requirement for H_2	-	-	-	-	ND	v	-	-	-	+	-	-	+	ND	-	-	-	-	+	-	+	v
Resistance to:																						
Nalidixic acid (30 µg)	-	-	-	-	-	-	-	v	-	(+)	+	v	+	(-)	+	ND	v	+	v	v	v	+
Cephalothin (30 µg)	-	(-)	-	+	+	+	-	-	+	-	ND	(+)	-	+	-	ND	-	-	-	(+)	-	(-)

Characteristic	*hyointestinalis lawsonii*	*iguaniorum*	*insulaenigrae*	*jejuni doylei*	*jejuni jejuni*	*lanienae*	*lari concheus*	*lari lari*	*mucosalis*	*ornithocola*	*peloridis*	*pinnipediorum caledonicus*	*pinnipediorum pinnipediorum*	*portucalensis*	*rectus*	*showae*	*sputorum*	*subantarcticus*	*taeniopygiae*	*ureolyticus*	*volucris*
Oxidase	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	v	+	+	+	+	+
Catalase	+	+	+	(+)	+	+	+	+	-	+	+	-	+	-	(-)	v	v*	+	+	(-)	+
Urease	-	-	-	-	-	-	-	v†	-	+	ND	+	+	-	-	-	v*	-	-	+	-
Alkaline phosphatase	(-)	ND	ND	-	-	+	ND	-	(+)	-	ND	ND	ND	ND	-	-	-	ND	-	-	-
Hydrolysis of:																					
Hippurate	-	-	-	+	(+)	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-
Indoxyl acetate	-	-	-	+	(+)	-	ND	(-)	-	-	ND	-	-	-	+	v	-	-	-	(-)	-
Reduction of:																					
Nitrate	+	+	+	-	+	+	+	+	(-)	v	ND	+	+	-	+	+	(+)	+	v	+	+
Selenite	+	ND	ND	-	(+)	ND	ND	v	(-)	ND	ND	ND	ND	ND	-	-	v	-	ND	-	+
TTC	-	ND	ND	v	(+)	ND	ND	(+)	-	(-)	ND	ND	ND	ND	-	-	-	ND	ND	-	-
H_2S production (TSI)	+	+	-	-	-	-	ND	-	+	-	ND	+	+	-	-	v	+	-	ND	-	-
α-haemolysis	v	+	ND	+	+	+	ND	+	-	-	ND	+	+	ND	+	+	+	+	ND	v	ND
Growth at/in/on:																					
25 °C (microaerobic)	-	+	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
30 °C (microaerobic)	+	+	-	-	(+)	-	-	+	+	-	-	+	+	ND	(-)	+	(+)	-	ND	+	-
37 °C (microaerobic)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	v	+	+	+	+	+
42 °C (microaerobic)	+	-	-	-	+	+	+	+	+	+	+	-	-	+	(-)	v	+	+	+	v	+
37 °C (anaerobic)	+	+	-	-	-	+	ND	-	+	+	ND	+	+	w	+	+	+	+	ND	+	+
37 °C (aerobic)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CCDA	+	+	ND	+	+	ND	+	+	+	ND	+	-	-	ND	-	+	(+)	ND	ND	+	ND
Glycine (1% w/v)	v	+	+	(-)	(+)	-	+	+	v	+	+	-	v	v	+	v	+	(+)	-	+	-
NaCl (2% w/v)	-	ND	-	-	-	-	+	(+)	(+)	ND	(+)	ND	ND	-	v	+	+	+	ND	+	-
Bile (1% w/v)	ND	ND	ND	+	+	ND	ND	ND	(+)	ND	ND	ND	ND	ND	-	-	v	+	ND	v	w
Requirement for H_2	v	-	ND	-	-	-	ND	-	+	-	ND	-	-	-	+	+	-	ND	ND	+	ND
Resistance to:																					
Nalidixic acid (30 µg)	+	+	+	-	-	+	-	v	(+)	ND	(+)	-	-	ND	(+)	-	(+)	+	-	-	+
Cephalothin (30 µg)	-	-	+	-	(+)	+	+	+	(-)	ND	(-)	-	-	ND	-	-	-	-	+	-	+

†: UPTC strains are urease positive; *: bv. fecalis strains are catalase positive, bv. paraureolyticus strains are urease positive; +, 90-100% positive; (+), 75-89% positive; v, 26-74% positive; (-), 11-25% positive; -, 0-10% positive; w, weak growth; ND, not determined; CCDA, charcoal cefoperazone deoxycholate agar; TTC, 2,3,5-triphenyltetrazolium chloride; TSI, triple sugar-iron agar. Data for reference taxa were taken from Piccirillo et al. [41], Van et al. [50], Cáceres et al. [9], Gilbert et al. [17], Boukerb et al. [6], Silva et al. [46], and Bryant et al. [8].

Table 2. Average amino acid and nucleotide identities among the *C. vulpis* sp. nov. strains and related *Campylobacter* taxa.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
***C. vulpis* sp. nov. 251/13^T**	**1**	---	99	99	99	99	99	100	99	99	100	99	100	99	100	90	85	74	75	74
***C. vulpis* sp. nov. 36/13**	**2**	99	---	99	99	99	99	99	99	100	99	99	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 43/13**	**3**	100	99	---	100	100	99	99	100	99	99	100	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 52/13**	**4**	100	99	100	---	99	99	99	100	99	99	100	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 73/13**	**5**	100	99	100	100	---	99	99	99	99	99	100	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 99/13**	**6**	100	99	100	100	100	---	99	99	99	99	99	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 108/13**	**7**	100	99	100	100	100	100	---	99	99	100	99	100	99	100	90	85	74	74	74
***C. vulpis* sp. nov. 131/13**	**8**	100	99	100	100	100	100	100	---	99	99	99	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 146/13**	**9**	99	100	99	99	99	99	100	99	---	99	99	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 147/13**	**10**	100	99	100	100	100	100	100	100	99	---	99	100	99	100	90	85	74	74	74
***C. vulpis* sp. nov. 154/13**	**11**	100	99	100	100	100	100	100	100	99	100	---	99	99	99	90	85	74	74	74
***C. vulpis* sp. nov. 166/13**	**12**	100	99	100	100	100	100	100	100	99	100	100	---	99	100	90	85	74	74	74
***C. vulpis* sp. nov. 227/13**	**13**	100	99	100	100	100	100	100	100	99	100	100	100	---	99	90	85	74	74	74
***C. vulpis* sp. nov. 250/13**	**14**	100	99	100	100	100	100	100	100	99	100	100	100	100	---	90	85	74	74	74
***C. upsaliensis* NCTC 11541^T**	**15**	95	95	95	95	95	95	95	95	95	95	95	95	95	95	---	85	74	74	74
***C. helveticus* ATCC 51209^T**	**16**	91	91	91	91	91	91	91	91	91	91	91	91	91	91	91	---	75	74	74
***C. jejuni jejuni* NCTC 11351^T**	**17**	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	82	---	96	84
***C. jejuni doylei* NCTC 11951^T**	**18**	81	80	81	81	81	81	81	81	80	81	81	81	80	81	81	81	98	---	84
***C. coli* ATCC 33559^T**	**19**	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	82	90	90	---

Average amino acid identity values are on the bottom left, and average nucleotide identity values are on the upper right. The average nucleotide identity (ANIb) values were calculated using JSpecies (v. 1.2.1)

Table 3. Description of *Campylobacter vulpis* sp. nov.

Genus name	*Campylobacter*
Species name	*Campylobacter vulpis*
Specific epithet	*vulpis*
Species status	sp. nov.
Species etymology	(*vulpis.* L. gen. n. *vulpis* of a fox)
Description of the new taxon and diagnostic traits	Cells are Gram-negative, sigmoid to allantoid in shape, 0.3-0.4 µm in width and 1.2-3.0 µm in length with a single flagellum at both poles and are motile with characteristic darting movements when observed by dark field microscopy. Cells appear coccoid after 5-6 days of incubation or when exposed to air. After incubation on Nutrient agar [amended with 5% (v/v) sheep blood] in a microaerobic atmosphere at 37 ± 1°C for 48 h, colonies appear 2-3 mm in diameter, α-haemolytic, grey, translucent, flat with an irregular edge, and show a tendency to spread along the direction of the streak and to swarm and coalesce. Strictly microaerophilic. Able to grow at 37 and 42 °C, but not at 25 and 30 °C or under anaerobic and aerobic conditions. Hydrogen is not required for growth. All isolates are oxidase positive, and catalase and urease negative. Hippurate is not hydrolysed, while all strains hydrolyse indoxyl acetate. Half of the isolates are alkaline phosphatase positive. Hydrogen sulfide is not produced on TSI agar. All strains are unable to reduce triphenyl tetrazolium chloride (TTC), while most reduce nitrate and selenite. All strains grow on Nutrient agar without blood and on CCDA, but not on MacConkey agar. Growth occurs in the presence of 1% (w/v) bile, but not in the presence of 2% (w/v) NaCl. Most strains grow in presence of 1% (w/v) glycine. Strains are susceptible to nalidixic acid (30 µg) and to cephalothin (30 µg) by disc diffusion tests. Pathogenicity is unknown. The type strain 251/13^{T} (=CCUG 70587^{T}, =LMG 30110^{T}) was isolated from the caecal contents of a wild red fox (*Vulpes vulpes*) in northern Italy in 2013. The DNA G+C content of the type strain is 34.62 mol %.
Country of origin	Italy
Region of origin	Emilia Romagna
Source of isolation	*Vulpes vulpes* – cecal content
Sampling date (dd/mm/yyyy)	17/06/2013
Latitude (xx°xx′xx″N/S)	44°33’10.5” N
Longitude (xx°xx′xx″E/W)	11°26’41.762” E
Altitude (meters above sea level)	28 m
16S rRNA gene accession nr.	KU855045
Genome accession number [RefSeq; EMBL; …]	CP041617
Genome status	Complete
Genome size	1,645 kbp
GC mol%	34.62 GC mol%
Number of strains in study	14
Source of isolation of non-type strains	*Vulpes vulpes* – cecal content
Designation of the Type Strain	251/13^{T}
Strain Collection Numbers	CCUG 70587^{T}; LMG 30110^{T}