Molecular characterization and differentiation of *Campylobacter fetus* subspecies

Linda van Bloois

2016

ISBN: 978-90-393-6659-2
Cover design: Kees van Vliet
Printing: Off Page

Financial support for publication of this thesis was kindly provided by:
Applied Maths NV, Becton Dickinson BV, BioTRADING Benelux BV, Infection & Immunity Utrecht, MART Microbiology BV, Thermo Fisher Scientific BV, World Courier Holland BV

Molecular characterization and differentiation of *Campylobacter fetus* subspecies

Moleculaire karakterisatie en differentiatie van *Campylobacter fetus* subspecies

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 16 november 2016 des middags te 2.30 uur

door

Linda van Bloois

geboren op 25 oktober 1981 te Rotterdam

Promotor: Prof.dr. J.A. Wagenaar

Copromotoren: Dr. B. Duim

Dr. A.L. Zomer

Contents

Chapter

General Introduction

General Introduction

History of nomenclature

The first publication implicating *Campylobacter fetus* (*C. fetus*), formerly named *Vibrio fetus* (*V. fetus*), as a cause of abortion in cattle and sheep was made in 1919 by Smith and Taylor (1). In 1959, Florent observed that venereal transmitted enzootic infertility in cattle was caused by a variant of *Vibrio fetus*, which he named *V. fetus* subsp. *venerealis* (2). A closely related intestinal variant associated with sporadic abortions in cattle was named *V. fetus* subsp. *intestinalis* (2). Florent described the biochemical assays that could differentiate the two variants; *V. fetus venerealis* did not produce hydrogen sulfide (H_2S) in cysteine-rich medium and was not able not grow in the presence of 1% glycine, whereas *V. fetus intestinalis* was able to produce H_2S and to grow in the presence of 1% glycine (2). Furthermore, a biovar of *V. fetus venerealis*, which was not 1% glycine tolerant, but able to produce H_2S was described by Florent in 1963 (3). This biovar was able to colonize not only the intestines, but also the genital tract of cattle and was named *V. fetus* subsp. *venerealis* biovar intermedius (3).

In 1973, Véron and Chatelain re-classified *V. fetus* into the newly described genus *Campylobacter* (4). *V. fetus* subsp. *intestinalis* was named *C. fetus* subsp. *fetus* (Cff), *V. fetus* subsp. *venerealis* was named *C. fetus* subsp. *venerealis* (Cfv) and *V. fetus* subsp. *venerealis* biovar intermedius was named *C. fetus* subsp. *venerealis* biovar intermedius (Cfvi) (4). This nomenclature of *C. fetus,* described by Veron *et al.* (4) was officially accepted in 1984 (5).

In 2014, an additional subspecies of *C. fetus* was described, designated *C. fetus* subsp. *testudinum* (Cft) (6). Cft strains were isolated from reptile species and humans and were shown to be different from Cff and Cfv strains based on MALDI-TOF MS results and genome comparisons (6).

General characteristics of Campylobacter fetus

Campylobacter fetus is one of the currently 27 recognized species within the genus *Campylobacter* (7). *C. fetus* is a Gram-negative, slender, rod-shaped bacterium (0.2 to 0.8 µm wide and 0.5 to 5.0 µm long) (8, 9). In old cultures (72-96h), cells usually change their shape and become coccoid (10). *C. fetus* organisms are motile and each cell has at least one polar flagellum. Examination using phase-contrast microscopy usually reveals a darting, corkscrew-like motion. *C. fetus* is catalase- and oxidase-positive and is micro-aerophilic with a respiratory type of metabolism (5). Optimal growth occurs at 37°C and

strains usually yield visible growth in 48-72h (5).

C. fetus cells express a proteinaceous surface-layer (S-layer). The capsule polysaccharides are found to be associated with the flagellum and outer cell envelope of strains (11, 12). *C. fetus* displays two major serotypes, based on variation of the O-antigen; Cfv strains belong to serotype A, Cff strains belong to either serotype A or serotype B, and Cft strains belong to serotype A, serotype B or serotype AB (13, 14). The genes encoding the surface layer proteins (sap) are located in a single genomic locus that contains *sap* homologs and one promoter that can switch the expression of *sap* homologs that are involved in antigenic variation (15). The antigenic variation is limited by the number of *sap* homologs, but can be sufficient to enable escape from adaptive immunity, thereby prolonging colonization of the *C. fetus* bacteria, resulting in infection and abortion or infertility (16). Furthermore, the S-layer of *C. fetus* is associated with resistance to phagocytosis (12, 17) and with *in vitro* serum resistance of strains (18); cross-reactive antiserum against the core region of the capsule polysaccharides demonstrated the presence of conserved antigens among serum-sensitive Cff strains, but these epitopes were not accessible in serum-resistant strains (19).

Clinical features of *C. fetus* su sp *fetus* and *C. fetus* su sp *venerealis*

Bovine Genital Campylobacteriosis

Campylobacter fetus subsp. *venerealis* is described as the causative agent of Bovine Genital Campylobacteriosis (BGC). BGC is a venereal disease, also known as Bovine Venereal Campylobacteriosis (BVC). BGC is sexually transmitted and characterized by infertility, early embryonic death and abortions in bovine (16).

BGC is currently classified by the World Organization for Animal Health (OIE) as a notifiable disease since it is deemed to have socio-economic and public health implications. All OIE Member states are obligated to monitor the presence of BGC in their countries, according to the OIE *Terrestrial Animal Health Code* (20) and report these findings to the OIE. These findings are publicly online available (21). Several countries have been successful in eradicating BGC, whereas in many countries BGC is still endemic. The incidence of BGC is highest in low and middle income countries (LMIC) where natural breeding of cattle is widely practiced, compared to high income countries where cattle are bred through artificial insemination (22).

Economic impact of BGC

BGC has a worldwide distribution (22) and causes high economic losses in endemic areas like La Pampa province in Argentina (23). BGC has been referred to as the "quiet profit taker", since infection in a herd can easily be overlooked (16). Infection is usually not suspected until low calving rates are noted within a herd, by which time extensive losses already have occurred. The primary economic losses incurred by the farmer are due to a decrease in calving percentages, delayed calving, culling of infertile animals and abortions (24). Research by Akhtar *et al.* (1993) indicated that infection with Cfv in dairy herds negatively influenced the production of milk, but this was not quantified (25). When BGC becomes established within a herd, profit margins may be as much as 36% below those of non-infected herds (26). The presence of the disease influences the international trade of animals and animal products (26).

The regulations of the trade in cattle aim to preserve the health status of the animal population and reduce the risk of spread of BGC. Freedom from infection with Cfv is a requirement of many countries for the import and export of bovine semen, embryos and cattle, as well as for health certification of bulls in semen production and distribution centers. The general animal health requirements governing the intra-EU trade and import of bovine semen are laid down in European Union Directive 88/407/EEC (27). This Directive harmonizes the animal health conditions for the trade within the EU and import to the EU from third countries, as well as the conditions of semen collection and storage. To ensure the health status of the bovine products and reduce the risk of spread of BGC, sensitive and specific diagnostic tests for the diagnosis of BGC are essential (28).

Host colonization of C. fetus subsp*. fetus and C. fetus* subsp. *venerealis strains*

Cfv is only isolated from the genital tract of cattle, whereas Cff can be isolated from different niches and animal species, including bovine, ovine, horses, goats, fowl, pigs, and human (29-33). Cfvi strains can be isolated from the genital tract as well from the intestines of cattle (34-36).

The mammal-associated *C. fetus* strains have different routes of transmission; Cff is described to be mainly orally transmitted although venereal transmission is possible and Cfv is only venereally transmitted (36). Venereal transmission of *C. fetus* between animals primarily occurs during coitus or artificial insemination procedures (16, 37). In bulls, it has been demonstrated that contact with contaminated bedding can also lead to infection and that transmission may also occur between bulls, during mounting, when large numbers of animals are enclosed together (37, 40).

Virulence of C. fetus subsp*. fetus and C. fetus* subsp. *venerealis infections*

C. fetus infection in bulls is asymptomatic and is not accompanied by either histological changes or modifications of the characteristics of the semen (37). Individual bulls vary in their susceptibility to infection; some animals become permanent carriers while others appear to be resistant to infection (38).

In cows, Cff infection is associated with abortion and Cfv infection with abortion and infertility (36). The diagnosis based on clinical signs in cows may be difficult in the acute stage. Breeding records often give the first indication, with many cows returning to estrus after repeated service: the repeat breeder syndrome (16). Cycles are longer than normal, usually more than a month, indicating early embryonic death (37). After variable periods of infertility, cows can recover and regain fertility (16). Cows can clear the infection, however some cows become permanent carriers (39).

Campylobacter fetus subsp. *fetus* is also recognized as an opportunistic pathogen in humans, particularly of compromised or immune-deficient patients (41). Infections are sporadic but a few food-borne outbreaks resulting from the consumption of contaminated raw beef, raw milk or cottage cheese have been reported (41).

Identification of *C. fetus* su sp *fetus* and *C. fetus* su sp *venerealis*

Strain isolation

For the isolation of *C. fetus* from bovine samples, methods are described in the OIE *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* (42). For the isolation of *C. fetus* from bovine samples, selective media with antimicrobial supplements are required. The recommended selective medium for isolation of *C. fetus* isolates is Skirrow's medium (43). This medium contains no cephalosporins, which can inhibit the growth of *C. fetus* (44). An alternative method is the filter technique with the use of non-selective media (45).

The success of culturing depends upon the collection and maintenance of the *C. fetus* bacterium, which has limited viability outside the host. Collection, use of transport media, isolation media and treatment of the samples effect the isolation of *C. fetus* strains and are of crucial importance, since isolation strongly depends on survival and propagation of the strains and the reduction of contaminating flora (46-48).

Phenotypic identification

The original described phenotypic tests to differentiate Cff and Cfv is tolerance to 1% glycine and H_2S production in medium containing 0.02% cysteine (2). Recommendations

for performing more phenotypic tests used for identifying *Campylobacter* spp. were made in 1996 (49). Four of these proposed extra phenotypic tests were evaluated with the 1% glycine tolerance test for the differentiation of Cff and Cfv isolates; susceptibility to the antibiotics metronidazole and cefoperazone, selenite reduction and growth at 42°C (50). It was shown that for all Cfv strains, the results of the selenite reduction and sensitivity to metronidazole and cefoperazone completely correlated with the results of the 1% glycine tolerance test. However for Cff strains, the results of these phenotypic tests only partly correlated with the results of the 1% glycine tolerance test, which led to the conclusion that the present traditional phenotypic characterization of *C. fetus* subspecies based on 1% glycine tolerance remained indispensable (50).

The current method recognized by the OIE for the differentiation of Cff and Cfv is tolerance to 1% glycine: Cff is able to grow in the presence of 1% glycine whereas the growth of Cfv is inhibited (42). The hydrogen sulphide (H_2S) test from medium containing 0.02% cysteine is able to differentiate Cfv (H_2S production negative) from Cff and Cfvi (both H_2S production positive) (4).

Molecular identification

Several molecular assays for *C. fetus* (sub)species identification and strain differentiation have been described and are presented in Table 1. Several assays are laborious and not suitable for standard routine diagnostic laboratories; hybridization assays (51-53), restriction fragment length polymorphism (RFLP) (54, 55), 16S sequencing (56) and pulsed field gel electrophoresis (PFGE) (57, 58). More practical assays for *C. fetus* identification have been described, such as nested PCR, targeting 16S and 23S rRNA (59) and loop-mediated isothermal amplification (LAMP) (60, 61). However, nested PCR is not preferred for routine diagnostics, since this type of assay is sensitive for contamination, whereas LAMP is relatively unknown and uncommon in diagnostic laboratories. The molecular typing methods Amplified Fragment Length Polymorphism (AFLP) (62) and Multi Locus Sequence Typing (MLST) (63) were shown to be able to differentiate the two *C. fetus* subspecies reliably and can be used for identification of *C. fetus* species and subspecies, however these tests are also laborious and therefore not preferred for routine diagnostics.

A preferred platform for routine *C. fetus* subspecies identification is PCR, since this type of assay is fast and easy. A number of PCR assays have been described to identify *C. fetus* species; PCR assays targeting the *cdt*-genes (64) and *cpn60*-gene (65) are described for *C. fetus* species identification, but these assays are not able to discriminate between the two

a le Molecular methods for *C. fetus* (sub)species identification differentiation

Reference	Method	Target	Identification	Comment
Wesley 1991	hybridization	16S rRNA	*C. fetus* species	Not suitable for routine diagnostics
Fujita 1994	hybridization	sapA	*C. fetus* species	Not suitable for routine diagnostics
Blom 1995	hybridization	16S rRNA	*C. fetus* species	Not suitable for routine diagnostics
Casademont 2000	hybridization	IG02	*C. fetus* species	Not suitable for routine diagnostics
Eaglesome 1995	RFLP	16S rRNA	*C. fetus* species	Not suitable for routine diagnostics
Cardarelli-Leite 1996	RFLP	16S rRNA	*C. fetus* species	Not suitable for routine diagnostics
Salama 1992	PFGE		*C. fetus* subspecies	Not suitable for routine diagnostics
Fujita 1995	PFGE		*C. fetus* species	Not suitable for routine diagnostics
Vargas 2003	RFLP-PFGE		*C. fetus* species	Not suitable for routine diagnostics
Gorkiewicz 2003	16S sequencing	16S rRNA	*C. fetus* species	Not suitable for routine diagnostics
Inglis 2003	Nested PCR	16S and 23S rRNA	*C. fetus* species	Sensitive for contamination
Yamazaki 2009	LAMP		*C. fetus* species	Uncommon for routine diagnostics
Yamazaki 2010	LAMP		Cfv	Uncommon for routine diagnostics
Wagenaar 2001	AFLP		*C. fetus* subspecies	Laborious and time consuming
Van Bergen 2005	MLST		*C. fetus* subspecies	Laborious and time consuming
Bastyns 1994	PCR	23S rRNA	*C. fetus* subspecies	*C. fetus* subspecies identification failed
Linton 1996	PCR	16S rRNA	*C. fetus* species	*C. fetus* species identification failed
Oyarzabal 1997	PCR	16S rRNA	*C. fetus* subspecies	Region is too conserved for subspecies differentiation
Hum 1997	Multiplex PCR	*cst* gene	*C. fetus* species	Sensitivity and specificity unknown
		parA gene	Cfv	Target parA not suitable for Cfv identification
Wang 2002	PCR	*sapB2* gene	Cff	Sensitivity and specificity unknown
Muller 2003	PCR	*parA* gene	Cfv	Target parA not suitable for Cfv identification
Van Bergen 2005	PCR	Unknown	Cfv	No Cfvi detection
Schulze 2006	PCR	*parA* gene	Cfv	Target parA not suitable for Cfv identification
McMillen 2006	Real-time PCR	*parA* gene	Cfv	Target parA not suitable for Cfv identification
Asakura 2007	PCR	*cdt* genes	*C. fetus* species	No subspecies differentiation
Abril 2007	Multiplex PCR	*nahE* gene	*C. fetus* species	Sensitivity and specificity unknown
		lscfe1	Cfv	Sensitivity and specificity unknown
Chaban 2009	PCR	*cpn60* gene	*C. fetus* species	No subspecies differentiation
Moolhuyzen 2009	PCR	Virulence genes	Cfv	Not present in all Cfv strains
Schmidt 2010	PCR	*parA* gene	Cfv	Target parA not suitable for Cfv identification
Iraola 2012	PCR	Virulence genes	Cfv	Not present in all Cfv strains
Chaban 2012	PCR	*parA* gene	Cfv	Target parA not suitable for Cfv identification

subspecies. A PCR assay targeting 23S rRNA identifies *C. fetus* species, but attempts to design primers on this target to discriminate between the two subspecies failed (66). A PCR assay targeting 16S rRNA (28) differentiated *C. fetus* from other *Campylobacter* species, but only a single base mismatch in the target 16S rRNA region differentiates the two *C. fetus* subspecies, making this region too conserved to enable adequate differentiation of the two subspecies by using a PCR assay (28). Two Cfv-specific PCRs have been described targeting virulence genes (67, 68), however, Gorkiewicz *et al.* (2010) showed that these virulence genes are not present in all Cfv strains (69) and therefore not suitable for Cfv identification. A PCR assay targeting a cryptic genomic DNA region with unknown function was able to identify Cfv strains correctly (70), but this assay did not detect Cfvi strains. The multiplex PCR described by Hum *et al*. (71) is currently the most cited and used PCR, targeting the gene *cstA* for *C. fetus* identification and the gene *parA* for Cfv identification. Several PCR assays using the same target genes have been described (26, 50, 68, 72-74), however, one of the PCR assays using *parA* as target (26) showed a positive result with a *C. hyointestinalis* strain isolated from a bull (75), which renders the *parA* gene unsuitable as target for Cfv identification. The most recently described *C. fetus* PCR assay is a multiplex PCR assay (76), targeting gene *nahE* for *C. fetus* identification and insertion sequence *ISCfe1* for Cfv identification. This multiplex PCR is promising, but has only been tested with a limited number of strains and has an unknown specificity and sensitivity.

Campylobacter fetus genomics

In 1992, the genome sizes of *C. fetus* strains were estimated by using PFGE, with Cff having a genomic size of 1.1 Mbp and Cfv of 1.3 Mbp (77). It was also described to use the observed genome size differences for subspecies differentiation (77). With PFGE, a physical map of the chromosome of Cff strain ATCC 27374 was constructed in 1995 (78). The estimated genome size of this strain was 1.17 Mbp and multiple genes were identified; a 2.8 kb *sapA* locus; three copies of rRNA genes; RNA polymerase genes *rpoA*, *rpoB* and *rpoD*; and flagellin genes *flaAB*, *gyrA* and *gyrB* (78).

The first complete Sanger sequenced and closed *C. fetus* genome was Cff strain 82-40 in 2006 by the Institute for Genomic Research (TIGR), with a genome size of 1.77 Mbp and with approximately 1820 predicted genes (Genbank accession number NC_008599). The first sequenced *C. fetus* subsp. *venerealis* genome of strain Azul-94 was published in 2009 by Moolhuyzen *et al.* (67), as unassembled genome consisting of multiple contigs.

However, with comparative analysis, they showed that an 80 Kb genomic sequence was unique to Cfv strain Azul-94 compared to Cff strain 82-40. This unique sequence contained type IV secretion system encoding genes, putative plasmid genes and hypothetical genes. It was suggested in this publication to use these Cfv-specific virulence genes as target for molecular identification of Cfv strains (67).

In 2010, Gorkiewicz *et al.* described a Cfv-specific genomic island (GI) containing a type IV secretion system (69). This GI was uniquely present in Cfv strains and it was suggested that the potential virulence functions of the island were involved in the pathogenicity of Cfv strains (69).

Aim and outline of the thesis

The aim of this thesis was to study the genomic characteristics and differences of the mammal-associated *C. fetus* subspecies *fetus* (Cff), *C. fetus* subspecies *venerealis* (Cfv) and *C. fetus* subsp. *venerealis* biovar intermedius (Cfvi) strains. Furthermore, we studied whether specific genomic characteristics could be associated with the virulence and the host specificity of *C. fetus* strains. A crucial element of this study was a reliable identification and differentiation of *C. fetus*.

In **Chapter 2**, we evaluate the currently available molecular assays that are described for the identification and differentiation of *C. fetus* strains.

In **Chapter 3**, the first closed genome sequence of a genotypically identified *C. fetus* subsp. *venerealis* biovar intermedius strain is presented.

In **Chapter 4**, the core genome and accessory genes of the *C. fetus* strains were defined and we characterized the *C. fetus* strains of both subspecies based on this whole-genome analysis. We compared the genomic characteristics of the strains with the current *C. fetus* subspecies identification based on phenotypic assays.

In **Chapter 5**, we performed a genome-wide single-nucleotide polymorphisms (SNPs) analysis of 42 *C. fetus* genomes. We performed a SNP-based phylogenetic analysis of the core genomes and a BEAST analysis to estimate the divergence dates of Cff and Cfv strains. Additionally, we investigated whether the genomes contain specific SNPs or genes that could be associated with the biochemical phenotypic identification and different clinical features of the *C. fetus* strains.

In **Chapter 6**, we examined the diversity of T4SS-encoding regions in 27 *C. fetus* strains using comparative genomics, and identified the location and composition of all T4SS encoding regions and their phylogeny. Furthermore, we studied whether the presence of specific T4SSs and *fic* genes could be associated with the *C. fetus* subspecies, their pathogenicity, the surface layer serotypes and geographic origin of the strains.

In **Chapter 7**, genome sequences were analyzed to find Cff and Cfv specific sequences for the development of a diagnostic PCR assay, to be used for Cff and Cfv differentiation. In this study, a comprehensive strain collection of 156 *C. fetus* strains and 15 non-fetus *Campylobacter* spp. was used to test the sensitivity and specificity of the newly developed PCR assays.

References

1. Smith, T, Taylor, MS. 1919. Some Morphological and Biological Characters of the Spirilla (*Vibrio Fetus*, N. Sp.) Associated with Disease of the Fetal Membranes in Cattle. J. Exp. Med. 30:299-311.

2. Florent, A. 1959. Les deux vibriosis génitales; la vibriose due à *V. fetus venerealis* et la vibriose d'origine intestinale due à *V. fetus intestinalis*. Mededelingen der Veeartsenijschool van de RijksUniversiteit te Gent. 1-60.

3. Florent, A. 1963. A propos dese vibrions responsables de la vibriose génitale des bovins et des ovins. Bull. Off. Int. Epizoot. 60:1063-1074.

4. Véron, M., Chatelain,R. 1973. Taxonomy study of the genus *Campylobacter* Sebald and Verón and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Verón. International Journal of Systematic Bacteriology. 23:122-134.

5. Smibert, RM. 1984. Genus II *Campylobacter* Sebald and Véron 1963, 907, p. 111-118. *In* Krieg, NR, Holt, HG (eds.), Bergey's manual of systematic bacteriologyvol. 1. Williams, & Wilkins Co., Baltimore.

6. Fitzgerald, C, Tu, ZC, Patrick, M, Stiles, T, Lawson, AJ, Santovenia, M, Gilbert, MJ, van Bergen, M, Joyce, K, Pruckler, J, Stroika, S, Duim, B, Miller, WG, Loparev, V, Sinnige, JC, Fields, PI, Tauxe, RV, Blaser, MJ, Wagenaar, JA. 2014. *Campylobacter fetus* subsp. *testudinum* subsp. nov., isolated from humans and reptiles. Int. J. Syst. Evol. Microbiol. 64:2944-2948.

7. List of Prokaroytic Names with standing in nomenclature. http://www.bacterio.net/index.html

8. Bryner, JH, Frank, AH, O'Berry, PA. 1962. Dissociation studies of Vibrios from the bovine genital tract. Am. J. Vet. Res. 23:32-41.

9. OGG, JE. 1962. Studies on the coccoid form of ovine *Vibrio fetus* I. Cultural and serologic investigations. Am. J. Vet. Res. 23:354-358.

10. NG, LK, Sherburne, R, Taylor, DE, Stiles, ME. 1985. Morphological forms and viability of *Campylobacter* species studied by electron microscopy. J. Bacteriol. 164:338-343.

11. McCoy, EC, Doyle, D, Wiltberger, H, Burda, K, Winter, AJ. 1975. Flagellar ultrastructure and flagella-associated antigens of *Campylobacter fetus*. J. Bacteriol. 122:307-315.

12. McCoy, EC, Doyle, D, Burda, K, Corbeil, LB, Winter, AJ. 1975. Superficial antigens of *Campylobacter (Vibrio) fetus*: characterization of antiphagocytic component. Infect. Immun. 11:517-525.

13. Tu, ZC, Dewhirst, FE, Blaser, MJ. 2001. Evidence that the *Campylobacter fetus* sap locus is an ancient genomic constituent with origins before mammals and reptiles diverged. Infect. Immun. 69:2237-2244.

14. Gilbert, MJ, Miller, WG, Yee, E, Zomer, AL, van der Graaf-van Bloois, L, Fitzgerald, C, Forbes, KJ, Meric, G, Sheppard, SK, Wagenaar, JA, Duim, B. 2016. Comparative genomics of *Campylobacter fetus* from reptiles and mammals reveals divergent evolution in host-associated lineages. Genome biology and evolution. 8:2006--2019.

15. Garcia, MM, Lutze-Wallace, CL, Denes, AS, Eaglesome, MD, Holst, E, Blaser, MJ. 1995. Protein shift and antigenic variation in the S-layer of *Campylobacter fetus* subsp. *venerealis* during bovine infection accompanied by genomic rearrangement of *sapA* homologs. J. Bacteriol. 177:1976-1980.

16. Thompson, SA, Blaser, MJ. 2008. Pathogenesis of *Campylobacter fetus*, p. 401--427. *In* Nachamkin, I, Szymanski, CM, Blaser, MJ (eds.), *Campylobacter*, 3rd ed. ASM Press, Washington, DC.

17. Corbeil, LB, Corbeil, RR, Winter, AJ. 1975. Bovine venereal vibriosis: activity of inflammatory cells in protective immunity. Am. J. Vet. Res. 36:403-406.

18. Blaser, MJ, Smith, PF, Hopkins, JA, Heinzer, I, Bryner, JH, Wang, WL. 1987. Pathogenesis of *Campylobacter fetus* infections: serum resistance associated with high-molecular-weight surface proteins. J. Infect. Dis. 155:696-706.

19. Perez Perez, GI, Blaser, MJ. 1985. Lipopolysaccharide characteristics of pathogenic campylobacters. Infect. Immun. 47:353-359.

20. OIE. 2013. Terrestrial Animal Health Code. 2013.

21. World Organisation for Animal Health. http://www.oie.int.

22. Mshelia, GD, Amin, JD, Woldehiwet, Z, Murray, RD, Egwu, GO. 2010. Epidemiology of bovine venereal campylobacteriosis: geographic distribution and recent advances in molecular diagnostic techniques. Reprod. Domest. Anim. 45:e221-30.

23. Molina, L, Perea, J, Meglia, G, Angon, E, Garcia, A. 2013. Spatial and temporal epidemiology of bovine trichomoniasis and bovine genital campylobacteriosis in La Pampa province (Argentina). Prev. Vet. Med. 110:388-394.

24. Irons, PC, Schutte, AP, van der Walt, ML, Bishop, GC. 2004. Genital campylobacteriosis in cattle, p. 1359-1468. *In* Coetzer, JAW, Trustin, RC (eds.), Infectious Diseases of Livestock, 2nd ed.

25. Akhtar, S, Riemann, HP, Thurmond, MC, Franti, CE. 1993. The association between antibody titres against *Campylobacter fetus* and milk production efficiency in dairy cattle. Vet. Res. Commun. 17:183-191.

26. McMillen, L, Fordyce, G, Doogan, VJ, Lew, AE. 2006. Comparison of culture and a novel 5' Taq nuclease assay for direct detection of *Campylobacter fetus* subsp. *venerealis* in clinical specimens from cattle. J Clin Microbiol. 44:938-45.

27. European Union. EUR-Lex, Directive 88/407/EEC.

28. Oyarzabal, OA, Wesley, IV, Harmon, KM, Schroeder-Tucker, L, Barbaree, JM, Lauerman, LH, Backert, S, Conner, DE. 1997. Specific identification of *Campylobacter fetus* by PCR targeting variable regions of the 16S rDNA. Vet. Microbiol. 58:61-71.

29. Dennis, SM. 1967. The possible role of the raven in the transmission of ovine vibriosis. Aust. Vet. J. 43:45-48.

30. Harvey, S, Greenwood, JR. 1985. Isolation of *Campylobacter fetus* from a pet turtle. J. Clin. Microbiol. 21:260-261.

31. Meinershagen, WA, Waldhalm, DG, Frank, FW, Scrivner, LH. 1965. Magpies as a reservoir of infection for ovine vibriosis. J. Am. Vet. Med. Assoc. 147:843-845.

32. Tu, ZC, Zeitlin, G, Gagner, JP, Keo, T, Hanna, BA, Blaser, MJ. 2004. *Campylobacter fetus* of reptile origin as a human pathogen. J. Clin. Microbiol. 42:4405-4407.

33. Watson, WA, Hunter, D, Bellhouse, R. 1967. Studies on vibrionic infection of sheep and carrion crows. Vet. Rec. 81:220-225.

34. Bryner, JH, O'Berry, PA, Frank, AH. 1964. Vibrio Infection of the Digestive Organs of Cattle. Am J Vet Res. 25:1048-50.

35. Elazhary, MASY. 1968. An assay of isolation and differential identification of some animal vibrios and of elucidation of their pathological significance. Meded. Veeartsschool. Rijksuniv. Gent. 1-80.

36. Garcia, MM, Eaglesome, MD, Rigby, C. 1983. Campylobacters important in veterinary medicine. Veterinary bulletin. 53:793-818.

37. Clark, BL. 1971. Review of bovine vibriosis. Aust. Vet. J. 47:103-107.

38. Kahn, CM. 2005. The Merck Veterinary Manual. Merck & Co, New Jersey.

39. Cipolla, AL, Casaro, AP, Terzolo, HR, Estela, ES, Brooks, BW, Garcia, MM. 1994. Persistence of *Campylobacter fetus* subspecies venerealis in experimentally infected heifers. Vet. Rec. 134:628.

40. Schutte, AP. 1969. Some aspects of *Vibrio fetus* in bulls, p. 13:88. *In* Anonymous Meded. Veeartsschool. Rijksuniv. Gent.

41. Wagenaar, JA, van Bergen, MA, Blaser, MJ, Tauxe, RV, Newell, DG, van Putten, JP. 2014. *Campylobacter fetus* infections in humans: exposure and disease. Clin. Infect. Dis. 58:1579-1586.

42. OIE. 2012. Bovine Genital Campylobacteriosis, p. 652. *In* Anonymous Manual of Diagnostic Tests and Vaccines for Terrestrial Animals (mammal, birds and bees), 7th ed. Office International des Epizooties, Paris.

43. Skirrow, MB. 1977. *Campylobacter enteritis*: a "new" disease. Br. Med. J. 2:9-11.

44. van Bergen, MAP, Linnane, S, van Putten, JP, Wagenaar, JA. 2005. Global detection and identification of *Campylobacter fetus* subsp. *venerealis*. Rev Sci Tech. 24:1017-26.

45. Lastovica, AJ, Le Roux, E. 2003. Optimal detection of *Campylobacter* spp in stools. J. Clin. Pathol. 56:480.

46. Hum, S, Brunner, J, McInnes, A, Mendoza, G, Stephens, J. 1994. Evaluation of cultural methods and selective media for the isolation of *Campylobacter fetus* subsp. *venerealis* from cattle. Aust. Vet. J. 71:184-186.

47. Atabay, HI, Corry, JE. 1998. Evaluation of a new arcobacter enrichment medium and comparison with two media developed for enrichment of *Campylobacter* spp. Int. J. Food Microbiol. 41:53-58.

48. Monke, HJ, Love, BC, Wittum, TE, Monke, DR, Byrum, BA. 2002. Effect of transport enrichment medium, transport time, and growth medium on the detection of *Campylobacter fetus* subsp. *venerealis*. J. Vet. Diagn. Invest. 14:35-39.

49. On, SL. 1996. Identification methods for campylobacters, helicobacters, and related organisms. Clin Microbiol Rev. 9:405-22.

50. Schulze, F, Bagon, A, Muller, W, Hotzel, H. 2006. Identification of *Campylobacter fetus* subspecies by phenotypic differentiation and PCR. J. Clin. Microbiol. 44:2019-2024.

51. Wesley, IV, Wesley, RD, Cardella, M, Dewhirst, FE, Paster, BJ. 1991. Oligodeoxynucleotide probes for *Campylobacter fetus* and *Campylobacter* hyointestinalis based on 16S rRNA sequences. J. Clin. Microbiol. 29:1812-1817.

52. Blom, K, Patton, CM, Nicholson, MA, Swaminathan, B. 1995. Identification of *Campylobacter fetus* by PCR-DNA probe method. J. Clin. Microbiol. 33:1360-1362.

53. Casademont, I, Bizet, C, Chevrier, D, Guesdon, JL. 2000. Rapid detection of *Campylobacter fetus* by polymerase chain reaction combined with non-radioactive hybridization using an oligonucleotide covalently bound to microwells. Mol. Cell. Probes. 14:233-240.

54. Eaglesome, MD, Sampath, MI, Garcia, MM. 1995. A detection assay for *Campylobacter fetus* in bovine semen by restriction analysis of PCR amplified DNA. Vet. Res. Commun. 19:253-263.

55. Cardarelli-Leite, P, Blom, K, Patton, CM, Nicholson, MA, Steigerwalt, AG, Hunter, SB, Brenner, DJ, Barrett, TJ, Swaminathan, B. 1996. Rapid identification of *Campylobacter* species by restriction fragment length polymorphism analysis of a PCR-amplified fragment of the gene coding for 16S rRNA. J. Clin. Microbiol. 34:62-67.

56. Gorkiewicz, G, Feierl, G, Schober, C, Dieber, F, Kofer, J, Zechner, R, Zechner, EL. 2003. Species-specific identification of campylobacters by partial 16S rRNA gene sequencing. J. Clin. Microbiol. 41:2537-2546.

57. Salama, SM, Tabor, H, Richter, M, Taylor, DE. 1992. Pulsed-field gel electrophoresis for epidemiologic studies of *Campylobacter hyointestinalis* isolates. J Clin Microbiol. 30:1982-4.

58. Vargas, AC, Costa, MM, Vainstein, MH, Kreutz, LC, Neves, JP. 2003. Phenotypic and molecular characterization of bovine *Campylobacter fetus* strains isolated in Brazil. Vet. Microbiol. 93:121-132.

59. Inglis, GD, Kalischuk, LD. 2003. Use of PCR for direct detection of *Campylobacter* species in bovine feces. Appl. Environ. Microbiol. 69:3435-3447.

60. Yamazaki, W, Taguchi, M, Ishibashi, M, Nukina, M, Misawa, N, Inoue, K. 2009. Development of a loop-mediated isothermal amplification assay for sensitive and rapid detection of *Campylobacter fetus*. Vet. Microbiol. 136:393-396.

61. Yamazaki, W, Taguchi, M, Misawa, N. 2010. Development of loop-mediated isothermal amplification and PCR assays for rapid and simple detection of *Campylobacter fetus* subsp. *venerealis*. Microbiol. Immunol. 54:398-404.

62. Wagenaar, JA, van Bergen, MAP, Newell, DG, Grogono-Thomas, R, Duim, B. 2001. Comparative study using amplified fragment length polymorphism fingerprinting, PCR genotyping, and phenotyping to differentiate *Campylobacter fetus* strains isolated from animals. J Clin Microbiol. 39:2283-6.

63. van Bergen, MAP, Dingle, KE, Maiden, MC, Newell, DG, van der Graaf-Van Bloois, L, van Putten, JP, Wagenaar, JA. 2005. Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. J Clin Microbiol. 43:5888-98.

64. Asakura, M, Samosornsuk, W, Taguchi, M, Kobayashi, K, Misawa, N, Kusumoto, M, Nishimura, K, Matsuhisa, A, Yamasaki, S. 2007. Comparative analysis of cytolethal distending toxin (*cdt*) genes among *Campylobacter jejuni*, *C. coli* and *C. fetus* strains. Microb. Pathog. 42:174-183.

65. Chaban, B, Musil, KM, Himsworth, CG, Hill, JE. 2009. Development of cpn60-based real-time quantitative PCR assays for the detection of 14 *Campylobacter* species and application to screening of canine fecal samples. Appl. Environ. Microbiol. 75:3055-3061.

66. Bastyns, K, Chapelle, S, Vandamme, P, Goossens, H, de Wachter, R. 1994. Species-specific Detection of Campylobacters Important in Veterinary Medecine by PCR Amplification of 23S rDNA Areas. System. Appl. Microbiol. 17:563-568.

67. Moolhuijzen, PM, Lew-Tabor, AE, Wlodek, BM, Aguero, FG, Comerci, DJ, Ugalde, RA, Sanchez, DO, Appels, R, Bellgard, M. 2009. Genomic analysis of *Campylobacter fetus* subspecies: identification of candidate virulence determinants and diagnostic assay targets. BMC Microbiol. 9:86.

68. Iraola, G, Hernandez, M, Calleros, L, Paolicchi, F, Silveyra, S, Velilla, A, Carretto, L, Rodriguez, E, Perez, R. 2012. Application of a multiplex PCR assay for *Campylobacter fetus* detection and subspecies differentiation in uncultured samples of aborted bovine fetuses. J. Vet. Sci. 13:371-376.

69. Gorkiewicz, G, Kienesberger, S, Schober, C, Scheicher, SR, Gully, C, Zechner, R, Zechner, EL. 2010. A genomic island defines subspecies-specific virulence features of the host-adapted pathogen *Campylobacter fetus* subsp. *venerealis*. J. Bacteriol. 192:502-17.

70. van Bergen, MAP, Simons, G, van der Graaf-van Bloois, L, van Putten, JP, Rombout, J, Wesley, I, Wagenaar, JA. 2005. Amplified fragment length polymorphism based identification of genetic markers and novel PCR assay for differentiation of *Campylobacter fetus* subspecies. J. Med. Microbiol. 54:1217-24.

71. Hum, S, Quinn, K, Brunner, J, On, SL. 1997. Evaluation of a PCR assay for identification and differentiation of *Campylobacter fetus* subspecies. Aust. Vet. J. 75:827-31.

72. Muller, W, Hotzel, H, Schulze, F. 2003. Identification and differentiation of *Campylobacter fetus* subspecies by PCR. Dtsch. Tierarztl. Wochenschr. 110:55-59.

73. Schmidt, T, Venter, EH, Picard, JA. 2010. Evaluation of PCR assays for the detection of *Campylobacter fetus* in bovine preputial scrapings and the identification of subspecies in South African field isolates. J. S. Afr. Vet. Assoc. 81:87-92.

74. Chaban, B, Chu, S, Hendrick, S, Waldner, C, Hill, JE. 2012. Evaluation of a *Campylobacter fetus* subspecies *venerealis* real-time quantitative polymerase chain reaction for direct analysis of bovine preputial samples. Can. J. Vet. Res. 76:166-173.

75. Spence, RP, Bruce, IR, McFadden, AM, Hill, FI, Tisdall, D, Humphrey, S, van der Graaf, L, van Bergen, MA, Wagenaar, JA. 2011. Cross-reaction of a *Campylobacter fetus* subspecies *venerealis* real-time PCR. Vet. Rec. 168:131.

76. Abril, C, Vilei, EM, Brodard, I, Burnens, A, Frey, J, Miserez, R. 2007. Discovery of insertion element ISCfe1: a new tool for *Campylobacter fetus* subspecies differentiation. Clin. Microbiol. Infect. 13:993-1000.

77. Salama, SM, Garcia, MM, Taylor, DE. 1992. Differentiation of the subspecies of *Campylobacter fetus* by genomic sizing. Int. J. Syst. Bacteriol. 42:446-50.

78. Salama, SM, Newnham, E, Chang, N, Taylor, DE. 1995. Genome map of *Campylobacter fetus* subsp. *fetus* ATCC 27374. FEMS Microbiol. Lett. 132:239-245.

79. Fujita, M, Amako, K. 1994. Localization of the *sapA* gene on a physical map of *Campylobacter fetus* chromosomal DNA. Arch. Microbiol. 162:375-380.

80. Fujita, M, Fujimoto, S, Morooka, T, Amako, K. 1995. Analysis of strains of *Campylobacter fetus* by pulsed-field gel electrophoresis. J. Clin. Microbiol. 33:1676-1678.

81. Linton, D, Owen, RJ, 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.

82. Willoughby, K, Nettleton, PF, Quirie, M, Maley, MA, Foster, G, Toszeghy, M, Newell, DG. 2005. A multiplex polymerase chain reaction to detect and differentiate *Campylobacter fetus* subspecies *fetus* and *Campylobacter fetus* subspecies *venerealis*: use on UK isolates of *C. fetus* and other *Campylobacter spp*. J. App.l Microbiol. 99:758-66.

83. Wang, G, Clark, CG, Taylor, TM, Pucknell, C, Barton, C, Price, L, Woodward, DL, Rodgers, FG. 2002. Colony multiplex PCR assay for identification and differentiation of *Campylobacter jejuni*, *C. coli*, *C. lari*, *C. upsaliensis*, and *C. fetus* subsp. *fetus*. J. Clin. Microbiol. 40:4744-7.

Chapter 2

Evaluation of molecular assays for identification *Campylobacter fetus* species and su species and development of a *C. fetus* specific real-time PCR assay

Linda van der Graaf - van Bloois[a,c], Marcel A. P. van Bergen[b,c], Fimme J. van der Wal[b,c], Albert G. de Boer[b,c], Birgitta Duim[a,c], Tracy Schmidt[d], Jaap A. Wagenaar[a,b,c]

[a] Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, P.O. Box 80165, 3508 TD Utrecht, The Netherlands
[b] Central Veterinary Institute of Wageningen UR, P.O. Box 65, 8200 AB Lelystad, The Netherlands
[c] WHO Collaborating Center for Campylobacter and OIE Reference Laboratory for Campylobacteriosis
[d] Allerton Provincial Veterinary Laboratory, Private Bag X2, Cascades, 3202 South Africa

Journal of Microbiological Methods. 2013, 95(1):93-97

A stract

Phenotypic differentiation between *Campylobacter fetus* (*C. fetus*) subspecies *fetus* and *C. fetus* subspecies *venerealis* is hampered by poor reliability and reproducibility of biochemical assays. AFLP (Amplified Fragment Length Polymorphism) and MLST (Multi Locus Sequence Typing) are the molecular standards for *C. fetus* subspecies identification, but these methods are laborious and expensive. Several PCR assays for *C. fetus* subspecies identification have been described, but a reliable comparison of these assays is lacking.

The aim of this study was to evaluate the most practical and routinely implementable published PCR assays designed for *C. fetus* species and subspecies identification. The sensitivity and specificity of the assays was calculated by using an extensively characterized and diverse collection *C. fetus* strains. AFLP and MLST identification was used as reference. Two PCR assays were able to identify *C. fetus* strains correctly at species level. The *C. fetus* species identification target, gene *nahE,* of one PCR assay was used to develop a real-time PCR assay with 100% sensitivity and 100% specificity, but the development of a subspecies venerealis specific real-time PCR (IS*CfeI*) failed due to sequence variation of the target insertion sequence and prevalence in other *Campylobacter* species. None of the published PCR assays was able to identify *C. fetus* strains correctly at subspecies level. Molecular analysis by AFLP or MLST is still recommended to identify *C. fetus* isolates at subspecies level.

Introduction

Bovine Genital Campylobacteriosis (BGC) is a syndrome characterized by fertility problems in cattle (1). The causative agent of BGC is *Campylobacter fetus* subspecies *venerealis* (Cfv). Cfv is restricted to the genital tract of both male and female cattle and includes a variant, Cfv biovar intermedius (Cfvi) (2). This subspecies is venereally transmitted only (3). The other *C. fetus* subspecies is *Campylobacter fetus* subspecies *fetus* (Cff), which occurs mainly in the intestinal tract of cattle and sheep (3). There are three transmission routes for this subspecies: faecal-oral, ascending genital infections most probably due to intestinal colonization, and venereal transmission (2, 3). Cff can cause sporadic infections in humans, abortion in cattle and sheep, and is incidentally isolated from a variety of sites in different hosts (2, 3).

A criterion for an effective BGC control program is the reliable differentiation of *Campylobacter fetus* subspecies *venerealis* from the closely related *Campylobacter fetus* subspecies *fetus.* The World Animal Health Organisation (OIE) Terrestrial Code (4) describes the health measures to be taken by the veterinary authorities. In addition to these standards and recommendations, countries can have their local or regional regulatory measures as well, for example the EU regulations of the European Union (http://eur-lex.europa.eu/nl/index.htm). Several countries have been successful in eradicating BGC, whereas in many other countries BGC is still endemic (5). Due to trade of livestock, individual cases and outbreaks with Cfv are occasionally detected in countries with the BGC-free status. Depending on the regulations concerning BGC, there are severe consequences for artificial insemination (AI) stations when a Cfv positive animal is identified: temporary closure and destruction of semen that was produced since the last negative monitoring results. Misidentification of Cfv as Cff may result into the spread of Cfv into cattle populations. Misidentification of Cff as Cfv may result in economic losses. Taken together, the consequences of subspecies misidentification can be severe, showing that indisputably *C. fetus* subspecies identification is crucial for BGC control programs.

The two *Campylobacter fetus* subspecies can be differentiated phenotypically by the 1% glycine tolerance test (2), but this test has a poor reproducibility (6). The hydrogen sulphide (H_2S) test is also described to differentiate Cff (H_2S positive) from Cfv (H_2S negative) (2). However, as Cfv biovar intermedius strains are positive in the H_2S test, this assay is not suitable to differentiate between the two subspecies (6). To overcome these problems, molecular assays for subspecies differentiation have been introduced.

In literature, several molecular assays for *C. fetus* species and subspecies identification have been described that are laborious or not suitable for standard routine diagnostic laboratories; hybridisation assays (7-9), restriction fragment length polymorphism (RFLP) (10), 16S sequencing (11) and pulsed field gel electrophoresis (PFGE) (12, 13). More practical assays for identification have been described, based on nucleic acid amplification such as nested PCR, targeting 16S and 23S rRNA (14) and loop-mediated isothermal amplification (LAMP) for *C. fetus* species and subspecies venerealis detection (15, 16). However, nested PCR is not preferred for routine diagnostics, since this type of assay is sensitive for contamination, whereas LAMP is relatively unknown and uncommon in diagnostic laboratories; importantly, the target of this LAMP assay (insertion element IS*Cfe1*) is not suitable to differentiation between the two *C. fetus* subspecies, as shown in this study. The molecular typing methods Amplified Fragment Length Polymorphism (AFLP) (17) and Multi Locus Sequence Typing (MLST) (6) proved to be able to differentiate the two subspecies reliably and can be used for identification of species and subspecies, but these tests are laborious and impractical for routine use.

A preferred platform for routine diagnostics is PCR, since this type of assay is fast, easy and reliable. A number of PCR assays have been developed to identify *C. fetus* species. A PCR assay targeting 16S rRNA (18) differentiated *C. fetus* from other *Campylobacter* species, but only a single base mismatch in the target 16S rRNA region differentiates the two *C. fetus* subspecies, making this region too conserved to enable successful differentiation of the two subspecies (18). A PCR assay targeting 23S rRNA identifies *C. fetus* species, but attempts to design primers on this target to discriminate between the two subspecies failed (19). PCR assays targeting the *cdt*-genes (20) and *cpn60*-gene (21) are described for *C. fetus* species identification, but these assays are not able to discriminate between the two subspecies. Since subspecies venerealis is described as the

a le . Evaluated PCR assays

Code	Reference	Identifcation	Target genes	Assay
A	Abril *et al*., 2007	*C. fetus* Cfv	*nahE* IS*Cfe1*	Multiplex PCR
B	van Bergen *et al.*, 2005b	Cfv	hypothetical protein	PCR
H	Hum *et al.*, 1997	*C. fetus* Cfv	*cstA* *parA*	Multiplex PCR
M	McMillen *et al.*, 2006	Cfv	*parA*	Real-time PCR
W	Wang *et al*., 2002	Cff	*sapB2*	PCR
	Tu *et al*., 2005	Cff / Cfv	not applicable	RAPD-PCR

causative agent of BGC, it is very important for routine diagnostic laboratories to differentiate between the two subspecies and assays that only identify *C. fetus* to the species level are not sufficient. Lastly, two subspecies venerealis specific PCRs have been described, targeting virulence genes (22, 23). However, Gorkiewicz *et al*., 2010 showed that these virulence genes are not present in all Cfv strains and are therefore not suitable for subspecies venerealis identification.

For this study, PCR assays that were reported to be able to differentiate the two subspecies were selected (Table 1). The multiplex PCR assay of Hum *et al*., 1997 is included to evaluate the suitability of the target genes *cstA* and *parA*, therefore the modified PCR assays using the same targets are excluded (23-26). Since McMillen *et al*., 2006 used real-time PCR instead of conventional PCR, this assay is included in this study.

The first goal of this study was to evaluate the selected PCR assays for both species and subspecies identification of *C. fetus*. The sensitivity and specificity of the assays was calculated by using an extensively characterized collection of strains from different sources and geographical regions. The second goal of this study was to improve *C. fetus* diagnostics. Based on the evaluation of published PCR assays, the gene *nahE* was used to develop a real-time PCR assay for *C. fetus* species identification and insertion sequence IS*CfeI* for subspecies venerealis identification.

2 Materials and methods

2.1 Strains

The PCR assays were tested with a collection of 143 *C. fetus* strains (83 Cff and 60 Cfv) (Supplemental Table 1) from different sources and geographical regions. Twelve non-fetus *Campylobacter* reference strains (Supplemental Table 1) were included to verify the specificity of the assays; these strains represent non-fetus *Campylobacter* species that are occasionally isolated from bovine or ovine samples.

Strains were grown for 48 h on heart-infusion agar supplemented with 5% sheep blood (Biotrading, Mijdrecht, The Netherlands) under micro-aerobic conditions (6%O_2, 7%CO_2, 7%H_2, 80%N_2, Anoxomat, Mart Microbiology, Lichtenvoorde, The Netherlands). Chromosomal DNA of all strains was isolated with the PureGene kit (Gentra PureGene DNA isolation kit, Qiagen). PCR A (see section 2.2) and the new real-time PCR assay (*nahE*) (see section 2.4) were also tested with boiled cell lysates. These boiled cell lysates were made by suspending cells (OD_{600nm} = 0.1) in sterile distilled water and subsequent heating at 95˚C for 10 minutes.

All *C. fetus* and non-fetus *Campylobacter* strains were characterized by AFLP as described (17). The *C. fetus* strains were also characterized by phenotypic assays, i.e. growth in presence of 1% glycine and H_2S production, and MLST (6).

2.2 PCR assays

The PCR assays, labeled A, B, H, and W (Table 1) were performed with the thermal cycling programs as described in the original publications (28-32) with the following reaction-mixture modifications: PCR assays were performed with 2.5 U AmpliTaq DNA polymerase (Applied Biosystems) and the supplied PCR buffer II and 2.5 mM $MgCl_2$. PCR H was performed with 2.5 U AmpliTaq DNA polymerase (Applied Biosystems) and PCR buffer as described in the original publication. All reactions were performed in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems). Real-time PCR M (33) was performed with the thermal cycling program as described in the original publication in a 7500 Fast Real-Time PCR System (Applied Biosystems), by using 2x TaqMan Fast Universal PCR Master Mix (Applied Biosystems) with primers, probe and DNA concentration as described in the original publication.

The new real-time PCR assays, described in section 2.4, were performed in 20 µl on a 7500 Fast Real-Time PCR System (Applied Biosystems) in Fast 7500 run mode, using the following thermal cycler protocol: 95˚C for 20 sec, followed by 40 cycles of 95˚C for 3 sec and 60˚C for 30 sec. Each PCR reaction contained 10 µl 2x TaqMan Fast Universal PCR Master Mix (Applied Biosystems), 1 µl (12 pmol/µl) of each primer, 0.5 µl (8 pmol/µl) probe, 2 µl of either DNA template or cell lysate, adjusted to 20 µl with nuclease free-water. Using a threshold of 0.1, a result was designated positive for Ct's below 30.

a le 2 Real-time PCR primers and probes for *C. fetus* species en subspecies *venerealis* identification

	5'→3'	
nahE-F	TGTTATGGTGATCAAAATAGCTGTTG	Forward primer
nahE-R	GAGCTGTTTTTATGGCTACTCTTTTTTTA	Reverse primer
nahE-P	*FAM*- TGTATATGCACTTTTAGCAACTT-*NFQ*	MGB probe
ISC1-F	AGGCGAAGAGAATGTTAAATTTGAA	Forward primer
ISC1-R	CCATAAAGCCTAGCTGAAAAAACTG	Reverse primer
ISC1-P	*VIC*- CCAAAGATGTCTTAGAAATA-*NFQ*	MGB probe
ISC2-F	TTCAAAAGCTCTTGGGGTTAC	Forward primer
ISC2-R	AAAGCCTTGTTTAGAACAATATAACTC	Reverse primer
ISC2-P	*VIC*- ACTCGTGGTGGAGAGCGTAG-*NFQ*	MGB probe

2.3 Sequencing of insertion sequence ISCfe1

The sequence of IS*Cfe1* of Cfv strain NCTC 10354 (Genbank AM260752.1) was used to design primers with Primer Express 3.0 (Applied Biosystems), outside the Cfv specific target region of PCR A (28). The resulting fragment, containing IS*Cfe1,* was sequenced for one Cfv strain (B27) using Sanger sequencing. Roche 454 sequence data of 16 Cfv strains (manuscript in preparation) was used to BLAST the insertion sequences, with the sequence of IS*Cfe1* of Cfv strain NCTC 10354 as scaffold. The IS sequences, obtained with BLAST analysis from whole genome sequences of 16 Cfv strains (including Zaf 3) and the IS*Cfe1* sequence of strain B27, were aligned and compared to serve as target for designing a new real-time PCR.

2.4 New real-time PCR assay for species and subspecies identification

To develop a new real-time PCR assay, the *C. fetus* species specific target gene *nahE* and the subspecies venerealis specific insertion sequence (IS) IS*Cfe1* (28) were used. With Primer Express 3.0 (Applied Biosystems), primers and a Taqman NFQ-MGB probe were designed for *C. fetus* species identification based on gene *nahE* (Table 2). For subspecies venerealis identification, real-time primers ISC1-F, ISC1-R and probe ISC1-P (Table 2) were developed within IS*Cfe1* from bp 585 to 668. Within the conserved region of the IS, a second set of primers ISC2-F, ISC2-R and probe ISC2-P (Table 2) were developed, covering the region from bp 132 to 387. The program and reaction mixtures for the new real-time PCRs are described in section 2.2.

a le 3 Estimated sensitivity and specificity of the PCR assays with AFLP and MLST subspecies identification as reference

Assay	Identification	Sensitivity		(95% CI)	Specificity		(95% CI)
Abril (A)	*C. fetus*	143/143[a]	100%	(97-100)	12/12[d]	100%	(76-100)
	Cfv	58/60[b]	97%	(87-99)	95/95[e]	100%	(96-100)
Van Bergen (B)	Cfv	27/60[b]	45%	(33-57)	95/95[e]	100%	(96-100)
Hum (H)	*C. fetus*	143/143[a]	100%	(97-100)	12/12[d]	100%	(76-100)
	Cfv	35/60[b]	58%	(46-70)	79/95[e]	83%	(74-89)
McMillen (M)	Cfv	32/60[b]	53%	(41-65)	95/95[e]	100%	(96-100)
Wang (W)	Cff	63/83[c]	76%	(66-84)	52/72[f]	72%	(61-81)
Real-time PCR (this study)	*C. fetus*	143/143[a]	100%	(97-100)	12/12[d]	100%	(76-100)

[a] Number of *C. fetus* strains: 143
[b] Number of *C. fetus* subsp. *venerealis* strains: 60
[c] Number of *C. fetus* subsp. *fetus* strains: 83
[d] Number of non-fetus *Campylobacter* strains: 12
[e] Number of non-Cfv strains: 95 (83[c] + 12[d])
[f] Number of non-Cff strains: 72 (60[b] + 12[d])

3 Results

3.1 PCR assays

One published RAPD-PCR assay (30) was excluded from the evaluation as a preliminary evaluation of this assay yielded a very low sensitivity with only 3/7 Cff and 3/8 Cfv strains correctly identified (data not shown).

The results of all evaluated PCR assays with 143 *C. fetus* strains (83 Cff and 60 Cfv) and 12 non-fetus *Campylobacter* strains are shown in Supplemental Table 1. The specificity and sensitivity and their 95% confidence intervals (CI) (34) of the assays were calculated, using AFLP and MLST subspecies identification as reference (Table 3).

Two PCR assays showed 100% sensitivity and 100% specificity for *C. fetus* species identification, PCR A (28) and PCR H (29). None of the evaluated PCR assays was able to identify all *C. fetus* strains correctly at subspecies level. For subspecies identification, PCR A showed the highest sensitivity (97%) and specificity (100%). Only two Cfv strains, B27 and Zaf 3, were misidentified as Cff with this PCR. To investigate this deviation, the IS*Cfe1* sequences of these strains were determined (see section 3.2).

For multiplex PCR A (28) and the newly developed *C. fetus* specific real-time PCR (*nahE*), the results using purified DNA were identical to PCR results using boiled cell lysates (data not shown).

a le 4. Estimated copy numbers of insertion sequences in genome sequences

Strain	Subspecies	Source	Country	Copy number Isa	Copy number Isb
97/608	Cfv	Bovine	AR	7	1
84/112	Cfv	Bovine	USA	3	-
B10	Cfv	Bovine	US	5	-
B27	Cfv	Bovine	US	-	>1*
CCUG 33872	Cfv	Unknown	CZ	3	-
CCUG 33900	Cfv	Bovine	FR	5	-
LMG 6570	Cfv	Bovine	BE	3	-
Zaf 3	Cfv	Bovine	ZA	-	2
01/165	Cfvi	Bovine	AR	6	-
02/298	Cfvi	Bovine	AR	4	-
03/293	Cfvi	Bovine	AR	6	-
03/596	Cfvi	Bovine	AR	5	-
92/203	Cfvi	Bovine	AR	7	-
97/532	Cfvi	Bovine	AR	7	-
98/25	Cfvi	Bovine	AR	5	-
WBT 011/09	Cfvi	Unknown	UK	5	-
Zaf 65	Cfvi	Bovine	ZA	4	-

* data obtained by Sanger sequencing, estimation of copy numbers of IS in genome is not possible

3.2 Sequencing of insertion sequence ISCfe1

Alignment of the sequences to the ISC*fe1* sequence of strain NCTC 10354 indicated the existence of at least two different types of insertion sequences, ISa and ISb, sharing 98.7% sequence homology. BLAST analysis indicated that all Cfv strains contained different copy numbers of these insertion sequences, where the majority contained ISa and two strains exclusively contained ISb (B27 and Zaf3) (Table 4). One of the Cfv specific primers of PCR A, primer Cven-L (28), is located on a region where the sequences of ISa and ISb differ. The sequence differences result in five mismatches of primer Cven-L with ISb, which is likely to affect the misidentification of PCR A for Cfv strains B27 and Zaf3.

3.3 Real-time PCR assay for species and subspecies identification

The newly developed real-time PCR assay for *C. fetus* species identification, targeting gene *nahE*, performed excellent with 100% sensitivity and 100% specificity (Table 3). For subspecies venerealis identification, real-time primers and probe (ISC1) targeting the same region of ISC*fe1* as PCR A, resulted in misidentification of two Cfv strains (B27 and Zaf 3) as Cff. An alternative approach with primers and a probe (ISC2), targeting the conserved region of the IS, failed as two *C. hyointestinalis* reference strains were misidentified as subspecies venerealis. It has not been possible to develop primers on target ISC*fe1* that result in 100% sensitivity and 100% specificity for subspecies venerealis identification.

4 Discussion

In this study, the molecular typing assays AFLP and MLST were used as reference to evaluate PCR assays for *C. fetus* species and subspecies identification, since AFLP and MLST proved to be able to differentiate the two subspecies reliably. AFLP analysis is a reliable typing method to differentiate between Cff, Cfv and Cfvi. The MLST housekeeping gene sequences of *C. fetus* are very stable (6), which makes MLST a very robust typing method. Nevertheless, the resolution is slightly lower as Cfv and Cfvi strains cannot be distinguished by MLST, since they belong to the same sequence type (ST), i.e. ST-4, ST-7 and ST-12. When the MLST result of a *C. fetus* isolate indicate an existing ST, the subspecies can be reliably identified, but when a new ST is found, the subspecies cannot be identified and AFLP needs to be performed (6).

PCR A (28) is a multiplex PCR assay, designed to simultaneously identify strains for *C. fetus* species and subspecies venerealis. The *C. fetus* species target of this PCR is gene

nahE. PCR A was able to identify *C. fetus* strains correctly at species level. For subspecies venerealis identification, this PCR targets the insertion sequence IS*CfeI*, which is present in Cfv, but absent in Cff. It was shown in this study that at least two different ISs (ISa and ISb) can occur in subspecies venerealis strains. In two Cfv strains containing only ISb, the venerealis specific primers were not able to amplify the target IS. The fact that strains from different continents (Africa and North America) carry only ISb, shows that the occurrence of the ISb variant does not depend on the geographic origin of the strains. This result also stresses the importance of inclusion of isolates from different geographical origin when an assay is validated for global use. The inability to identify all Cfv strains correctly renders this PCR unsuitable for subtyping *C. fetus* strains.

PCR B (31) was designed to identify subspecies venerealis. With PCR B, all Cfv *sensu stricto* strains were correctly identified, but the assay did not detect any of the Cfv biovar intermedius strains. Therefore, PCR B is not suitable for diagnostic purposes. If, in the future, a reliable PCR assay becomes available that detects all Cfv including biovar intermedius strains, PCR B may be used as an addition epidemiological tool to distinguish Cfv *sensu strict* strains from biovar intermedius strains.

PCR H (29) is a multiplex PCR assay, designed to identify *C. fetus* species and subspecies venerealis. The *C. fetus* species identification of this PCR, based on target gene *cstA*, performs excellent. For subspecies venerealis identification, PCR H and PCR M (33) use the same target, gene *parA*. The target gene *parA* can be located on a transferable genomic island or on a plasmid, and strains can contain multiple copy numbers of this gene (35, 36). The sensitivity and specificity of the Cfv subspecies identification of PCR H and PCR M with target *parA* are poor, confirming the results from a previous publication (37). In a recent study, PCR M showed a positive result with a *Campylobacter hyointestinalis* strain isolated from a bull (38). This observation and the low sensitivity observed in this study render the *parA* gene and consequently all other PCR assays using this target (23-26); (27) unsuitable for Cfv identification.

PCR W (32) was designed to identify subspecies fetus by using *sapB2* as target. All Cfv strains are serotype A (with *sapA* genes) while Cff strains can be either serotype A or B (with either *sapA* or *sapB*). The reverse primer of this PCR is positioned in a variable region in the *sapB2* gene that remains highly homologous with *sapA2* (37), resulting in incorrect subspecies fetus identification. In the present study, PCR W showed low specificity which supported that this target is not suitable for identification of *C. fetus* subspecies.

One should realize that the sensitivity and specificity of the assays have been calculated

for the strain set used in this study. Evaluation with another set of strains may result into a different sensitivity and specificity of the evaluated assays.

A goal of this study was to improve *C. fetus* diagnostics. As real-time PCR is rapid, easy and less labour-intensive than traditional PCR assays, it is the preferred PCR format for routine diagnostics. Two PCR assays, PCR A and PCR H, were able to identify *C. fetus* strains correctly at species level. Since the subspecies identification of PCR A showed the highest sensitivity (97%) and specificity (100%), the targets of this PCR were used to develop a real-time PCR assay, targeting gene *nahE* for species identification and insertion sequence IS*Cfe1* for subspecies venerealis identification.

Development of a real-time PCR assay for subspecies venerealis identification failed, as it was not possible to develop primers and probe on target gene IS*Cfe1* with 100% sensitivity and 100% specificity. The *C. fetus* species identification of the newly developed real-time PCR (*nahE*) assay performed excellent with 100% sensitivity and 100% specificity, with both purified DNA and boiled cell lysates, resulting in a reliable and rapid assay for *C. fetus* species identification. This real-time PCR (*nahE*) assay can be used to identify the species fetus when a *Campylobacter* strain is isolated. As obtaining isolates from samples still requires a significant effort, the use of the new real-time PCR (*nahE*) assay directly on DNA material from clinical samples (e.g. boiled) would be the next step to improve *C. fetus* diagnostics. Research towards the development of new PCR assays based on recent genome sequences is needed in order to improve molecular diagnostic tools (e.g. real-time PCR) for identification of *C. fetus* subspecies.

5 Conclusion

Two PCR assays were able to correctly identify all *C. fetus* strains. The best performing PCR assay was converted into a real-time PCR assay, targeting gene *nahE* and insertion sequence IS*CfeI.* The *C. fetus* species specific real-time PCR (*nahE*) assay showed 100% sensitivity and 100% specificity, but the development of a subspecies venerealis specific real-time PCR (IS*CfeI*) failed due to sequence variation of the target insertion sequence and prevalence in other *Campylobacter* species. None of the PCR assays tested were able to identify *C. fetus* strains correctly at subspecies level. Of the described methods, molecular analysis by AFLP or MLST is the most effective method for identification of *C. fetus* isolates at subspecies level; a persistent need for an easy, rapid and reliable molecular assay for subspecies identification remains.

Acknowledgements

We are very grateful to Dr. W. Miller and E. Yee (Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA, USA) for generating the Roche 454 sequence data. Furthermore, we thank Dr. K. Clothier (California Animals Health and Food Safety Lab System, U.C. Davis, Davis, CA, USA) for kindly supplying *C. fetus* strains.

Supplemental data

Supplemental Table S1. PCR assay results of 143 *C. fetus* strains and 12 non-fetus *Campylobacter* strains
Supplemental Table S1 can be found online at https://goo.gl/NDrkIs

References

1. Dekeyser, J. 1984. *Campylobacter* infections in man and animals. J.P. Butzler (ed.) CRC Press Inc., Boca Raton, Florida. 181-191.

2. Véron, M., Chatelain,R. 1973. Taxonomy study of the genus *Campylobacter* Sebald and Verón and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Verón. International Journal of Systematic Bacteriology. 23:122-134.

3. Garcia, MM, Eaglesome, MD, Rigby, C. 1983. Campylobacters important in veterinary medicine. Veterinary bulletin. 53:793-818.

4. OIE. 2013. Terrestrial Animal Health Code. http://www.oie.int/international-standard-setting/terrestrial-code. Accession date: 23/04/2013.

5. OIE. 2013. OIE Wahid - Disease Information.
http://www.oie.int/wahis_2/public/wahid.php/Wahidhome/Home. Accession date: 23/04/2013.

6. van Bergen, MAP, Dingle, KE, Maiden, MC, Newell, DG, van der Graaf-van Bloois, L, van Putten, JP, Wagenaar, JA. 2005. Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. J. Clin. Microbiol. 43:5888-98.

7. Blom, K, Patton, CM, Nicholson, MA, Swaminathan, B. 1995. Identification of *Campylobacter fetus* by PCR-DNA probe method. J. Clin. Microbiol. 33:1360-1362.

8. Casademont, I, Bizet, C, Chevrier, D, Guesdon, JL. 2000. Rapid detection of *Campylobacter fetus* by polymerase chain reaction combined with non-radioactive hybridization using an oligonucleotide covalently bound to microwells. Mol. Cell. Probes. 14:233-240.

9. Wesley, IV, Wesley, RD, Cardella, M, Dewhirst, FE, Paster, BJ. 1991. Oligodeoxynucleotide probes for *Campylobacter fetus* and *Campylobacter hyointestinalis* based on 16S rRNA sequences. J. Clin. Microbiol. 29:1812-1817.

10. Eaglesome, MD, Sampath, MI, Garcia, MM. 1995. A detection assay for *Campylobacter fetus* in bovine semen by restriction analysis of PCR amplified DNA. Vet. Res. Commun. 19:253-263.

11. Gorkiewicz, G, Feierl, G, Schober, C, Dieber, F, Kofer, J, Zechner, R, Zechner, EL. 2003. Species-specific identification of campylobacters by partial 16S rRNA gene sequencing. J. Clin. Microbiol. 41:2537-2546.

12. Salama, SM, Garcia, MM, Taylor, DE. 1992. Differentiation of the subspecies of *Campylobacter fetus* by genomic sizing. Int. J. Syst. Bacteriol. 42:446-50.

13. Vargas, AC, Costa, MM, Vainstein, MH, Kreutz, LC, Neves, JP. 2003. Phenotypic and molecular characterization of bovine *Campylobacter fetus* strains isolated in Brazil. Vet. Microbiol. 93:121-132.

14. Inglis, GD, Kalischuk, LD. 2003. Use of PCR for direct detection of *Campylobacter* species in bovine feces. Appl. Environ. Microbiol. 69:3435-3447.

15. Yamazaki, W, Taguchi, M, Ishibashi, M, Nukina, M, Misawa, N, Inoue, K. 2009. Development of a loop-mediated isothermal amplification assay for sensitive and rapid detection of *Campylobacter fetus*. Vet. Microbiol. 136:393-396.

16. Yamazaki, W, Taguchi, M, Misawa, N. 2010. Development of loop-mediated isothermal amplification and PCR assays for rapid and simple detection of *Campylobacter fetus* subsp. *venerealis*. Microbiol. Immunol. 54:398-404.

17. Wagenaar, JA, van Bergen, MAP, Newell, DG, Grogono-Thomas, R, Duim, B. 2001. Comparative study using amplified fragment length polymorphism fingerprinting, PCR genotyping, and phenotyping to differentiate *Campylobacter fetus* strains isolated from animals. J. Clin. Microbiol. 39:2283-6.

18. Oyarzabal, OA, Wesley, IV, Harmon, KM, Schroeder-Tucker, L, Barbaree, JM, Lauerman, LH, Backert, S, Conner, DE. 1997. Specific identification of *Campylobacter fetus* by PCR targeting variable regions of the 16S rDNA. Vet. Microbiol. 58:61-71.

19. Bastyns, K, Chapelle, S, Vandamme, P, Goossens, H, de Wachter, R. 1994. Species-specific Detection of

Campylobacters Important in Veterinary Medecine by PCR Amplification of 23S rDNA Areas. System. Appl. Microbiol. 17:563-568.

20. Asakura, M, Samosornsuk, W, Taguchi, M, Kobayashi, K, Misawa, N, Kusumoto, M, Nishimura, K, Matsuhisa, A, Yamasaki, S. 2007. Comparative analysis of cytolethal distending toxin (*cdt*) genes among *Campylobacter jejuni*, *C. coli* and *C. fetus* strains. Microb. Pathog. 42:174-183.

21. Chaban, B, Musil, KM, Himsworth, CG, Hill, JE. 2009. Development of cpn60-based real-time quantitative PCR assays for the detection of 14 *Campylobacter* species and application to screening of canine fecal samples. Appl. Environ. Microbiol. 75:3055-3061.

22. Moolhuijzen, PM, Lew-Tabor, AE, Wlodek, BM, Aguero, FG, Comerci, DJ, Ugalde, RA, Sanchez, DO, Appels, R, Bellgard, M. 2009. Genomic analysis of *Campylobacter fetus* subspecies: identification of candidate virulence determinants and diagnostic assay targets. BMC Microbiol. 9:86.

23. Iraola, G, Hernandez, M, Calleros, L, Paolicchi, F, Silveyra, S, Velilla, A, Carretto, L, Rodriguez, E, Perez, R. 2012. Application of a multiplex PCR assay for *Campylobacter fetus* detection and subspecies differentiation in uncultured samples of aborted bovine fetuses. J. Vet. Sci. 13:371-376.

24. Muller, W, Hotzel, H, Schulze, F. 2003. Identification and differentiation of *Campylobacter fetus* subspecies by PCR. Dtsch. Tierarztl. Wochenschr. 110:55-59.

25. Schmidt, T, Venter, EH, Picard, JA. 2010. Evaluation of PCR assays for the detection of *Campylobacter fetus* in bovine preputial scrapings and the identification of subspecies in South African field isolates. J. S. Afr. Vet. Assoc. 81:87-92.

26. Schulze, F, Bagon, A, Muller, W, Hotzel, H. 2006. Identification of *Campylobacter fetus* subspecies by phenotypic differentiation and PCR. J. Clin. Microbiol. 44:2019-2024.

27. Chaban, B, Chu, S, Hendrick, S, Waldner, C, Hill, JE. 2012. Evaluation of a *Campylobacter fetus* subspecies venerealis real-time quantitative polymerase chain reaction for direct analysis of bovine preputial samples. Can. J. Vet. Res. 76:166-173.

28. Abril, C, Vilei, EM, Brodard, I, Burnens, A, Frey, J, Miserez, R. 2007. Discovery of insertion element IS*fe1*: a new tool for *Campylobacter fetus* subspecies differentiation. J. Clin. Microbiol. Infect. 13:993-1000.

29. Hum, S, Quinn, K, Brunner, J, On, SL. 1997. Evaluation of a PCR assay for identification and differentiation of *Campylobacter fetus* subspecies. Aust. Vet J. 75:827-31.

30. Tu, ZC, Eisner, W, Kreiswirth, BN, Blaser, MJ. 2005. Genetic divergence of *Campylobacter fetus* strains of mammal and reptile origins. J. Clin. Microbiol. 43:3334-40.

31. van Bergen, MAP, Simons, G, van der Graaf-van Bloois, L, van Putten, JP, Rombout, J, Wesley, I, Wagenaar, JA. 2005. Amplified fragment length polymorphism based identification of genetic markers and novel PCR assay for differentiation of *Campylobacter fetus* subspecies. J. Med. Microbiol. 54:1217-24.

32. Wang, G, Clark, CG, Taylor, TM, Pucknell, C, Barton, C, Price, L, Woodward, DL, Rodgers, FG. 2002. Colony multiplex PCR assay for identification and differentiation of *Campylobacter jejuni*, *C. coli*, *C. lari*, *C. upsaliensis*, and *C. fetus* subsp. *fetus*. J. Clin. Microbiol. 40:4744-7.

33. McMillen, L, Fordyce, G, Doogan, VJ, Lew, AE. 2006. Comparison of culture and a novel 5' Taq nuclease assay for direct detection of *Campylobacter fetus* subsp. *venerealis* in clinical specimens from cattle. J. Clin. Microbiol. 44:938-45.

34. Dohoo, I, Martin, W, Stryhn, H. 2003. Veterinary Epidemiologic Research. Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Canada.

35. Abril, C, Brodard, I, Perreten, V. 2010. Two novel antibiotic resistance genes, tet(44) and ant(6)-Ib, are located within a transferable pathogenicity island in *Campylobacter fetus* subsp. *fetus*. Antimicrob. Agents Chemother. 54:3052-5.

36. Gorkiewicz, G, Kienesberger, S, Schober, C, Scheicher, SR, Gully, C, Zechner, R, Zechner, EL. 2010. A genomic island defines subspecies-specific virulence features of the host-adapted pathogen *Campylobacter fetus* subsp.

venerealis. J. Bacteriol. 192:502-17.

37. Willoughby, K, Nettleton, PF, Quirie, M, Maley, MA, Foster, G, Toszeghy, M, Newell, DG. 2005. A multiplex polymerase chain reaction to detect and differentiate *Campylobacter fetus* subspecies *fetus* and *Campylobacter fetus* species *venerealis*: use on UK isolates of *C. fetus* and other *Campylobacter spp*. J. Appl. Microbiol. 99:758-66.

38. Spence, RP, Bruce, IR, McFadden, AM, Hill, FI, Tisdall, D, Humphrey, S, van der Graaf, L, van Bergen, MA, Wagenaar, JA. 2011. Cross-reaction of a *Campylobacter fetus* subspecies *venerealis* real-time PCR. Vet. Rec. 168:131.

Chapter 3

First closed genome sequence of *Campylobacter fetus* su sp *venerealis* iovar intermedius

Linda van der Graaf – van Bloois[a,b], William G. Miller[c], Emma Yee[c], James L. Bono[d], Martine Rijnsburger[e], Carlos Campero[f], Jaap A. Wagenaar[a,b,g], Birgitta Duim[a,b]

[a] Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands
[b] WHO Collaborating Center for Campylobacter / OIE Reference Laboratory for Campylobacteriosis
[c] Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA, USA
[d] Meat Safety and Quality Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Clay Center, NE, USA
[e] VU University Medical Center, Dept. of Medical Microbiology and Infection Control, Amsterdam, the Netherlands
[f] Patología Veterinaria, Instituto Nacional de Tecnología Agropecuaria (INTA), Balcarce, Argentina
[g] Central Veterinary Institute of Wageningen UR, Lelystad, The Netherlands

Genome Announcements. 2014, 2:10.1128/genomeA.01246-13

A stract

Campylobacter fetus subsp. *venerealis* biovar intermedius is a variant of *Campylobacter fetus* subsp. *venerealis*, the causative agent of Bovine Genital Campylobacteriosis, a venereal disease associated with abortion and infertility in cattle. We report the first closed whole-genome sequence of this biovar.

Genome announcement

The pathogen *Campylobacter fetus* contains two subspecies: *Campylobacter fetus* subsp. *fetus* and *Campylobacter fetus* subsp. *venerealis*. The niches of the subspecies are different: *Campylobacter fetus fetus* can be isolated from a variety of different hosts (1), whereas *Campylobacter fetus venerealis* is restricted to the genital tract of cattle (2). Both subspecies can cause disease in cattle, but only *Campylobacter fetus venerealis* is described as the causative agent of Bovine Genital Campylobacteriosis, a disease characterized by fertility problems in cattle (2). A variant of *Campylobacter fetus venerealis*, designated *Campylobacter fetus venerealis* biovar intermedius has been identified (3). A *Campylobacter fetus venerealis* biovar intermedius genome sequence is available (4), consisting of 218 unassembled contigs. This unassembled genome does not allow the full identification of the core and accessory genome and the reconstruction of pathways and surface structures that might contribute to the pathogenicity of this biovar. In this study, we report the first closed whole-genome sequence of *Campylobacter fetus venerealis* biovar intermedius strain 03/293, isolated in Argentina from the lung of an aborted bovine foetus (5).

For sequencing, three platforms were combined: Roche GS-FLX Titanium, Illumina MiSeq and PacBio RS. A total of 248,123 Roche 454 reads were assembled using the Newbler assembler (v2.6) into four scaffolds with 82 contigs, with a 50x coverage. All 454 base calls were validated using 1,490,018 Illumina MiSeq reads, which added 156x coverage. A circular, high resolution AflII restriction map of the genome was generated by optical mapping (Argus Optical Mapper, OpGen Inc, Gaithersburg, MD) to determine the orientation and order of the contigs and validate the assembly. Assembly of regions with insertion sequences or repeats in the S-layer and PAI locus were confirmed with PacBio Continuous Long Reads, adding 117x coverage. All base calls and polymeric tracts were validated using the high-depth MiSeq reads.

Campylobacter fetus venerealis biovar intermedius strain 03/293 has a circular genome of 1,866,009 bp with an average GC content of 33%. Protein-, rRNA- and tRNA-encoding genes were identified as described (6). Annotation was based on *Campylobacter fetus* strain 82-40 (accession number NC_008599) and 03-427^{T} (accession number CP006833), using Artemis (7) and the identification of Pfam domains (v27.0) (8). The genome encodes 1773 putative protein-coding genes, 48 pseudogenes, 43 tRNA genes and 3 rRNA operons. The genome contains 13 insertion sequence elements and 31 homopolymeric GC tracts ($\geq$ 8bp). This strain is a carrier of two megaplasmids (91,400 bp and 35,326 bp) and a small cryptic plasmid (3,993 bp).

The biovar intermedius genome contains a S-layer coding region, which is also found in *Campylobacter fetus fetus* and *Campylobacter fetus venerealis*. The genomes of *Campylobacter fetus* subspecies were compared with BLASTP analysis. A high degree of both synteny and similarity in gene content between the *Campylobacter fetus* subspecies was shown. The closed genome sequence of *Campylobacter fetus venerealis* biovar intermedius provides an important reference genome for the comparison of virulence and host association of the different *Campylobacter fetus* subspecies.

Nucleotide sequence accession num ers

The complete genome sequence of *Campylobacter fetus venerealis* biovar intermedius strain 03/293 has been deposited in GenBank under accession numbers CP006999-CP007002.

Acknowledgements

We thank Mary Chapman and Nathaniel Simon for the generation of Illumina MiSeq reads and Sandy Fryda Bradley, Bob Lee and Renee Godtel for the generation of PacBio RS reads. This work was supported in part by USDA-ARS CRIS project 5325-42000-047-00D.

References

1. Garcia, MM, Eaglesome, MD, Rigby, C. 1983. Campylobacters important in veterinary medicine. Veterinary Bulletin. 53:793-818.

2. Dekeyser, J. 1984. *Campylobacter* infections in man and animals. J.P. Butzler (Ed.) CRC Press Inc., Boca Raton, Florida.

3. Véron, M., Chatelain,R. 1973. Taxonomy study of the genus *Campylobacter* Sebald and Verón and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Verón. International Journal of Systematic Bacteriology. 23:122-134.

4. Iraola, G, Perez, R, Naya, H, Paolicchi, F, Harris, D, Lawley, TD, Rego, N, Hernandez, M, Calleros, L, Carretto, L, Velilla, A, Morsella, C, Mendez, A, Gioffre, A. 2013. Complete Genome Sequence of *Campylobacter fetus* subsp. *venerealis* Biovar Intermedius, Isolated from the Prepuce of a Bull. Genome Announcements. 1: 00526-13.

5. Campero, CM, Moore, DP, Odeon, AC, Cipolla, AL, Odriozola, E. 2003. Aetiology of bovine abortion in Argentina. Vet. Res. Commun. 27:359-369.

6. Merga, JY, Winstanley, C, Williams, NJ, Yee, E, Miller, WG. 2013. Complete Genome Sequence of the *Arcobacter butzleri* Cattle Isolate 7h1h. Genome Announcements. 1:00655-13.

7. Rutherford, K, Parkhill, J, Crook, J, Horsnell, T, Rice, P, Rajandream, MA, Barrell, B. 2000. Artemis: sequence visualization and annotation. Bioinformatics. 16:944-945.

8. Punta, M, Coggill, PC, Eberhardt, RY, Mistry, J, Tate, J, Boursnell, C, Pang, N, Forslund, K, Ceric, G, Clements, J, Heger, A, Holm, L, Sonnhammer, EL, Eddy, SR, Bateman, A, Finn, RD. 2012. The Pfam protein families database. Nucleic Acids Res. 40:D290-301.

Chapter 4

Inconsistency of phenotypic and genomic characteristics of *Campylobacter fetus* su species requires reevaluation of current diagnostics

Linda van der Graaf – van Bloois[a,b], William G. Miller[c], Emma Yee[c], Martine Rijnsburger[d], Jaap A. Wagenaar[a,b,e], Birgitta Duim[a,b]

[a] Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands
[b] WHO Collaborating Center for Campylobacter / OIE Reference Laboratory for Campylobacteriosis
[c] Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA, USA
[d] VU University Medical Center, Dept. of Medical Microbiology and Infection Control, Amsterdam, the Netherlands
[e] Central Veterinary Institute of Wageningen UR, Lelystad, The Netherlands

Journal of Clinical Microbiology. 2014, 52(12):4183-4188

A stract

Classification of the *Campylobacter fetus* subspecies *fetus* and *venerealis* was first described in 1959 and was based on the source of isolation (intestinal versus genital) and the ability of the strains to proliferate in the genital tract of cows. Two phenotypic assays (1% glycine tolerance and H_2S production) were described to differentiate the subspecies. Multiple molecular assays have been applied to differentiate the *C. fetus* subspecies, but none of these tests are consistent with the phenotypic identification. In this study, we defined the core genome and accessory genes of *C. fetus*, based on the closed genomes of five *C. fetus* strains. Phylogenetic analysis using the core genomes of 23 *C. fetus* strains of both subspecies showed a division into two clusters. The phylogenetic core genome clusters were not consistent with the phenotypic classification of the *C. fetus* subspecies. However, they were consistent with the molecular characteristics of the strains, determined by multilocus sequence typing, sap-typing, and the presence/absence of insertion sequences and a type I restriction-modification system. The fact that three of the phenotypically-defined *C. fetus* subsp. *fetus* strains have the same genomic characteristics as *C. fetus* subsp. *venerealis* strains, when considering the core genome and accessory genes, requires a critical evaluation of the clinical relevance of *C. fetus* subspecies identification by phenotypic assays.

Introduction

Campylobacter fetus is an important veterinary pathogen, associated with genital infections in cattle and sheep, resulting in abortion and infertility (1). Based on clinical and phenotypic observations, the species *C. fetus* was subdivided into two subspecies: *C. fetus* subsp. *venerealis* (Cfv) and *C. fetus* subsp. *fetus* (Cff) (2). Cfv was of venereal origin, had a high ability to cause abortions and could persist in the genital tracts, whereas Cff was of intestinal origin, caused only sporadic abortions and was cleared from the genital tract of the cow following the abortion. The two subspecies were phenotypically differentiated with the 1% glycine tolerance and H_2S production tests: Cff is positive in both tests (glycine tolerant and H_2S production) and Cfv is negative in both tests (no growth in presence of 1% glycine and no H_2S production) (2). *C. fetus* strains were isolated that were able to establish themselves in the genital tract of a non-pregnant cow, like the venereal Cfv strains (3); however, the glycine resistance of these strains was reduced, compared to most Cff strains, and they were positive in the H_2S test, like the intestinal Cff strains (4). They were classified in an intermediate group (4) and designated *Campylobacter fetus* subsp. *venerealis* biovar intermedius (1).

For molecular subspecies identification, several PCR-assays have been described, but these lack specificity (5). The subspecies can be genetically differentiated with multilocus sequence typing (MLST) (6) and Amplified Fragment Length Polymorphism (AFLP) analyses (7), but these methods are laborious and therefore not useful for routine diagnostic methods. The rationale behind the differentiation between the *C. fetus* subspecies is the supposed difference in pathogenicity and disease epidemiology. Cfv is described as the causative agent of Bovine Genital Campylobacteriosis (BGC), a statutory disease in many countries of the world and listed by the OIE (World Organisation for Animal Health), in contrast with Cff, which is associated with sporadic abortions (8).

Comparative genomics of two *C. fetus* strains revealed several unique regions for both subspecies, as shown by Kienesberger et al. (9): *C. fetus* subsp. *venerealis* contained multiple unique regions representing insertion sequences and genomic islands with Type IV secretion system components and phage-related/hypothetical proteins (9); and *C. fetus* subsp. *fetus* contained CRISPR-cas loci and unique genes involved in LPS biosynthesis (9). These data suggested that *C. fetus* subspecies can be distinguished on genomic features, but comparative genomics of a larger set of *C. fetus* strains is lacking. The aim of this study is to characterize *C. fetus* strains of both subspecies based on whole genome sequencing

data and to compare the results of classification, based on core and accessory genome analysis, with the current *C. fetus* subspecies identification based on phenotypic assays.

Materials and methods

Bacterial strains; phenotyping and genotyping

C. fetus strains (listed in Table 1) were grown on heart-infusion agar supplemented with 5% sheep blood (Biotrading, Mijdrecht, the Netherlands) for two days under microaerobic conditions (6% O_2, 7% CO_2, 7% H_2, 80% N_2, (Anoxomat, Mart Microbiology, Lichtenvoorde, the Netherlands)). The subspecies of the strains were phenotypically identified with the 1% glycine tolerance test and hydrogen sulphide (H_2S) production in medium with 0.02% cysteine-HCl test, as described (1). Molecular identification was performed using MLST (6) and AFLP analysis (7).

Whole genome sequencing

Whole genome sequence data of 21 *C. fetus* strains was obtained using Roche GS-FLX Titanium sequencing. Roche 454 reads were assembled into contigs using the Newbler assembler (v2.6). The genomes of two *C. fetus* strains, 04/554 and 97/608 were closed through assembly of the Roche 454 contigs into scaffolds by using Perl scripts. To validate the assembly of the contigs and to determine the orientation and order of the scaffolds, a circular, high resolution *Afl*II restriction map of the genome was generated by optical mapping (Argus Optical Mapper, OpGen Inc, Gaithersburg, MD. Assembly of the *sap* locus, genomic islands, regions with insertion sequences and repeats were confirmed with PacBio Continuous Long Reads (Keygene N.V., Wageningen, the Netherlands). All base calls and polymeric tracts were validated using the high-depth Illumina MiSeq reads. The genomes of strain 04/554 and 97/608 were annotated as described (10).

Phylogenetic analysis of core and accessory genomes

Three available closed *C. fetus* genomes were used as a reference: strain 82-40 (GenBank accession number CP000487), strain 84-112 (GenBank accession numbers HG004426-HG004427) and strain 03/293 (GenBank accession numbers CP006999-CP007002). The amino acid sequences of the open reading frames (ORFs) encoded by five genomes (the three reference genomes plus two genomes sequenced in this study (04/554 and 97/608)) were used as input for an all-versus-all sequence similarity search using BLASTp (-e 0.0001, >80% similarity cutoff). ORFs that exist in each of the five strains (>80% identity

a le General characteristics of *C. fetus* strains

			Phenotype[b]			*Genotype*[c]			*Accessory genes*[d,e]					
Strain	**Country**[a]	**Source**	**1% Glycine tolerance**	**H_2S production**	**Phenotypic ID**	**AFLP**	**MLST (ST)**	**Core genome cluster**	***sap* type**	**Prophage in *sap* locus**	**GI with T4SS**	**IS**	**Type I RM system**	**CRISPR-cas**
82-40	US	Human	+	+	Cff	Cff	6	**B**	A	-	-	-	+/-	+
110800-21-2	NL	Bovine (bull)	+	+	Cff	Cff	2	**B**	A	-	+	-	+/-	-
BT 10/98	UK	Ovine	+	+	Cff	Cff	2	**B**	A	-	-	-	+/-	-
04/554	AR	Bovine (foetus)	+	+	Cff	Cff	5	**B**	B	-	-	-	-	-
98/v445	UK	Bovine (bull)	+	+	Cff	Cff	3	**B**	B	-	+	-	-	+
03/293	AR	Bovine (foetus)	+	+	Cff	Cfvi	4	**A**	A	-	+	+	+	-
ADRI 1362	AR	Bovine	+	+	Cff	Cfvi	4	**A**	A	+	+	+	+	-
Zaf 65	SA	Bovine	+	+	Cff	Cfvi	4	**A**	A	-	-	+	+	-
01/165	AR	Bovine (mucus)	-	+	Cfvi	Cfvi	4	**A**	A	-	+	+	+	-
02/298	AR	Bovine (foetus)	-	+	Cfvi	Cfvi	4	**A**	A	-	+	+	+	-
03/596	AR	Bovine (foetus)	-	+	Cfvi	Cfvi	4	**A**	A	-	+	+	+	-
92/203	AR	Bovine (placenta)	-	+	Cfvi	Cfvi	4	**A**	A	-	+	+	+	-
97/532	AR	Bovine (mucus)	-	+	Cfvi	Cfvi	4	**A**	A	-	+	+	+	-
98/25	AR	Bovine (foetus)	-	+	Cfvi	Cfvi	4	**A**	A	-	+	+	+	-
WBT 011/09	UK	unknown	-	+	Cfvi	Cfvi	4	**A**	A	-	+	+	+	-
Zaf 3	SA	Bovine (foetus)	-	+	Cfvi	Cfvi	4	**A**	A	-	-	+	+	-
ADRI 513	AU	unknown	-	+	Cfvi	Cfv	4	**A**	A	-	-	+	+	-
CCUG 33872	CZ	unknown	-	+	Cfvi	Cfv	4	**A**	A	+	-	+	+	-
84-112	US	Bovine	-	-	Cfv	Cfv	4	**A**	A	+	+	+	+	-
97/608	AR	Bovine (placenta)	-	-	Cfv	Cfv	4	**A**	A	+	+	+	+	-
B10	US	Bovine	-	-	Cfv	Cfv	4	**A**	A	+	+	+	+	-
CCUG 33900	F	Bovine (abortion)	-	-	Cfv	Cfv	4	**A**	A	+	+	+	+	-
LMG 6570	BE	Bovine	-	-	Cfv	Cfv	4	**A**	A	+	+	+	+	-

Abbreviations: MLST (ST); Multi Locus Sequence Typing (Sequence Type)
GI with T4SS; Genomic Island with type IV secretion system
IS; Insertion Sequence
RM system; Restriction-Modification system
CRISPR; Clustered regularly interspaced short palindromic repeats

Country code: AR Argentina, AU Australia, BE Belgium, CZ Czech Republic, F France, NL Netherlands, SA South Africa, UK United Kingdom, US United States

Phenotype: + positive in assay - negative in assay

Accessory: + genes are present - genes are absent

Taxon: Cff *C. fetus fetus* Cfv *C. fetus venerealis* Cfvi *C. fetus venerealis* biovar intermedius

over at least 80% of the protein length) were considered to be part of the *C. fetus* core genome. The ORF sequences of strain 82-40 were used as reference sequences of the core genes. Regions encoding the *sap* locus, genomic islands, restriction-modification systems, prophages and insertion sequences were considered as accessory genes.

The accessory genes in the Roche 454 contigs of 21 *C. fetus* strain were identified with a local BLASTn analysis (-e 0.0001, >80% similarity cutoff) against the identified accessory genes of the five closed *C. fetus* strains. The strains were considered positive for the specific accessory regions if the BLASTn match was >80% over at least 80% of the region.

The phylogenetic analysis of the core genomes was performed as follows: nucleotide sequences of the predicted genes of the Roche 454 contigs were generated using GeneMark (v2.8) (11). For each core gene, the corresponding nucleotide sequence of each strain was extracted and aligned on a gene-by-gene basis using MUSCLE (12). The alignments were concatenated into a contiguous sequence for each *C. fetus* strain. From this concatenated alignment, a phylogenetic maximum likelihood tree was built using RAxML (v7.2.8) under the GTRCAT model.

Results

Genome Features

The genome of *C. fetus* subsp. *fetus* 04/554 is a circular chromosome of 1,800,764 bp with an average G+C content of 33.2%, and one megaplasmid of 25,862 bp. The genome of *C. fetus* subsp. *venerealis* 97/608 has a circular chromosome of 1,935,028 bp with an average G+C content of 33.3% and contains two megaplasmids of 38,272 bp and 27,124 bp. The general features of the assembled genomes are shown in Table 2. The genome

a le 2 Features of assembled *C. fetus* genomes

	Data for strain (reference or source):				
Feature	**04/554 (this study)**	**97/608 (this study)**	**03/293 (10)**	**82-40 (9)[a]**	**84-112 (9)[a]**
Genome size (bp)	1,800,764	1,935,028	1,866,009	1,773,615	1,926,886
G+C content (%)	33.2	33.3	33.3	33.3	33.3
No. of rRNA genes	3	3	3	3	3
No. of tRNA genes	43	43	43	43	43
No. of homopolymeric G+C tracts (> 8bp)	29	24	31	30	34
No. of open reading frames (no. of pseudogenes)	1,684 (68)	1,879 (60)	1,773 (48)	1769	1992
Plasmids	1	2	3	0	1
Size (bp)	25,862	38,272 / 27,124	91,400 / 35,326 / 3,993		61,141
G+C content (%)	29.0	31.3 / 28.1	29.4 / 33.0 / 31.4		31.5
Sap locus type	B	A	A	A	A
Insertion elements (no. of copies)	0	14	13	0	5
Restriction or modification locus type	none	I	I	none	I
CRISPR-*cas* system	no	no	no	yes	no

[a] With modifications from original publication

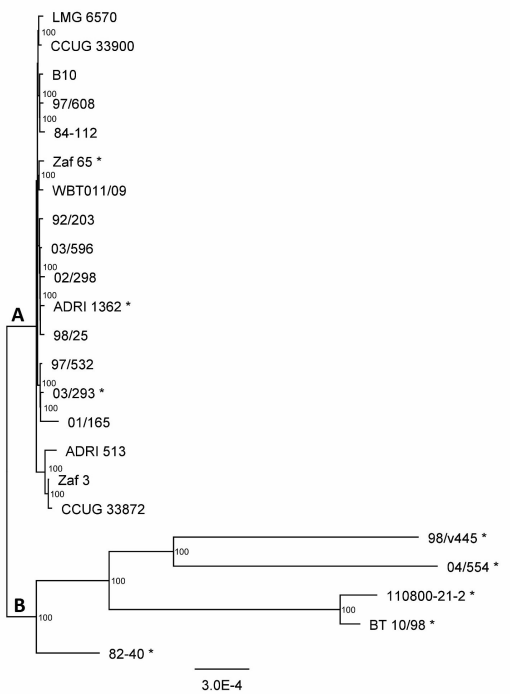

Figure . Phylogenetic tree of *C. fetus* strains based on the core genomes. Strains that are phenotypically identified as Cff are marked with an asterisk (*). Bootstrap supports are indicated on the branches. The scale represents the mean number of nucleotide substitutions per site.

features of strains 03/293, 82-40 and 84-112 have been described previously (9, 10) and are summarized in Table 2.

Nucleotide sequence accession numbers

In GenBank, the genome sequence of *C. fetus* subsp. *fetus* 04/554 has been deposited under accession numbers CP008808-CP008809 and *C. fetus* subsp. *venerealis* 97/608 under accession numbers CP008810-CP008812.

Phylogeny of the C. fetus core genome

Comparison of five closed *C. fetus* genomes (82-40, 84-112, 04/554, 97/608 and 03/293) revealed highly syntenic genomes, sharing >90% sequence identity. The core genome of *C. fetus* was defined on the ORFs present in the closed genomes of five *C. fetus* strains, and was determined to consist of 1,409,454 bp and 1509 ORFs. The core genome was then identified in an additional 18 *C. fetus* strains and the genetic distance of all 23 core

genomes was visualized with a phylogenetic maximum likelihood tree (Figure 1). The phylogenetic tree is arranged in two clusters of strains, designated Cluster A and B. The majority (n=18) of strains were located in Cluster A, whereas Cluster B consisted of five strains with a higher genetic distance.

Accessory genes of C. fetus strains

Major differences between the five closed *C. fetus* genomes were found in the accessory genes. Table 1 shows the identified genes belonging to the *sap* loci, insertion sequences, genomic islands, type I restriction-modification systems, prophages and CRISPR-*cas* systems of the 23 analyzed *C. fetus* strains.

***Sap* locus** All analyzed *C. fetus* strains contained a *sap* locus, but the composition of this region differed between strains. *Sap*-type B strains belonged to the core genome Cluster B, wheareas *sap*-type A strains were found in both genome Clusters (Table 1). The *sap* locus of strain 84-112 and strain 97/608 contained transposable elements and a set of phage-related genes and hypothetical genes, indicating the presence of a prophage. These prophage sequences shared 100% identity, but were inserted at different positions in the *sap*-locus.

Genomic islands Two genomic islands (GI) encoding a Type IV secretion system (T4SS), as defined by Kienesberger et al. (9, 13), were identified in the chromosomes of the closed *C. fetus* genomes. These two chromosomally-located T4SS regions were present in strains that were distributed over the two core genome clusters (Table 1).

Insertion Sequences. Insertion sequences (IS) were found in the chromosomes as well as in plasmids of the closed genomes of strains 84-112, 97/608 and 03/293. The identified IS belonged to the IS*605*, IS*607* and IS*200* families. The IS are only found in *C. fetus* strains belonging to core genome Cluster A (Table 2). Each IS positive strain contained all of the identified IS families.

Restriction-modification system Three of the closed *C. fetus* genomes, 84-112, 97/608 and 03/293 contained a type I restriction-modification (RM) system. This type I RM system consists of *hsd* genes with intervening ORFs, similar to the described type I RM systems in *C. jejuni* (14). The complete type I RM system is only found in *C. fetus* strains belonging to core genome Cluster A (Table 2). The three other *C. fetus* strains of *sap*-type A, 82-40, 110800-21-2 and BT 10/98, contained a remnant of this type as the type I RM system of these strains lacked the *hsdS2* gene. The genomes of the *sap*-type B strains did not contain any type I RM-encoding genes.

CRISPR-*cas* system CRISPR repeats were present in all *C. fetus* strains, but only two *C.*

fetus subsp. *fetus* strains, 82-40 and 98/v445, contained *cas* genes. These strains were not linked with the same core genome cluster.

Core genome clusters compared with accessory genes

The presence of specific components of the accessory genes encoding prophages, genomic islands and CRISPR-*cas* system were not associated with a specific core genome cluster. The IS elements and complete type I RM system were exclusively found in the strains of Cluster A. Strains of Cluster B did not contain IS elements or a complete type I RM system and have different *sap* types.

Core genome clusters compared to subspecies identification

The subspecies were not consistently phenotypically- and genotypically-subdivided (Table 1). The genotypic method MLST was consistent with the obtained core genome clustering: strains of Cluster A are all MLST ST-4, whereas Cluster B consisted of strains with other MLST STs. Cluster B included two strains with a similar MLST ST-2, and these strains were determined to have a lower genetic distance than the other strains in Cluster B. AFLP is able to distinguish the strains within Cluster A with a minor difference in fingerprint as Cfv and Cfvi(6), but this discrimination is not observed with the phylogenetic analysis of the core genomes. Strains that were phenotypically classified as Cfv and Cfvi belonged to core genome Cluster A, but the eight phenotypically-classified Cff strains were dispersed among both genome clusters. This is represented in Figure 1, where all phenotypically identified Cff strains are marked with an asterisk.

Discussion

The original classification of the *C. fetus* subspecies is based on differences in the colonization of different niches and phenotypic characteristics (2-4). The two *C. fetus* subspecies are highly syntenic, sharing 92.9% sequence identity (9), and the subspecies cannot be distinguished by DNA-DNA hybridization (15), which questions the validity of subspecies differentiation and hampers an adequate taxonomic positioning of the subspecies. Furthermore, the reliability of the 1% glycine tolerance test can be influenced by the fact that glycine tolerance can be transduced by phages (16). Several molecular assays for the identification of the *C. fetus* subspecies have been published (5),: however, none of the molecular assays corresponded fully to the phenotypic identification of the *C. fetus* subspecies (5, 6).

In this study, phylogenetic analysis of the core genomes subdivided the *C. fetus* strains in

two clusters. All strains that are phenotypically identified as Cfv (including biovar intermedius) clustered in one core genome cluster, containing only strains with MLST ST-4 and harboring IS elements and a type I RM system. The strains phenotypically identified as Cff were assigned to both clusters. Three Cff strains, 03/293, Zaf 65 and ADRI 1362 were assigned to the core genome cluster with Cfv and Cfvi strains, despite their phenotypic identification of Cff. The similarity of the MLST identification and core genome clusters can be explained by the fact that MLST is a small-scale reflection of the core genome. The phylogenetic analysis showed an obvious resemblance with the MLST STs of the strains; the genetic distances between strains with the same STs are very low as shown for strains of ST4 and ST2, and the genetic distances increased for strains with different STs.

Campylobacter fetus venerealis biovar intermedius is described as a phenotypic variant of *C. fetus* subsp. *venerealis* (1). The phenotypically-identified Cfvi strains are all positioned together with Cfv strains in cluster A of the phylogenetic tree. The accessory genes of Cfv and Cfvi strains showed no consistent presence of a Cfv or Cfvi specific region. However, it is remarkable that, except strain ADRI 513, all of the with AFLP identified Cfv strains contain a prophage in the *sap*-locus and that this prophage is absent in the majority of Cfvi strains. Almost all proteins of this prophage are hypothetical with unknown function, but one may speculate that the presence of this prophage influences the phenotypic difference, such as the H_2S production between the Cfv and Cfvi strains.

The differentiation between *C. fetus* subspecies goes beyond only taxonomic interest. Clinically, the subspecies have been described to be different. Cfv (including biovar intermedius) is described as the causative agent of Bovine Genital Campylobacteriosis (BGC). There is a generally accepted association between the *C. fetus* subspecies and their specific clinical features, epidemiological characteristics and host niche specificity. Bovine products for trade must be checked for the absence of Cfv as stated in the Terrestrial Animal Health Code by the World Organisation of Animal Health (OIE) (17). When in such a screening *C. fetus* is detected, subspecies identification is generally done by phenotypic assays as described in the OIE Manual of Diagnostic Tests and Vaccines for Terrestrial Animals (18). However, there is no evidence that the phenotypic markers (glycine tolerance and H_2S production) are linked to virulence characteristics of *C. fetus* subspecies. Future diagnostics of *C. fetus* should preferably detect genomic characteristics associated with virulence and different host niches. The virulence genes of genomic islands present in *C. fetus* strains were described as Cfv-specific and proposed as targets for diagnostic assays (19, 20). However these genes are not consistently present in Cfv

genomes (13, 21) and therefore not useful as a diagnostic assay. Pending the identification of virulence-associated genes, one should be aware that the current association between phenotype and virulence is questionable, since several phenotypically-defined Cff strains have the same genomic characteristics as Cfv strains, on the basis of core genome and accessory gene similarity, as shown in this study. The inconsistency of phenotypes and genomic characteristics of *C. fetus* strains encourages a critical evaluation of the clinical relevance of *C. fetus* subspecies identification by phenotypic assays.

Acknowledgements

We like to thank Dr. Brian Brooks and Dr. John Devenish (Canadian Food Inspection Agency) for providing strains. We thank Mary Chapman and Nathaniel Simon for the generation of Illumina MiSeq reads and thank James Bono for the generation of PacBio RS reads.

References

1. Véron, M, Chatelain, R. 1973. Taxonomy study of the genus *Campylobacter* Sebald and Verón and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Verón. International Journal of Systematic Bacteriology. 23:122-134.

2. Florent, A. 1959. Les deux vibriosis génitales; la vibriose due à *V. fetus venerealis* et la vibriose d'origine intestinale due à *V. fetus intestinalis*. Mededelingen der Veeartsenijschool van de RijksUniversiteit te Gent. 1-60.

3. Park, RWA, Munro, IB, Melrose, DR, Stewart, DL. 1962. Observations on the ability of two biochemical types of *Vibrio fetus* to proliferate in the genital tract of cattle and their importance with respect to infertility. Br. Vet. J. 118:411.

4. Florent, A. 1963. A propos dese vibrions responsables de la vibriose génitale des bovins et des ovins. Bull. Off. Int. Epizoot. 60:1063-1074.

5. van der Graaf-van Bloois, L, van Bergen, MA, van der Wal, FJ, de Boer, AG, Duim, B, Schmidt, T, Wagenaar, JA. 2013. Evaluation of molecular assays for identification *Campylobacter fetus* species and subspecies and development of a *C. fetus* specific real-time PCR assay. J. Microbiol. Methods. 95:93-97.

6. van Bergen, MAP, Dingle, KE, Maiden, MC, Newell, DG, van der Graaf-Van Bloois, L, van Putten, JP, Wagenaar, JA. 2005. Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. J. Clin. Microbiol. 43:5888-98.

7. Wagenaar, JA, van Bergen, MAP, Newell, DG, Grogono-Thomas, R, Duim, B. 2001. Comparative study using amplified fragment length polymorphism fingerprinting, PCR genotyping, and phenotyping to differentiate *Campylobacter fetus* strains isolated from animals. J. Clin. Microbiol. 39:2283-6.

8. Garcia, MM, Eaglesome, MD, Rigby, C. 1983. Campylobacters important in veterinary medicine. Veterinary bulletin. 53:793-818.

9. Kienesberger, S, Sprenger, H, Wolfgruber, S, Halwachs, B, Thallinger, GG, Perez-Perez, GI, Blaser, MJ, Zechner, EL, Gorkiewicz, G. 2014. Comparative Genome Analysis of *Campylobacter fetus* Subspecies Revealed Horizontally Acquired Genetic Elements Important for Virulence and Niche Specificity. PLoS One. 9:e85491.

10. van der Graaf-van Bloois, L, Miller, WG, Yee, E, Bono, JL, Rijnsburger, M, Campero, C, Wagenaar, JA, Duim, B. 2014. First Closed Genome Sequence of *Campylobacter fetus* subsp. *venerealis* bv. intermedius. Genome Announc. 2:10.1128.

11. Lukashin, AV, Borodovsky, M. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. 26:1107-1115.

12. Edgar, RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 5:113.

13. Gorkiewicz, G, Kienesberger, S, Schober, C, Scheicher, SR, Gully, C, Zechner, R, Zechner, EL. 2010. A genomic island defines subspecies-specific virulence features of the host-adapted pathogen *Campylobacter fetus* subsp. *venerealis*. J. Bacteriol. 192:502-17.

14. Miller, WG, Pearson, BM, Wells, JM, Parker, CT, Kapitonov, VV, Mandrell, RE. 2005. Diversity within the *Campylobacter jejuni* type I restriction-modification loci. Microbiology. 151:337-351.

15. Roop, RM, Smibert, RM, Johnson, JL, Krieg, NR. 1984. Differential characteristics of catalase-positive Campylobacters correlated with DNA homology groups. Can. J. Microbiol. 30:938-51.

16. Chang, W, Ogg, JE. 1971. Transduction and mutation to glycine tolerance in *Vibrio fetus*. Am. J. Vet. Res. 32:649-653.

17. OIE. 2013. Terrestrial Animal Health Code. 2013.

18. OIE. 2012. Bovine Genital Campylobacteriosis, p. 652. *In* Anonymous Manual of Diagnostic Tests and Vaccines for Terrestrial Animals (mammal, birds and bees), 7th ed. Office International des Epizooties, Paris.

19. Iraola, G, Hernandez, M, Calleros, L, Paolicchi, F, Silveyra, S, Velilla, A, Carretto, L, Rodriguez, E, Perez, R. 2012. Application of a multiplex PCR assay for *Campylobacter fetus* detection and subspecies differentiation in

uncultured samples of aborted bovine fetuses. J. Vet. Sci. 13:371-376.

20. Moolhuijzen, PM, Lew-Tabor, AE, Wlodek, BM, Aguero, FG, Comerci, DJ, Ugalde, RA, Sanchez, DO, Appels, R, Bellgard, M. 2009. Genomic analysis of *Campylobacter fetus* subspecies: identification of candidate virulence determinants and diagnostic assay targets. BMC Microbiol. 9:86.

21. Abril, C, Brodard, I, Perreten, V. 2010. Two novel antibiotic resistance genes, tet(44) and ant(6)-Ib, are located within a transferable pathogenicity island in *Campylobacter fetus* subsp. *fetus*. Antimicrob. Agents Chemother. 54:3052-5.

Chapter 5

Whole genome sequence analysis indicates recent diversification of mammal-associated *Campylobacter fetus* and implicates a genetic factor associated with H_2S production

Linda van der Graaf – van Bloois[a,b], Birgitta Duim[a,b], William G. Miller[c], Ken J. Forbes[d], Jaap A. Wagenaar[a,b,c], Aldert L. Zomer[a,b]

[a] Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands
[b] WHO Collaborating Center for Campylobacter / OIE Reference Laboratory for Campylobacteriosis
[c] Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA, USA
[d] School of Medicine and Dentistry, University of Aberdeen, Aberdeen, United Kingdom
[e] Central Veterinary Institute of Wageningen UR, Lelystad, The Netherlands

BMC Genomics. 2016, 17:713

A stract

Background: *Campylobacter fetus* (*C. fetus*) can cause disease in both humans and animals. *C. fetus* has been divided into three subspecies: *C. fetus* subsp. *fetus* (Cff), *C. fetus* subsp. *venerealis* (Cfv) and *C. fetus* subsp. *testudinum* (Cft). Subspecies identification of mammal-associated *C. fetus* strains is crucial in the control of Bovine Genital Campylobacteriosis (BGC), a syndrome associated with Cfv. The prescribed methods for subspecies identification of the Cff and Cfv isolates are: tolerance to 1% glycine and H_2S production.

Results: In this study, we observed the deletion of a putative cysteine transporter in the Cfv strains, which are not able to produce H_2S from L-cysteine. Phylogenetic reconstruction of the core genome single nucleotide polymorphisms (SNPs) within Cff and Cfv strains divided these strains into five different clades and showed that the Cfv clade and a Cff clade evolved from a single Cff ancestor.

Conclusions: Multiple *C. fetus* clades were observed, which were not consistent with the biochemical differentiation of the strains. This suggests the need for a closer evaluation of the current *C. fetus* subspecies differentiation, considering that the phenotypic differentiation is still applied in BGC control programs.

Background

The pathogen *Campylobacter fetus* (*C. fetus*) can cause disease in both animals and humans. In humans, *C. fetus* infections vary from acute diarrhea to systemic illness [1]. In animals, *C. fetus* infections can cause abortion and infertility, mainly in cattle and in sheep [2]. The mammal-associated *C. fetus* subspecies are *C. fetus* subsp. *fetus* (Cff) and *C. fetus* subsp. *venerealis* (Cfv) [3], whereas *C. fetus* subsp. *testudinum* (Cft) is associated with reptiles [4]. *C. fetus* subsp. *venerealis* includes a biochemical variant, designated *C. fetus* subsp. *venerealis* biovar intermedius (Cfvi) [5].

C. fetus subsp. *venerealis* has been described to be the causative agent of Bovine Genital Campylobacteriosis (BGC), associated with infertility and abortion in cattle [6]. BGC is notifiable to the World Organisation for Animal Health (OIE). A crucial element in the BGC control program relies on the subspecies identification of *C. fetus* isolates. Currently, the methods prescribed by the OIE to differentiate Cff, Cfv and Cfvi are tolerance to 1% glycine and H_2S production [7]: Cff is tolerant to 1% glycine and able to produce H_2S, Cfv is not tolerant to 1% glycine and not able to produce H_2S and Cfvi is not tolerant to 1% glycine (like Cfv) and able to produce H_2S (like Cff). The biochemical tests are hampered by poor reproducibility [8], and the phenotypes are not completely consistent with the genomic characteristics of the *C. fetus* strains, since phenotypically-identified Cff strains were genotypically identical with Cfv strains [9]. An obvious distinguishing and important feature of *C. fetus* cells is the surface layer (S-layer), which is considered to be associated with the pathogenicity of *C. fetus* strains [10]. *C. fetus* cells can express two types of surface array proteins (Sap), which correlates with the serotypes of the bacterium; Cfv strains are serotype A; Cff strains may be either serotype A or serotype B and rarely serotype AB; and Cft strains are serotype A, serotype B or serotype AB [11, 12]. The molecular method multi-locus sequence typing (MLST) was recommended to differentiate Cff and Cfv strains [8]. However, a recent study showed that the current MLST scheme was not able to reliably differentiate the *C. fetus* subspecies, as a Cff strain was isolated with the Cfv-associated MLST ST-4 genotype [13].

Whole-genome analysis provides fine-scale resolution of bacterial genomes and allows the calculation of evolutionary events, as shown for *C. coli* and *C. jejuni* [14]. Whole-genome analysis has improved, and will continue to improve, our understanding of the features that distinguish *C. fetus* subspecies and the evolutionary forces that have acted on *C. fetus* over time. In this study, we performed a core genome single-nucleotide polymorphisms (SNPs) analysis of 42 *C. fetus* Cff and Cfv genomes to identify subspecies-specific SNPs. We performed a SNP-based phylogenetic analysis of the core genomes and

a BEAST analysis to estimate the divergence dates of Cff and Cfv strains. Additionally, we investigated whether the genomes contain specific SNPs or genes that could be associated with the biochemical tests and different clinical features of the *C. fetus* strains.

Methods

Bacterial strains and whole genome sequencing

In this study, 42 *C. fetus* strains from different countries and sources were included (Table 1). The strains were biochemically characterized (except the NCBI GenBank strains H1-UY, 642-21, B6, and 99/541), using H_2S production in medium amended with 0.02% cysteine-HCl and 1% glycine tolerance, as described before [15]. Genotypic subspecies characterisation was performed using MLST [8] and AFLP [16]. The whole genome sequence data of all *C. fetus* genomes was published previously [9, 12, 17] . Briefly, the *C. fetus* genomes (except the genomes with accession numbers starting with ERR and the genomes downloaded from NCBI Genbank) were sequenced using a Roche 454 GS-FLX+ Genome sequencer with Titanium chemistry and assembled into contigs using the Newbler Assembler (version 2.6). The genomes with accession numbers starting with ERR were sequenced using an Illumina Hiseq and assembled into contigs with the Velvet assembler in an established pipeline at the Sanger Institute. The genomes of four *C. fetus* strains (04/554, 97/608, 03/293 and 01/165) were closed by using a PacBio RS sequencer and assembled into contigs using Quiver (Pacific Biosciences, CA, USA) with the base calls validated using Illumina MiSeq reads.

The sequences with accession numbers starting with ERR are available from the European Nucleotide Archive (ENA) [18] and the remaining sequences are available from NCBI GenBank (Table 1). The following genomes were also included; Cff strain H1-UY (GenBank accession number JYCP00000000), Cfv strain 642-21 (GenBank accession number AJSG00000000), Cfv strain B6 (GenBank accession number AJMC00000000), Cfv strain NCTC 10354 (GenBank accession number AFGH00000000) and Cfvi strain 99/541 (GenBank accession number ASTK00000000). The quality of the whole genome sequence data was assessed with the Checkm tool [19], showing a suitable completeness score of >96% with 600 tested marker genes.

Genome alignment and phylogenetic core genome SNP analysis

For phylogenetic core genome SNP analysis, whole genome sequences of 42 *C. fetus* isolates were aligned using Parsnp v1.2 [20]. We included Cft strain 03-427 as an outgroup for the phylogenetic core genome SNP analysis, but excluded this strain from visualization

a le Strain characteristics

Strain number	Accession number	Serotype	Bovine abortion	Isolation year	Clade (Fig. 1)	Subspecies		1% glycine tolerance	H_2S production
						Genotype	Phenotype		
98/v445	LMBH00000000	B	No	1998	1	Cff	Cff	+	+
B0066	ERR419610	B	No	2013	1	Cff	Cff	+	+
B0130	ERR419638	B	No	2013	1	Cff	Cff	+	+
B0129	ERR419637	B	No	2013	1	Cff	Cff	+	+
S0478D	ERR419653	B	No	2011	1	Cff	Cff	+	+
B0042	ERR419595	B	No	2013	2	Cff	Cff	+	+
S0693A	ERR419284	B	No	2012	2	Cff	Cff	+	+
B0167	ERR460866	B	No	2013	2	Cff	Cff	+	+
B0168	ERR460867	B	No	2013	2	Cff	Cff	+	+
04/554	CP008808-CP008809	B	Yes	2004	2	Cff	Cff	+	+
B0047	ERR419600	B	No	2013	2	Cff	Cff	+	+
B0151	ERR419648	B	No	2013	2	Cff	Cff	+	+
B0152	ERR419649	B	No	2013	2	Cff	Cff	+	+
110800-21-2	LSZN00000000	A	No	2000	3	Cff	Cff	+	+
B0097	ERR419623	A	No	2013	3	Cff	Cff	+	+
BT 10/98	LRAL00000000	A	Unknown	1998	3	Cff	Cff	+	+
H1-UY	JYCP00000000	A	No	2013	4	Cff	n.a.	n.a.	n.a.
82-40	CP000487	A	No	1982	4	Cff	Cff	+	+
B0131	ERR419639	A	No	2013	4	Cff	Cff	+	+
Zaf 3	LREZ00000000	A	Yes	2006	5	Cfvi	Cfvi	-	+
CCUG 33872	LREU00000000	A	Unknown	1995	5	Cfvi	Cfvi	-	+
642-21	AJSG00000000	A	Unknown	n.a.	5	Cfvi	n.a.	n.a.	n.a.
ADRI 513	LRFA00000000	A	Unknown	n.a.	5	Cfv	Cfvi	-	+
CCUG 33900	LREV00000000	A	Yes	1995	5	Cfv	Cfv	-	-
LMG 6570	LREW00000000	A	Unknown	1953	5	Cfv	Cfv	-	-
B6	AJMC00000000	A	Unknown	n.a.	5	Cfv	n.a.	n.a.	n.a.
NCTC 10354	AFGH00000000	A	Unknown	1962	5	Cfv	Cfv	-	-
84-112	HG004426-HG004427	A	Unknown	1984	5	Cfv	Cfv	-	-
97/608	CP008810-CP008812	A	Yes	1997	5	Cfv	Cfv	-	-
B10	LRET00000000	A	Unknown	n.a.	5	Cfv	Cfv	-	-
WBT 011/09	LMBI00000000	A	Unknown	2009	5	Cfvi	Cfvi	-	+
Zaf 65	LREY00000000	A	Unknown	2007	5	Cfvi	Cff	+	+
97/532	LRER00000000	A	No	1997	5	Cfvi	Cfvi	-	+
01/165	CP014568-CP014570	A	No	2001	5	Cfvi	Cfvi	-	+
03/293	CP0006999-CP007002	A	Yes	2003	5	Cfvi	Cff	+	+
92/203	LRVL00000000	A	Yes	1992	5	Cfvi	Cfvi	-	+
03/596	LRAM00000000	A	Yes	2003	5	Cfvi	Cfvi	-	+
02/298	LRVK00000000	A	Yes	2002	5	Cfvi	Cfvi	-	+
ADRI 1362	LREX00000000	A	Unknown	1989	5	Cfvi	Cff	+	+
98/25	LRES00000000	A	Yes	1998	5	Cfvi	Cfv	-	-
INTA 99/541		A	Unknown	1999	5	Cfvi	n.a.	n.a.	n.a.
03/427	CP006833	A	No	2003	n.a.	Cft	Cft	+	+

n.a. not available + positive - negative
Cff *C. fetus* subsp. *fetus*
Cfv *C. fetus* subsp. *venerealis*
Cfvi *C. fetus* subsp. *venerealis* biovar intermedius
Cft *C. fetus* subsp. *testudinum*

of the phylogenetic core genome SNP tree to get a better resolution of the Cff and Cfv branches. SNP discovery was focused on a comparison between Cff and Cfv strains; therefore, Cft strain 03-427 was replaced by Cff strain 82-40 as a reference for this analysis. Recombination regions in the core genome alignment were detected and visualized using Gubbins [21]. A phylogenetic tree was constructed using FastTree2 [22] with a generalized time-reversible model and gamma correction on the recombination-filtered SNPs in the core genome of all isolates, including Cft strain 03-427. The resulting tree was rooted on Cft strain 03-427 using Dendroscope [23] prior to visualization using iTOL [24].

BEAST analysis

Recombination-filtered non-synonymous SNPs of the mammal-associated *C. fetus* isolates were extracted from the Gubbins results and used for divergence dating in BEAST [25], using the isolation dates as tip dates in the phylogenetic tree, as outlined in the BEAST manual with the following modifications: 10,000,000× sampling and a general time reversible (GTR) model plus gamma correction as the distance model. Combinations of a strict clock, log-normal clock, exponential clock and random local clock as the clock model and a constant population, exponentially-growing population, and a Bayesian skyline plot with six groups as demographic models were used. Maximum ESS values and lowest 95% confidence intervals (CI) in the predicted divergence dates of the clades were obtained with a random local clock and a Bayesian skyline plot with four groups as the demographic model.

Calculation of branch-specific dN/dS ratios

To calculate the dN/dS ratio per branch in the phylogenetic tree, we aligned the genomes of the mammal-associated *C. fetus* isolates with Parsnp v1.2 [20] and Cft strain 03-427 as outgroup. Synonymous and non-synonymous SNPs were determined on the basis of their location in the coding regions of the Cff strain 82-40 reference genome. Recombination regions were detected using Gubbins [21] and excluded from the alignment. Ancestral state reconstruction of the node sequences was performed using FastML [26] with a generalized time-reversible model and gamma correction. The dN/dS ratios were determined per branch between node sequences.

Comparison of SNPs and genes with phenotypic characteristics of the strains

As traditional differentiation of *C. fetus* subspecies are based on the biochemical 1%

glycine tolerance and H_2S production tests, the genomes were screened for genes and SNPs that were associated with these phenotypic characteristics, and tested for 1% glycine tolerance and H_2S production in cysteine-rich medium as described above. The protein-encoding gene presence and absence were determined using BLAST-based all vs all comparisons with Prokka-annotated genomes [22] using Roary [27], which clustered the proteins using MCL-edge [28]. A Fisher's exact test was used to calculate the two-tail probability value (*p*) of respectively the detected SNPs and genes versus the outcome of the biochemical tests.

Calculation of clinical association

A Fisher's exact test was used to calculate the two-tail probability value (*p*) of SNPs and genes which were specifically found in strains that were isolated from bovine abortions. Significantly-associated SNPs or genes were checked for their presence in Cff strain 04/554, as this is a non-Cfv bovine clinical strain. SNPs and genes that were not present in this strain were excluded from this analysis, to separate clinical-associated from phylogenetically-associated SNPs and genes.

Availability of data

Genome sequences are available from the European Nucleotide Archive (ENA) and from NCBI GenBank, with the accession numbers listed in Table 1.

Results

Phylogenetic analysis of the core genome SNPs

Phylogenetic analysis, based on core genome SNPs using Cft as the outgroup, showed that the mammal-associated Cff and Cfv genomes group into five distinct clades (Figure 1). The division of clades was consistent with the serotypes of the strains; clades 1 and 2 consist of serotype B strains and clades 3, 4 and 5 consist of serotype A strains. The division of clades was also consistent with the classification of MLST sequence types (STs), except for Cff strain H1-UY, which has the Cfv-associated MLST ST-4 genotype [13]. Interestingly, SNP phylogeny showed the divergence of clade 4 and clade 5 from a common Cff ancestor. Clade 5 consists of phenotypically-identified Cff, Cfvi and Cfv strains (Table 1) without a clear separation in the phylogeny. However, when the genotypic characterization is used [8, 16], clade 5 consists exclusively of Cfv and Cfvi strains and clade 1-4 of Cff strains.

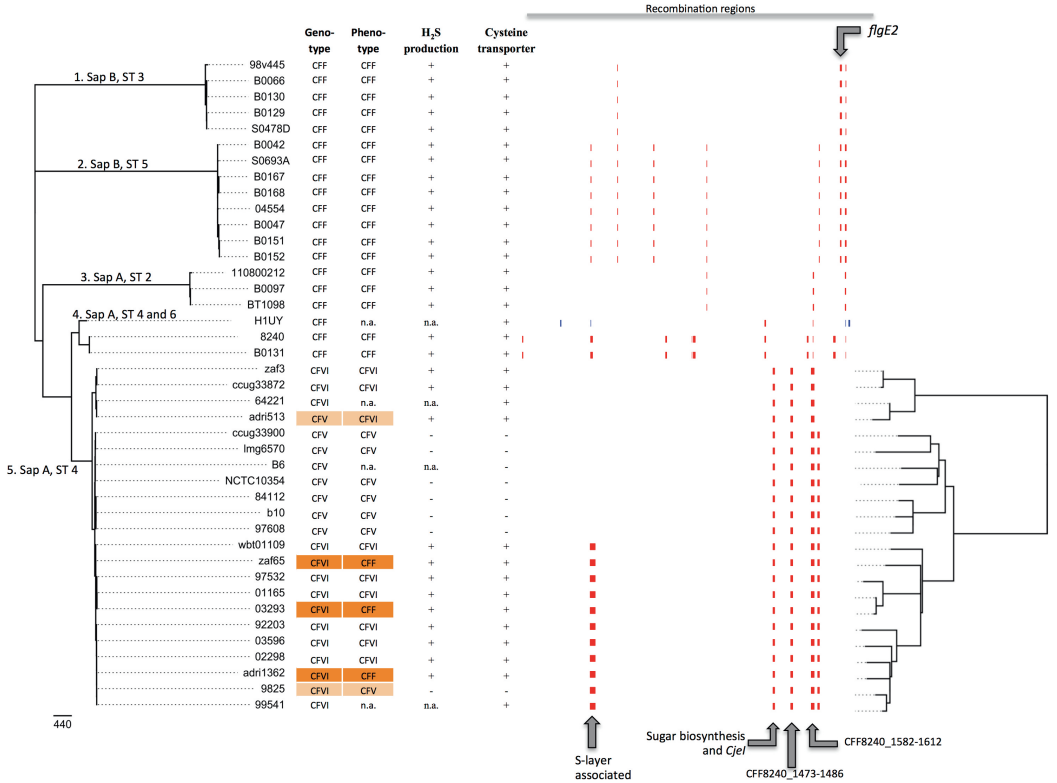

Figure ML tree of C fetus core genome and mammalian associated C fetus recom ination regions Cft 03-427 was excluded from the visualization. Tree branches are labeled with the clade numbers, serotype and MLST STs of the *C. fetus* strains. Subspecies identification (genotype and phenotype), H_2S production test results (+ positive, - negative), the presence of the putative cysteine transporter (+ present, - absent) and the Gubbins predicted recombination regions in the mammalian associated *C. fetus* core genome are shown, with on the right side a zoomed ML tree of clade 5. Gene *flgE2* encodes a flagellar hook protein and *cjeI* a type II R-M system protein. Scale bar at the bottom left represents the nucleotide substitutions per site.

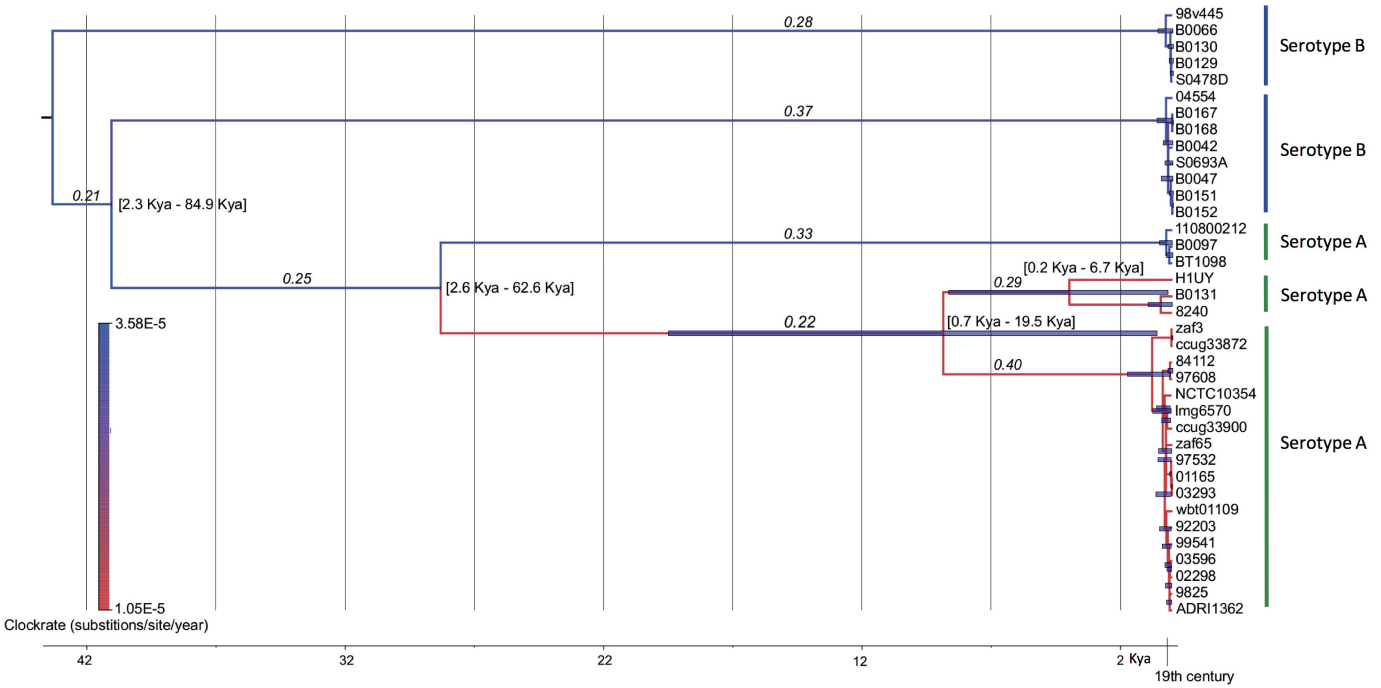

Figure 2 BEAS tree with divergence dates BEAST analysis 95% CI divergence dates are shown within brackets and with blue bars on branch nodes. dN/dS ratios are shown in italics on the branches. Branches are colored according the clock rates, represented as substitutions per site per year. Scale bar at the bottom represents thousand years ago (Kya).

a le 2 Core genome SNPs specific for *C. fetus* clades

Clade	Representing	Number of clade-specific SNPs	Number of clade-specific orthologs (annotation)
1	Cff serotype B	1547	2 (hypothetical)
2	Cff serotype B	1703	41 (fic, transposase, hypothetical)
3	Cff serotype A	1430	6 (R-M system type I and III)
4	Cff serotype A	121	0
5	Cfv and Cfvi	282	0
5	Cfv	5	6
1+2	Cff serotype B, total	571	14 (toxin/antitoxin, R-M, S-layer associated)
3-4	Cff serotype A, total	83	14 (glycosyltransferase, methyltransferase, ketoreductase)

Cff *C. fetus* subsp. *fetus*
Cfv *C. fetus* subsp. *venerealis*
Cfvi *C. fetus* subsp. *venerealis* biovar intermedius

BEAST analysis determined that diversification of the mammal-associated C. fetus is a relatively recent event under disruptive selection pressure

We used BEAST to determine the divergence date of the different clades (Figure 2). The topology of the BEAST tree was slightly different than the maximum likelihood (ML) SNPs tree (Figure 1), especially for the serotype B strains. This may be either the result of using only synonymous SNPs in the BEAST analysis or because of the differences between ML methods and methods based on coalescence.

The 95% confidence interval (CI) of the BEAST analysis was too large to estimate the divergence date of clade 2 and clades 3, 4 and 5 [95% CI 2.3 Kya – 84.9 Kya] and the divergence date of clades 4 and 5 [95% CI 0.7 Kya – 19.5 Kya] (Figure 2). Interestingly, the BEAST analysis showed that all currently circulating strains in clades 1, 2, 3 and 5 diverged very recently from each other.

Cfv is restricted to the bovine genital tract, but Cff can be isolated from different animal species, e.g. cattle, sheep, goats, pigs, horses, and humans [2]. Therefore, we attempted to investigate whether this change in niche preference has resulted in increased evolutionary pressure on Cfv. We estimated the rate of molecular evolution by comparing the ratio of non-synonymous (dN) versus synonymous (dS) SNPs (Figure 2). The dN/dS ratio was the highest in the strains of clade 5 (0.40) compared to the ratio of the strains of clade 4 (0.29) and in all branches preceding, based on node sequences reconstructed using FastML (Figure 2). An increased dN/dS ratio (>0.2) is indicative for recently diversifying genomes, whereas lower dN/dS ratios (0.04-0.20) correspond to older SNPs and more diverged genomes [29]. All clades have a dN/dS ratio > 0.2, with the highest dN/dS ratio of clade 5 of 0.40, showing that the strains of clade 5 are evolving under

relatively more diversifying selection than the strains of the other clades.

The estimated molecular clock rate for the mammal-associated *C. fetus* strains was 1.5-3.5x10^{-2} substitutions per kb per year, which corresponds with the estimated molecular clock rate of *C. jejuni* of 1.86-5.81x10^{-2} per kb/year [30].

Clade-specific SNPs and orthologs

The numbers of SNPs and orthologous genes (orthologs) specifically present in each clade (Figure 1) are presented in Table 2, as well as the specific SNPs and orthologs for clades representing the serotypes and subspecies of the strains. The details and annotation of the detected SNPs and orthologs are listed in Supplemental Table 1.

Clades 1-2 consisted only of Cff serotype B strains. For clade 1, 1547 SNPs and 2 orthologs were found, specifically present in clade 1 strains (Table 2, Supplemental Table 1). The two clade 1-specific orthologs were hypothetical proteins with unknown function. Specific for clade 2, 1703 SNPs and 41 orthologs were found (Table 2, Supplemental Table 1). The specific orthologs were located in a region encoding mainly hypothetical proteins, as well as a Fic domain protein and a transposase. This clade 2-specific region is chromosomally located in Cff strain 04/554 (CFF04554_0637 - CFF04554_0684). All strains of both clades 1 and 2 (Cff serotype B) contain 571 specific SNPs and 14 specific orthologs. The clade 1 and 2-specific orthologs included a toxin/antitoxin system (CFF04554_0478 and CFF04554_0479), restriction-modification (R-M) system-associated genes (CFF04554_1799, CFF04554_0318 and CFF04554_0319) and S-layer associated genes (CFF04554_1746, CFF04554_0481 and CFF04554_0484).

Clades 3 and 4 consist of Cff serotype A strains. Clade 3 contains 1430 specific SNPs and six specific orthologs (Table 2, Supplemental Table). The specific orthologs encoded hypothetical proteins of unknown function, as well as R-M type I (CFF8240_0988) and R-M type III genes. For clade 4, 121 specific SNPs and no specific orthologs are found. All strains of clades 3 – 4 (Cff serotype A) contain 83 SNPs and 14 specific orthologs that are mainly located in one region (CFF8240_1591 – CFF8240_1608) that includes a glycosyltransferase, a methyltransferase and a ketoreductase.

Clade 5 consists of both genotypically-identified Cfv and Cfvi strains. For this clade, 282 specific SNPs were found (Supplemental Table 1). The genotypically-identified Cfv strains of clade 5 contain five specific SNPs and six specific orthologs. The specific orthologs encoded hypothetical proteins, a transcriptional regulator (CFV97608_1300), a resolvase and a transposase.

Analysis of SNP regions that are subject to recombination

Interestingly, 42 % (119/282) of the SNPs differing between clades 1-4 and clade 5 are located in regions <1 kb apart. This is suggestive of recombination, since on average a SNP would be expected to occur approx. every 6.5 kb, based on the number of SNPs identified here in genomes of approx. 1.85 mbp. (1.85 mbp./282 SNPs = 6.5 kb/SNP). We predicted potential recombination regions in the mammal-associated *C. fetus* core genome alignment with Gubbins [21] and visualized the results (Figure 1). Using this analysis, one region showed recombinations in all strains of clades 1-4 and three recombination regions were found in all strains of clade 5.

The recombination region common to all strains of clades 1-4 encoded the flagellar hook gene *flgE2* (CFF8240_1769) (Figure 1). All strains of clade 5 contained a recombination region (CFF8240_1393-CFF8240_1398) with a type II R-M gene (*cjeI*) (CFF8240_1393) and genes involved in sugar biosynthesis, including: a NAD-dependent epimerase/dehydratase (CFF8240_1396); a nucleotide sugar dehydrogenase (CFF8240_1397); and a polysaccharide biosynthesis protein (CFF8240_1398). Another clade 5-specific recombination region is CFF8240_1473-CFF8240_1486 that encodes: leader peptidase A (LepA; CFF8240_1473); ribose-phosphate pyrophosphokinase (Prs; CFF8240_1474); a subunit of aspartate carbamoyltransferase (PyrB; CFF8240_1475), a formate dehydrogenase subunit (FdhC; CFF8240_1482) and a glutamate synthase subunit (GltD; CFF8240_1486). In all strains of clade 5, a recombination region was found (CFF8240_1582-CFF8240_1612), encoding: several radical SAM domain proteins; two transketolase subunits (CFF8240_1587 and CFF8240_1589); a methyltransferase (CFF8240_1590); and two glycosyltransferases (CFF8240_1607 and CFF8240_1612). Additionally, the S-layer (*sap)* region (CFF8240_0455-CFF8240_0501) was also identified by Gubbins as a recombination region, although the *sap* locus itself is not included in the core genome alignment due to assembly issues of this region [12].

Deletion of a putative cysteine transporter is associated with H_2S production-negative strains

Currently, the OIE-prescribed biochemical tests to differentiate Cff, Cfv and Cfvi consist of determining tolerance to 1% glycine and production of hydrogen sulfide (H_2S) from L-cysteine [7]. We analyzed the genomes of the *C. fetus* strains for the presence of specific genes associated with these phenotypic characteristics. Interestingly, H_2S negative strains (Table 1) did not encode two subunits of an amino-acid ABC transporter (CFF8240_0780 and CFF8240_0781), that together with an ATP-binding subunit (CFF8240_0779)

putatively form an amino-acid ABC transport system involved in cysteine transport (Kegg module cff_M00234) (Additional file 1). The absence of the two ABC transporter genes in the *C. fetus* H_2S negative strains possibly explains the phenotypic characteristic of these strains. Cfvi strain 98/25 was described before as H_2S positive [8, 9]; however, in this study, the isolates that were sequenced were biochemically characterized and the sequenced isolate of strain 98/25 was H_2S production negative, in accordance with the absence of the putative cysteine transporter in this genome.

Conjugative transfer region associated with strains from bovine abortions

The *C. fetus* strains were arranged according to their clinical features: nine strains were from bovine abortions; seventeen strains were from screenings; and for fifteen strains, the clinical features were unknown (Table 1). No SNPs were found that were both present in the only known Cff strain from a bovine abortion (strain 04/554) and significantly present in the Cfv/Cfvi strains from bovine abortions. However, one region was significantly present in all strains from bovine abortions ($p < 0.05$); this region contains genes encoding a filamentation induced by cyclic AMP (Fic) domain protein (CFV97608_b0017), a DNA-binding protein (CFV97608_b0010) and conjugative transfer (*tra*) proteins (CFV97608_b00014, CFV97608_b0015, CFV97608_b0021). This region is located in Cff strain 04/554, Cfv strain 97/608 and Cfv strain 84-112 on a megaplasmid or an extra-chromosomal element [17]. This *tra/trb* region is present in 11 *C. fetus* strains isolated from non-abortion cases; Cff strain 98/v445 and Cfv/Cfvi strains CCUG 33872, 642-21, ADRI 513, B10, 84-112, LMG 6570, NCTC 10354, Zaf 65, ADRI 1362 and 99-541 (Additional file 2).

Discussion

The core genome SNP analysis of the mammal-associated *C. fetus* strains identified a large number of SNPs. Since available PCR methods and MLST typing schemes are not able to identify *C. fetus* strains correctly to subspecies level [13,15], Cfv/Cfvi-specific SNPs could be used to develop new diagnostic methods, such as a SNP probe-PCR, or to improve the current MLST scheme. However, two drawbacks of a PCR based on Cfv/Cfvi-specific SNPs are that the stability of such SNPs is unknown and that the design of a reliable PCR based on specific SNPs requires a high degree of optimization [31, 32]. In a previous study, we presented the division of the mammal-associated *C. fetus* strains into two different clusters, based on a core gene alignment of five Cff and 18 Cfv/Cfvi strains [9]. In this

study, 19 Cff strains and 22 Cfv/Cfvi were included, which gave a better resolution of the branches with Cff strains, when compared to a previous analysis with five Cff strains [9]. The current core genome SNP analysis separated the *C. fetus* strains into five different clades, that contain either the genotypically-identified Cff (four clades) or Cfv/Cfvi (one clade) strains. We demonstrated that the biochemical differentiation of the Cff, Cfv and Cfvi strains is not supported by the core genome SNP phylogeny: clade 5 consists of phenotypically-identified Cff, Cfv and Cfvi strains without a congruent separation in the phylogeny of these strains, showing that the phenotypic separation of the mammal-associated *C. fetus* strains is not supported by the core genome SNP phylogeny, which gives rise to the consideration if the current phylogenetic subspecies differentiation is still reasonable, as previously mentioned [9].

Forces of Cff and Cfv diversification

Phylogenetic analysis indicated that Cfv/Cfvi clade 5 and Cff clade 4 have a common Cff ancestor. No association of the clades with geographic origin and/or host specificity was observed. Interestingly, the short terminal branches and recent diversification suggest that the isolates within each clade, except for strain H1-UY, have diverged recently. This is in contrast to what is observed in Cft, where terminal branches of the isolates are much longer and a much higher genome diversity is observed [11]. The world-wide spread of the mammal-associated *C. fetus* strains is potentially associated with the improvements in cattle breeding by selection and cross-breeding which started around the 1700s and 1800s, followed by extensive spread of high-producing dairy cows and beef cattle [33], which may have carried only a very limited number of *C. fetus* clones. We suggest that this spread from a limited number of sources has resulted in our current observation of the highly clonal nature of mammal-associated *C. fetus*. Alternatively, the observation of only recently diverged strains is a result of selective sweeps, when a population member with an advantageous trait will take over the population before the trait can spread to other members [34]. It is unlikely that selective sweeps occurred in the mammal-associated *C. fetus* population, since cattle are intensively monitored worldwide for the presence of *C. fetus* subsp. *venerealis* and only one non-clonal *C. fetus* strain was found (H1-UY). Cff strain H1-UY was isolated from the blood of a rural worker in 2013, who was diagnosed with cellulitis and was in daily contact with cattle [13].

The rates of changes at non-synonymous and synonymous SNPs indicate whether a gene is under purifying or diversifying selection [35, 36] and is expressed in the dN/dS ratio. The dN/dS ratio of bacterial genes under stabilizing selection falls generally within the

range 0.04-0.2 [37]. In the mammal-associated *C. fetus*, the dN/dS ratios were all > 0.2, showing that all genomes are under diversifying selection. The highest dN/dS ratio was found for the clade 5 strains (0.40) (Figure 2). This high ratio may be due to the genetic features of these strains: they have no CRISPR-cas systems and the clade 5-strains have 24 specific SNPs in the type II R-M system gene *cjeI* compared to the other strains (Supplemental Table). Type II R-M systems can undergo major changes in specificity by recombination events as shown for *Helicobacter pylori* [38] and R-M type II systems may play a role in plasmid transformation, as described for *C. jejuni* where knockout mutagenesis of gene *cjeI* resulted in a strain with a 1,000-fold-enhanced transformation efficiency [39]. It is unknown whether these SNPs influence the functionality or specificity of this type II R-M system in *C. fetus* strains, but the clade 5-strains are possibly more susceptible to insertion of foreign DNA, since the genomes contain genomic islands (GIs) encoding T4SSs, phages and insertion sequences [9, 17].

H_2S negative phenotype of Cfv strains associated with loss of a putative cysteine transporter

The original classification of the mammal-associated *C. fetus* subspecies Cff and Cfv are based on the 1% glycine tolerance test [40]. Cfv strains can be discriminated from Cfvi strains with the H_2S production test [5]. We show that the H_2S negative *C. fetus* strains have lost a putative cysteine transporter. Without the encoded transporter, the cells are possibly less capable of importing cysteine, which under normal situations is reduced while forming H_2S. The H_2S-negative Cfv strains have a niche restriction to the genital tract of cattle, whereas the H_2S-positive Cfvi and Cff strains are not restricted to the genital tract and assumed to be able to colonize the intestines as well [5]. One may speculate that the Cfv strains have a defect causing the incapability to grow outside of the genital tract, but it is unknown if this partial deletion of the putative cysteine transporter can be associated with the niche restriction of these strains.

Virulence-associated genes specific for clades and strains from bovine abortions

Since the first description of the *C. fetus* subspecies in 1959, it is presumed that Cfv cause disease in the genital tract of cows, like enzootic venereal sterility and abortions in pregnant cows, and that Cff only cause sporadic abortions [40]. It is unknown which genomic features are responsible for the pathogenicity of the *C. fetus* strains, but potential candidates are the surface layer of the *C. fetus* cells [10, 41] and type IV secretion systems (T4SSs) [17, 42, 43]. We studied the genomes of nine strains from

bovine abortions and observed that, in addition to the S-layer proteins and T4SSs, all the strains from bovine abortions contain a region encoding conjugative transfer (*tra*) proteins. This *tra* region is located in the closed genomes of strains 04/554, 97/608 and 84-112 on a plasmid/ICE, which also contain a *trb*-T4SS and *fic* genes. The fic domain proteins encoded by this region are of potential interest, because fic domain proteins have an immune-modulatory function by influencing the cytoskeleton organization of the host cells [44] and are translocated by T4SSs in *Legionella pneumophila*, *Coxiella burnetii* and *Bartonella henselae* [43], which has also been proposed for *C. fetus* strains [45]. The *trb/tra* genes are also present in 11 *C. fetus* strains isolated from non-abortion cases (Additional File 2). It is possible that these strains can also cause bovine abortions; however, more studies, specifically animal experiments investigating the virulence of *C. fetus* strains are required to assess this hypothesis.

Conclusion

Phylogenetic core genome SNP analysis divided the mammal-associated *C. fetus* strains into five different clades, which were consistent with the serotypes, but not with the phenotypes of the strains. BEAST analysis showed that the clade with genotypically-identified Cfv/Cfvi strains has evolved from a Cff ancestor under diversifying selection. Phylogenetic analysis of the core genome SNPs did not differentiate H_2S-negative Cfv from H_2S positive Cfvi strains. The partial deletion of a putative cysteine transporter is observed in all H_2S negative Cfv strains.

Acknowledgements

We like to thank Emma Yee (U.S. Department of Agriculture) for the generation of sequence data, we thank James Bono (U.S. Department of Agriculture) for the generation of PacBio RS reads and thank Dr. Brian Brooks and Dr. John Devenish (Canadian Food Inspection Agency) for providing C. fetus strains and for critical review of this manuscript.

Declarations

Publication charges for this article have been funded by Utrecht University, the Netherlands.

Availability of data and materials

Sequencing data is available from the European Nucleotide Archive (ENA) and NCBI

GenBank under the accession numbers listed in Table 1. Datasets supporting the results of this article are also included in the additional files. The used files for tree generation were uploaded to Treebase repository (Access URL for Figure 1: http://purl.org/phylo/treebase/phylows/study/TB2:S19749 and for Figure 2: http://purl.org/phylo/treebase/phylows/study/TB2:S19751)

Authors' contributions

LG, BD, JW, AZ conceived this study. LG, WG, KF, AZ performed the experiments. LG, BD, JW, AZ analyzed the data. LG wrote the manuscript. BD, WM, KF, JW, AZ were involved in discussions of the work and manuscript revision. All authors approved the final version.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The bacterial isolates were already part of an anonymized strain collection. Evaluation of human patient data has not been performed, therefore no ethics approval was considered necessary.

Supplemental data

Additional file 1: Table S1. Table with details of *C. fetus* SNPs and genes, specific for clades and bovine abortion strains
Additional file 1 can be found online at https://goo.gl/4eaGCs

Additional file 2: Figure S1. Schematic representation of the putative cysteine transporter genes CFF8240_0779 - CFF8240_0781 (white) in strain Cff 8240 with flanking genes *rarD* and *psgA* (grey). The start and stop positions of respectively *rarD* and *psgA* are shown in italics. Cfv strain 97/608 lacks one complete ORF and contain a remnant of one ORF of the putative cysteine transporter. Shown are the similar deletion sites sequences in the Cfv strains containing the incomplete transporter.
Additonal file 2 can be found online at https://goo.gl/g3v4mb

References

1. Wagenaar JA, van Bergen MA, Blaser MJ, Tauxe RV, Newell DG, van Putten JP: *Campylobacter fetus* infections in humans: exposure and disease. *Clin. Infect. Dis.* 2014, 58(11);1579-1586.

2. Thompson SA, Blaser MJ: Pathogenesis of *Campylobacter fetus*. In *Campylobacter*. 3rd edition. Edited by Nachamkin I, Szymanski CM, Blaser MJ. Washington, DC: ASM Press; 2008. p. 401-427.

3. Smibert RM: Genus II *Campylobacter* Sebald and Véron 1963, 907. In Bergey's manual of systematic bacteriology. *Volume 1.* Edited by Krieg NR, Holt HG. Baltimore: Williams, & Wilkins Co.; 1984. p. 111-118.

4. Fitzgerald C, Tu ZC, Patrick M, Stiles T, Lawson AJ, Santovenia M, Gilbert MJ, van Bergen M, Joyce K, Pruckler J, Stroika S, Duim B, Miller WG, Loparev V, Sinnige JC, Fields PI, Tauxe RV, Blaser MJ, Wagenaar JA: *Campylobacter fetus* subsp. *testudinum* subsp. nov., isolated from humans and reptiles. *Int. J. Syst. Evol. Microbiol* 2014, 64(Pt 9);2944-2948.

5. Florent A: A propos dese vibrions responsables de la vibriose génitale des bovins et des ovins. *Bull. Off. Int. Epizoot.* 1963, 60;1063-1074.

6. Garcia MM, Eaglesome MD, Rigby C: Campylobacters important in veterinary medicine. *Veterinary bulletin* 1983, 53(9);793-818.

7. OIE: Bovine Genital Campylobacteriosis. In Manual of Diagnostic Tests and Vaccines for Terrestrial Animals (mammal, birds and bees). 7th edition. Edited by Anonymous Paris: Office International des Epizooties; 2012. p. 652.

8. van Bergen MAP, Dingle KE, Maiden MC, Newell DG, van der Graaf-van Bloois L, van Putten JP, Wagenaar JA: Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. *J. Clin. Microbiol.* 2005, 43(12);5888-98.

9. van der Graaf-van Bloois L, Miller WG, Yee E, Rijnsburger M, Wagenaar JA, Duim B: Inconsistency of phenotypic and genomic characteristics of *Campylobacter fetus* subspecies requires re-evaluation of current diagnostics. *J. Clin. Microbiol.* 2014, 52(12);4183-4188.

10. Blaser MJ, Smith PF, Hopkins JA, Heinzer I, Bryner JH, Wang WL: Pathogenesis of *Campylobacter fetus* infections: serum resistance associated with high-molecular-weight surface proteins. *J. Infect. Dis.* 1987, 155(4);696-706.

11. Tu ZC, Dewhirst FE, Blaser MJ: Evidence that the *Campylobacter fetus* sap locus is an ancient genomic constituent with origins before mammals and reptiles diverged. *Infect. Immun.* 2001, 69(4);2237-2244.

12. Gilbert MJ, Miller WG, Yee E, Zomer AL, van der Graaf-Van Bloois L, Fitzgerald C, Forbes KJ, Meric G, Sheppard SK, Wagenaar JA, Duim B: Comparative genomics of *Campylobacter fetus* from reptiles and mammals reveals divergent evolution in host-associated lineages. *Genome biology and evolution* 2016, 8(6);2006-2019.

13. Iraola G, Betancor L, Calleros L, Gadea P, Algorta G, Galeano S, Muxi P, Greif G, Perez R: A rural worker infected with a bovine-prevalent genotype of *Campylobacter fetus* subsp. *fetus* supports zoonotic transmission and inconsistency of MLST and whole-genome typing. *Eur. J. Clin. Microbiol. Infect Dis* 2015, 34(8);1593-1596.

14. Sheppard SK, Maiden MC: The evolution of *Campylobacter jejuni* and *Campylobacter coli*. *Cold Spring Harb. Perspect. Biol.* 2015, 7(8);a018119.

15. van der Graaf-van Bloois L, van Bergen MA, van der Wal FJ, de Boer AG, Duim B, Schmidt T, Wagenaar JA: Evaluation of molecular assays for identification *Campylobacter fetus* species and subspecies and development of a *C. fetus* specific real-time PCR assay. *J. Microbiol. Methods* 2013, 95(1);93-97.

16. Wagenaar JA, van Bergen MAP, Newell DG, Grogono-Thomas R, Duim B: Comparative study using amplified

fragment length polymorphism fingerprinting, PCR genotyping, and phenotyping to differentiate *Campylobacter fetus* strains isolated from animals. *J. Clin. Microbiol.* 2001, 39(6);2283-6.

17. van der Graaf-van Bloois L, Miller WG, Yee E, Gorkiewicz G, Forbes KJ, Zomer AL, Wagenaar JA, Duim B: *Campylobacter fetus* subspecies contain conserved type IV secretion systems on multiple genomic islands and plasmids. *PLoS One* 2016, 11(4);e0152832.

18. Leinonen R, Akhtar R, Birney E, Bower L, Cerdeno-Tarraga A, Cheng Y, Cleland I, Faruque N, Goodgame N, Gibson R, Hoad G, Jang M, Pakseresht N, Plaister S, Radhakrishnan R, Reddy K, Sobhany S, Ten Hoopen P, Vaughan R, Zalunin V, Cochrane G: The European Nucleotide Archive. *Nucleic Acids Res* 2011, 39;D28-31.

19. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW: CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015, 25(7);1043-1055.

20. Treangen TJ, Ondov BD, Koren S, Phillippy AM: The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol.* 2014, 15(11);524.

21. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR: Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 2015, 43(3);e15.

22. http://meta.microbesonline.org/fasttree/. Accessed 04/12 2016.

23. Huson DH, Scornavacca C: Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* 2012, 61(6);1061-1067.

24. Letunic I, Bork P: Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 2007, 23(1);127-128.

25. Drummond AJ, Rambaut A: BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 2007, 7;214.

26. Pupko T, Pe'er I, Hasegawa M, Graur D, Friedman N: A branch-and-bound algorithm for the inference of ancestral amino-acid sequences when the replacement rate varies among sites: Application to the evolution of five gene families. *Bioinformatics* 2002, 18(8);1116-1123.

27. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J: Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015, 31(22);3691-3693.

28. Enright AJ, Van Dongen S, Ouzounis CA: An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* 2002, 30(7);1575-1584.

29. Castillo-Ramirez S, Harris SR, Holden MT, He M, Parkhill J, Bentley SD, Feil EJ: The impact of recombination on dN/dS within recently emerged bacterial clones. *PLoS Pathog* 2011, 7(7);e1002129.

30. Wilson DJ, Gabriel E, Leatherbarrow AJ, Cheesbrough J, Gee S, Bolton E, Fox A, Hart CA, Diggle PJ, Fearnhead P: Rapid evolution and the importance of recombination to the gastroenteric pathogen *Campylobacter jejuni*. *Mol. Biol. Evol.* 2009, 26(2);385-397.

31. Hur D, Kim MS, Song M, Jung J, Park H: Detection of genetic variation using dual-labeled peptide nucleic acid (PNA) probe-based melting point analysis. *Biol. Proced. Online* 2015, 17;14-015-0027-5. eCollection 2015.

32. Syvanen AC: Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* 2001, 2(12);930-942.

33. Orozco-ter Wengel P, Barbato M, Nicolazzi E, Biscarini F, Milanesi M, Davies W, Williams D, Stella A, Ajmone-

Marsan P, Bruford MW: Revisiting demographic processes in cattle with genome-wide population genetic analysis. *Front. Genet.* 2015, 6;191.

34. Bendall ML, Stevens SL, Chan LK, Malfatti S, Schwientek P, Tremblay J, Schackwitz W, Martin J, Pati A, Bushnell B, Froula J, Kang D, Tringe SG, Bertilsson S, Moran MA, Shade A, Newton RJ, McMahon KD, Malmstrom RR: Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *ISME J.* 2016.

35. Kimura M: Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. *Proc. Natl. Acad. Sci. U S A* 1991, 88(14);5969-5973.

36. Nei M: Selectionism and neutralism in molecular evolution. *Mol. Biol. Evol.* 2005, 22(12);2318-2342.

37. Rocha EP, Smith JM, Hurst LD, Holden MT, Cooper JE, Smith NH, Feil EJ: Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J. Theor. Biol.* 2006, 239(2);226-235.

38. Nobusato A, Uchiyama I, Kobayashi I: Diversity of restriction-modification gene homologues in *Helicobacter pylori*. *Gene* 2000, 259(1-2);89-98.

39. Holt JP, Grant AJ, Coward C, Maskell DJ, Quinlan JJ: Identification of Cj1051c as a major determinant for the restriction barrier of *Campylobacter jejuni* strain NCTC11168. *Appl. Environ. Microbiol.* 2012, 78(22);7841-7848.

40. Florent A: Les deux vibriosis génitales; la vibriose due à *V. fetus venerealis* et la vibriose d'origine intestinale due à *V. fetus intestinalis*. *Mededelingen der Veeartsenijschool van de RijksUniversiteit te Gent* 1959, (3);1-60.

41. Blaser MJ: Role of the S-layer proteins of *Campylobacter fetus* in serum-resistance and antigenic variation: a model of bacterial pathogenesis. *Am. J. Med. Sci.* 1993, 306(5);325-329.

42. Kienesberger S, Sprenger H, Wolfgruber S, Halwachs B, Thallinger GG, Perez-Perez GI, Blaser MJ, Zechner EL, Gorkiewicz G: Comparative genome analysis of *Campylobacter fetus* subspecies revealed horizontally acquired genetic elements important for virulence and niche specificity. *PLoS One* 2014, 9(1);e85491.

43. Gorkiewicz G, Kienesberger S, Schober C, Scheicher SR, Gully C, Zechner R, Zechner EL: A genomic island defines subspecies-specific virulence features of the host-adapted pathogen *Campylobacter fetus* subsp. *venerealis*. *J. Bacteriol.* 2010, 192(2);502-17.

44. Worby CA, Mattoo S, Kruger RP, Corbeil LB, Koller A, Mendez JC, Zekarias B, Lazar C, Dixon JE: The fic domain: regulation of cell signaling by adenylylation. *Mol. Cell.* 2009, 34(1);93-103.

45. Kienesberger S, Schober Trummler C, Fauster A, Lang S, Sprenger H, Gorkiewicz G, Zechner EL: Interbacterial macromolecular transfer by the *Campylobacter fetus* subsp. *venerealis* type IV secretion system. *J. Bacteriol.* 2011, 193(3);744-758.

Chapter 6

Campylobacter fetus su species contain conserved type IV secretion systems on multiple genomic islands and plasmids

Linda van der Graaf – van Bloois[a,b], William G. Miller[c], Emma Yee[c], Gregor Gorkiewicz[d], Ken J. Forbes[e], Aldert L. Zomer[a,b], Jaap A. Wagenaar[a,b,f], Birgitta Duim[a,b]

[a] Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands
[b] WHO Collaborating Center for Campylobacter / OIE Reference Laboratory for Campylobacteriosis
[c] Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA, USA
[d] Institute of Pathology, Medical University of Graz, Graz, Austria
[e] School of Medicine and Dentistry, University of Aberdeen, United Kingdom
[f] Central Veterinary Institute of Wageningen UR, Lelystad, The Netherlands

PLoS One. 2016, 11 (4):e0152832

A stract

The features contributing to differences in pathogenicity of the *Campylobacter fetus* subspecies are unknown. Putative factors involved in pathogenesis are located in genomic islands that encode a type IV secretion system (T4SS) and fic domain (filamentation induced by cyclic AMP) proteins, which may disrupt host cell processes. In the genomes of 27 *C. fetus* strains, three phylogenetically-different T4SS-encoding regions (T4SSs) were identified: one was located in both the chromosome and in extra-chromosomal plasmids; one was located exclusively in the chromosome; and one exclusively in extra-chromosomal plasmids. We observed that *C. fetus* strains can contain multiple T4SSs and that homologous T4SSs can be present both in chromosomal genomic islands (GI) and on plasmids in the *C. fetus* strains. The GIs of the chromosomally located T4SS differed mainly by the presence of *fic* genes, insertion sequence elements and phage-related or hypothetical proteins. Comparative analysis showed that T4SS sequences, inserted in the same locations, were conserved in the studied *C. fetus* genomes. Using phylogenetic analysis of the T4SSs, it was shown that *C. fetus* may have acquired the T4SS regions from other *Campylobacter* species by horizontal gene transfer. The identified T4SSs and *fic* genes were found in Cff and Cfv strains, although the presence of T4SSs and *fic* genes were significantly associated with Cfv strains. The T4SSs and *fic* genes could not be associated with S-layer serotypes or geographical origin of the strains.

Introduction

Campylobacter fetus (*C. fetus*) contains currently three subspecies: *C. fetus* subsp. *fetus* (Cff), *C. fetus* subsp. *venerealis* (Cfv) and *C. fetus* subsp. *testudinum* (Cft), and the Cfv variant *C. fetus* subsp. *venerealis* biovar intermedius (Cfvi) [1,2]. Cff and Cfv are primarily associated with mammals [1,3], whereas Cft is associated with reptiles [2,4]. Cff and Cfv are highly related at the genome level [5,6], but are adapted to distinct hosts. *C. fetus* subsp. *fetus* can cause sporadic infections in humans, abortion in cattle and sheep and can be isolated from a variety of sites in different hosts [7]. *C. fetus* subsp. *venerealis* is restricted to the genital tract of cattle and is the causative agent of Bovine Genital Campylobacteriosis (BGC), a syndrome characterized by fertility problems in cattle [8].

A 57 kb genomic island encoding a type IV secretion system (T4SS) was identified in Cfv by Gorkiewicz *et. al*.[9]. This T4SS is analogous to the T4SS of *Agrobacterium tumefaciens*, and is considered to function as a type IVa class T4SS. In *A. tumefaciens*, the type IV translocation pilus is encoded by the *virB* operon that consists of eleven genes (*virB1-virB11*); translocation is also dependent on an additional gene *virD4*, which encodes the type IV coupling protein T4CP. The previously-identified genomic island of Cfv contains the *virB/virD4* T4SS, plasmid-related genes and two fic (filamentation induced by cAMP) domain-encoding genes [9]. The T4SS of Cfv has been shown to be functional and supports intra- and interspecies conjugative DNA transfer [10]. The fic domain proteins have critical roles in multiple cellular processes, including disrupting the host cell processes that are important to pathogen survival and replication, after transmission into eukaryotic cells [11]. It was hypothesized that this genomic island is responsible for the pathogenicity and clinical symptoms manifested during Cfv infections [10].

Other features responsible for the pathogenicity of *C. fetus* strains are the surface layer proteins (SLPs) that cover *C. fetus* cells [12-15]. The *C. fetus* SLPs undergo antigenic variation and protect the cell against the host immune system. The mammalian *C. fetus* strains can be serotyped into two major groups, serotype A or serotype B [16]. As both the S-layer proteins and T4SS regions are suggested to have a role in the pathogenicity of *C. fetus*, it might be possible that these features have a synergistic role in immune escape.

In the first description of a *C. fetus* genomic island harboring a T4SS, it was concluded that this genomic island was specific for *C. fetus* subsp. *venerealis* [9]. From the recently published *C. fetus* genomes [17-19], it has become clear that the genome of Cff strain 04/554 contains a T4SS on a megaplasmid and that some *C. fetus* strains can even harbor

multiple T4SSs. The genome of Cfv strain 84-112 harbors four T4SSs; two genomic islands contain a T4SS and two T4SSs were located in an extra-chromosomal element [17]. It is unknown if *C. fetus* commonly harbors multiple T4SS-encoding regions and how dispersed the different T4SSs are among *C. fetus* strains and the *C. fetus* subspecies.

In this study, we examined the diversity of T4SS-encoding regions in 27 *C. fetus* strains using comparative genomics, and identified the location and composition of all T4SS encoding regions and their phylogeny. Furthermore, we studied whether the presence of specific T4SSs and *fic* genes could be associated with the *C. fetus* subspecies, their pathogenicity, the S-layer serotypes and geographic origin of the strains. Phylogenetic analysis with T4SSs of other *Campylobacter* species suggested that the *C. fetus* T4SS regions did not evolve from the same ancestor, but were acquired from different donors.

Materials and Methods

Bacterial strains

In this study, 27 *C. fetus* strains from different countries and sources were analysed (Table 1). The phenotypic and genotypic characteristics of the strains were described previously [19].

Whole genome sequencing

The *C. fetus* strains (except strains B0066, B0097, B0131 and B0167) were sequenced using a Roche 454 GS-FLX+ Genome sequencer with Titanium chemistry. Roche 454 reads were assembled into contigs using the Newbler Assembler (version 2.6). The remaining four Cff strains (B0066, B0097, B0131 and B0167) from the UK were sequenced according to the following procedure; the isolation of genomic DNA for whole genome sequencing (WGS) used the Promega Wizard Genomic DNA Purification Kit. All of the DNA samples went through a genomic library prep which is similar to the Illumina Truseq protocol, but which was developed at the Sanger Institute. The libraries were sequenced on Illumina HiSeq 2000 analysers on 100bp paired end runs. The paired read files were de novo assembled using the Velvet assembler in an established pipeline at the Sanger Institute.

To span repeat regions, four *C. fetus* genomes (04/554, 97/608, 03/293 and 01/165) were sequenced with a PacBio RS sequencer (Keygene N.V., Wageningen, the Netherlands). PacBio RS reads were assembled into contigs using Quiver (Pacific Biosciences, CA, USA) and the base calls were validated with Illumina MiSeq reads.

The reference genomes of strain 82-40 genome (Genbank accession number CP000487),

a le Characterization and prevalance of T4SS regions in *C. fetus* strains

Strain	Accession number	Country	Source	Identification			Pfam* (VirD4)	Pfam* (Fic)	T4SS (sub)clusters**										
				Phenotypic ID [19]	Genotypic ID [19]	Sap serotype [19]			1						2		3		
									A	B	C	D	E	F	A	B	A	B	C
BT 10/98	LRAL00000000	UK	Ovine	Cff	Cff	A	-	-	-	-	-	-	-	-	-	-	-	-	-
B0097	ERR419623	UK	Bovine (faeces)	Cff	Cff	A	-	-	-	-	-	-	-	-	-	-	-	-	-
110800-21-2	LSZN00000000	NL	Bovine (bull)	Cff	Cff	A	1	2	+	-	-	-	-	-	-	-	-	-	-
98/v445	LMBH00000000	UK	Bovine (bull)	Cff	Cff	B	4	4	+	-	-	-	-	-	-	+	-	+	+
B0066	ERR419610	UK	Bovine (faeces)	Cff	Cff	B	-	-	-	-	-	-	-	-	-	-	-	-	-
04/554	CP008808-008809	AR	Bovine (foetus)	Cff	Cff	B	1	4	-	-	-	-	-	-	-	-	-	-	+
B0167	ERR460866	UK	Bovine (faeces)	Cff	Cff	B	-	4	-	-	-	-	-	-	-	-	-	-	-
82-40	CP000487	US	Human (blood)	Cff	Cff	A	-	-	-	-	-	-	-	-	-	-	-	-	-
B0131	ERR419639	UK	Bovine (faeces)	Cff	Cff	A	-	-	-	-	-	-	-	-	-	-	-	-	-
03/293	CP0006999-007002	AR	Bovine (foetus)	Cff	Cfvi	A	3	3	+	-	+	-	-	-	+	-	-	-	-
ADRI 1362	LREX00000000	AR	Bovine	Cff	Cfvi	A	4	3	-	-	+	+	-	-	+	-	-	-	+
Zaf 65	LREY00000000	SA	Bovine	Cff	Cfvi	A	3	6	-	+	-	-	+	-	-	-	-	+	-
01/165	CP014568-014570	AR	Bovine (mucus)	Cfvi	Cfvi	A	3	2	-	-	-	-	-	+	+	-	+	-	-
02/298	LRVK00000000	AR	Bovine (foetus)	Cfvi	Cfvi	A	4	4	+	-	+	-	-	-	+	-	-	-	+
03/596	LRAM00000000	AR	Bovine (foetus)	Cfvi	Cfvi	A	2	3	-	-	-	-	-	-	+	-	-	-	+
92/203	LRVL00000000	AR	Bovine (placenta)	Cfvi	Cfvi	A	3	4	+	-	-	-	-	+	+	-	-	-	-
97/532	LRER00000000	AR	Bovine (mucus)	Cfvi	Cfvi	A	3	8	-	-	-	-	-	+	+	-	-	-	+
98/25	LRES00000000	AR	Bovine (foetus)	Cfvi	Cfvi	A	4	4	-	-	+	-	-	-	+	-	+	-	+
WBT 011/09	LMBI00000000	UK	unknown	Cfvi	Cfvi	A	5	4	+	+	-	-	-	+	+	+	-	-	-
Zaf 3	LREZ00000000	SA	Bovine (foetus)	Cfvi	Cfvi	A	2	6	-	-	-	-	+	-	-	-	-	-	+
ADRI 513	LRFA00000000	AU	unknown	Cfvi	Cfv	A	4	7	-	-	-	-	+	-	-	-	+	+	+
CCUG 33872	LREU00000000	CZ	unknown	Cfvi	Cfv	A	4	8	-	-	-	-	+	-	-	-	+	+	+
84-112	HG004426-004427	US	Bovine	Cfv	Cfv	A	4	4	+	-	+	-	-	-	+	-	-	-	+
97/608	CP008810-008812	AR	Bovine (placenta)	Cfv	Cfv	A	4	3	+	-	+	-	-	-	+	-	-	-	+
B10	LRET00000000	US	Bovine	Cfv	Cfv	A	3	7	-	-	-	-	+	-	+	-	-	-	+
CCUG 33900	LREV00000000	FR	Bovine (abortion)	Cfv	Cfv	A	2	4	-	-	-	-	-	-	+	-	-	-	+
LMG 6570	LREW00000000	BE	Bovine	Cfv	Cfv	A	3	3	+	-	-	-	-	-	+	-	-	-	+

* Numbers refer to present proteins

** Classification according to Fig 1. + ; region is present, - ; region is absent

Country code: AR, Argentina; AU, Australia; BE, Belgium; CZ, Czech Republic; FR, France, NL, Netherlands; SA, South Africa; UK, United Kingdom; US, United States

Abbreviations: Cff, *Campylobacter fetus fetus*; Cfv, *Campylobacter fetus venerealis*; Cfvi, *Campylobacter fetus venerealis* biovar intermedius

strain 84-112 (Genbank accession numbers HG004426-HG004427), strain 03/293 (Genbank accession numbers CP0006999-CP007002), 04/554 (Genbank accession numbers CP008808-CP008809) and strain 97/608 (Genbank accession numbers CP008810-CP008812) were used to determine the exact locations of T4SS encoding regions. The sequences of strains B0066, B0097, B0131 and B0167 (with accession numbers starting with ERR) are available from the Wellcome Trust Sanger Institute, and the remaining sequences are available from NCBI Genbank (Table 1).

Pfam search and phylogenetic analysis of virD4 and fic genes

From the Roche 454 and Illumina contigs, predicted gene nucleotide and protein sequences were generated using GeneMark.hmm version 3.25 [20]. The predicted protein sequences were used in a Pfam search (version 27.0; Pfam.xfam.org) [21] to identify the matching Pfam families of the proteins. For each genome, proteins that matched with Pfam family T4SS-DNA_transf (PF02534.9) and Fic (PF02661.13) were selected. Genes matching Pfam family T4SS-DNA_transf were annotated as *virD4*.

A Fisher's exact test was used to calculate the two-tail probability value (p) between the presence of *virD4* and *fic* genes versus the different serotypes and subspecies of the strains.

Phylogenetic analysis of the *virD4* and *fic* genes was performed by alignment of these genes with MUSCLE [22] and building a maximum likelihood tree using RAxML (v7.2.8) under the GTRCAT model. Four *fic* genes were included in the phylogenetic analysis as reference: *fic1* (Genbank accession number ACS15152), *fic2* (Genbank accession number ACA64462), *fic3* (Genbank accession number CDF65920) and *fic4* (Genbank accession number CDF65967). The positions of the *fic* genes of the in this study analysed sequences are listed in supplemental table S3 (S3 Table).

Analysis of T4SS encoding regions

The WGS contigs containing the *virD4* gene sequences, as defined above, were sorted out for each genome. The location of each T4SS region was identified by tracing the core genes adjacent to the T4SS genes on the contigs. Using the reference genomes of strains 84-112, 01/165, 04/554, BRIG alignments were created with a 70% upper identity threshold and 50% lower identity threshold [23]. Phylogenetic analysis of the *virB9* genes and complete T4SSs was performed as described for the *virD4* and *fic* genes.

Comparison with T4SS proteins from other Campylobacter species

To calculate the homology of T4SS proteins with proteins from other *Campylobacter* species, a BLASTP comparison to the proteins in the NCBI non-redundant database was performed. The phylogenetic comparison of the *C. fetus* VirD4 proteins with proteins from other *Campylobacter* species was performed by alignment of these genes with MUSCLE [22] and building a maximum likelihood tree using RAxML (v7.2.8) under the GTR model with gamma correction. For *C. fetus*, one VirD4 protein from each phylogenetic cluster was included. The VirD4 protein sequences of other *Campylobacter* species were obtained from GenBank submissions (Fig 3).

Results

Analysis of the fic and virD4 genes and T4SS encoding regions

For each strain, the number of genes that matched with Pfam families T4SS-DNA_transf (*virD4*) and Fic (*fic*) are listed in Table 1.

Phylogenetic analysis of the *fic*-encoding genes demonstrated that the *fic* genes are highly diverse and can be divided into multiple clusters (S1 Fig). Almost all analysed *C. fetus* strains contained multiple *fic* genes, except strains 82-40, BT10/98, B0066, B0097 and B0131. These strains also lacked *virB* and *virD4* genes. Cff strains 04/554 and B0167 contained four *fic* genes in a genomic island, as shown with the BRIG analysis (S2 Fig). The presence of

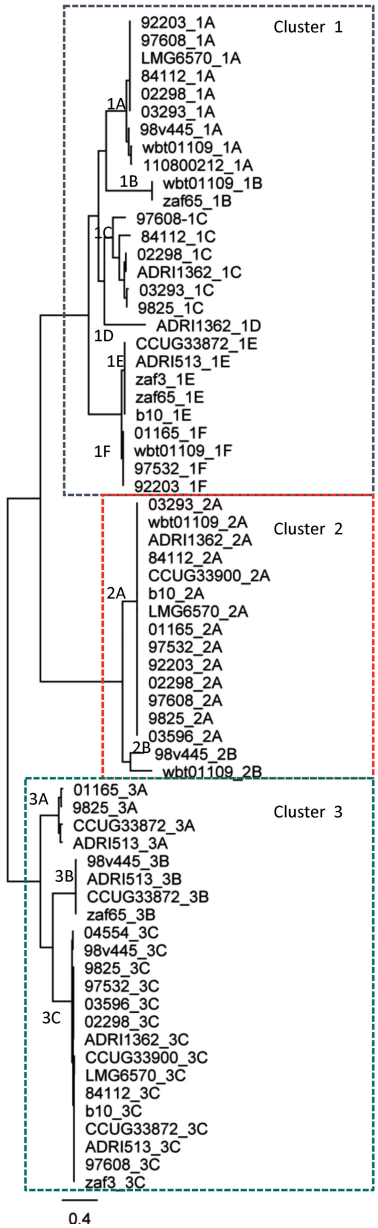

Figure . Phylogenetic analysis of virD4 genes. The scale bar represents the mean number of nucleotide substitutions per site. Related virD4 genes are indicated in boxes.

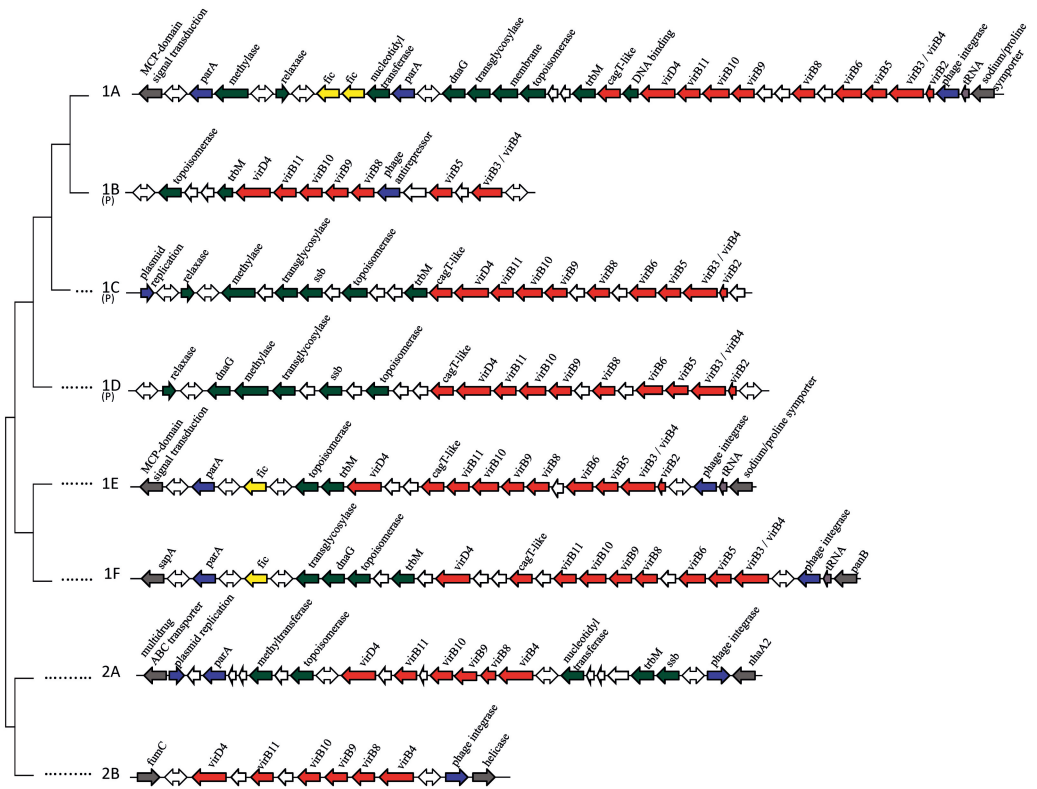

Figure 2 Schematic representation of 4SS regions in genomic islands and plasmids. The T4SS regions and their adjacent genes are listed according to the phylogenetic clustering as shown in Fig 1. The T4SS regions in genomic islands or plasmids (P) contain: vir/cag genes (red), phage integrases and other phage-related genes (blue), genes of plasmid origin (green), tRNA genes (purple), chromosomal integration sites (grey) and genes of unknown function (white). Double-sided arrows indicate genes of unknown function in varying orientations. The genomic islands of cluster 2 are in the chromosomes located in the opposite orientation as shown in Fig 2.

a le 2 Typical features of *C. fetus* T4SS encoding regions

Cluster		Location	Transfer-associated genes					T4SS genes		T4SS with highest homology
			parA	*dnaG*	*trbM*	*cagT*-like	*fic* genes (n)	present	absent	
1	A	GI	+	+	+	+	+ (2)	*virB2-virD4*		*C. hominis*, *C. ureolyticus*
	B	plasmid	-	-	+	-	-	*virB3-virD4*	*virB2, virB6*	
	C	plasmid	-	-	+	+	-	*virB2-virD4*		
	D	plasmid	-	+	-	+	-	*virB2-virD4*		
	E	GI	+	-	+	+	+ (1)	*virB2-virD4*		
	F	GI	+	+	+	+	+ (1)	*virB3-virD4*	*virB2*	
2	A	GI	+	-	+	-	-	*virB4-virD4*	*virB2, virB3, virB5, virB6*	*C. concisus*, *C. showae, C. rectus*
	B	GI	-	-	-	-	-	*virB4-virD4*	*virB2, virB3, virB5, virB6*	
3		plasmid						*tra, trb*		*C. coli* plasmid

GI; genomic island, + ; gene is present, - ; gene is absent

the *fic* genes is significantly associated with *C. fetus* subsp. *venerealis* (p = 0.001). The presence of *fic* genes was not significantly associated with the geographical origin and serotypes of the strains (p = 1.0).

The identified *virD4* genes were classified into three main phylogenetic clusters (Fig 1: clusters 1-3). These clusters were divided further into multiple sub-clusters, designated 1A-1F, 2A-2B and 3A-3C (Fig 1). The *virD4* genes of clusters 1 and 2 are all located within T4SS regions consisting of *virB2-virB11* genes. The *virD4* genes of cluster 3 are located on plasmids, in T4SS regions encoding *tra* and *trb* conjugative transfer genes. The *tra/trb* regions of cluster 3 are more diverse in composition than the sequences of the *virB/virD4* regions (data not shown). The highly diverse *tra/trb* regions of cluster 3 were excluded from assemblies of the T4SS regions and their adjacent genes.

The T4SS regions of cluster 1 are located either in a genomic island (GI) or on a plasmid, whereas the T4SS regions of cluster 2 are located only in a GI. The adjacent genes of the T4SS regions in the GIs and the plasmids were identified (Fig 2). This enabled the assembly of the T4SS-encoding region, which showed extensive diversity in gene content. The typical features of the T4SS regions and their accompanying genes are shown in Table 2. Additionally, through alignment with the reference genomes of strains 84-112 and 01/165, the location of the chromosomal T4SS regions 1A, 1F and 2A in the genomes is shown in supplemental figure S2 (S2 Fig).

Phylogenetic clustering of the *virD4* genes is similar to the clustering of both the *virB9* genes and the complete T4SSs (data not shown), demonstrating not only conservation of

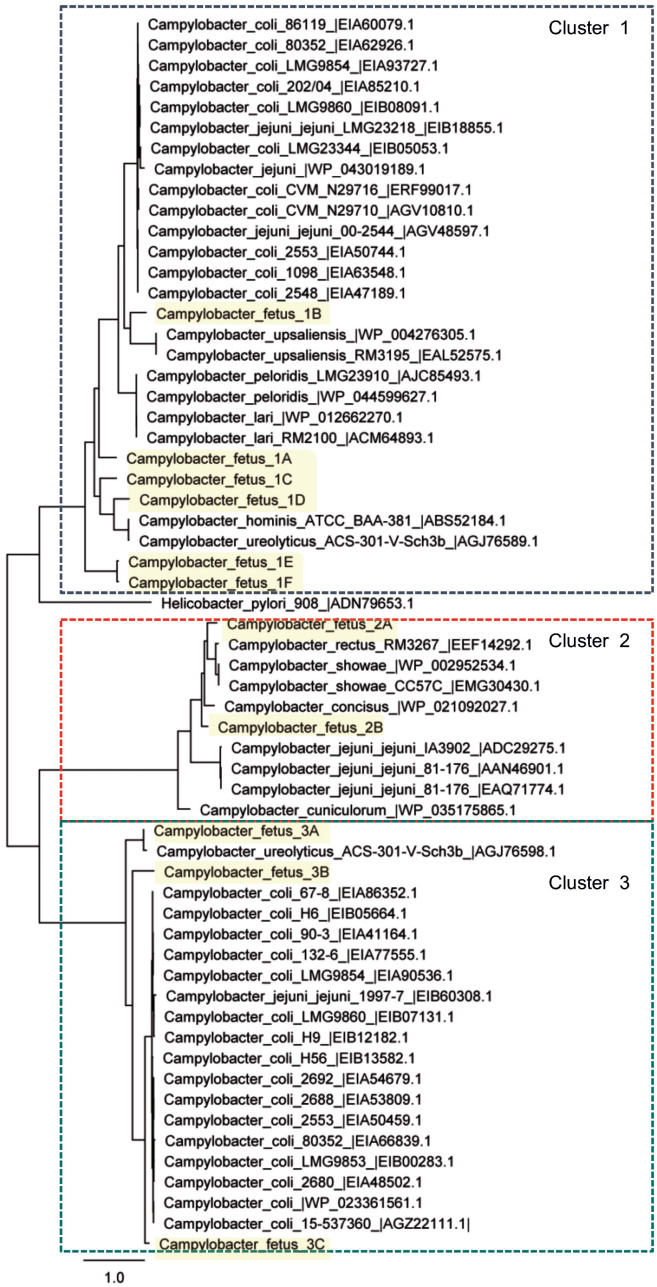

Fig 3 Phylogenetic analysis of virD4 proteins of different *Campylobacter* species The scale bar represents branch length (number of amino acid substitutions/100 residues). Related VirD4 proteins are indicated in boxes.

the sequences of *virD4* gene sequences, but also conservation of the entire T4SS-encoding regions within *C. fetus*.

Location and gene content of the cluster 1 T4SS regions

The T4SS regions of cluster 1 are located in genomic islands (regions 1A, 1E and 1F) as well as on plasmids (regions 1B, 1C and 1D) (Fig 2). The T4SS region 1A is the most studied T4SS of *C. fetus* [9,10,17], and is located in the most complex genomic island of cluster 1 (Fig 2).

The T4SS regions 1A and 1E are phylogenetically positioned in different sub-clusters (Fig 1), but the genomic islands of these T4SS regions are integrated in the same location in the chromosome, between an MCP-domain signal transduction protein (GenBank accession no. YP_892387.1) and a sodium/proline symporter (GenBank accession no. YP_892386.1). The plasmid-associated T4SS region 1B is smaller and is lacking the *virB6* gene, possibly due to the insertion of the phage anti-repressor gene (Fig 2). T4SS region 1F was found in two strains and was located in both strains in a genomic island in the *sap* locus, between a *sapA* homolog and *panB* (GenBank accession no. YP_891665.1) (Fig 2).

The GI of T4SS region 1A contains two *fic* genes, whose products are possibly secreted by the T4SS [10]. The genomic islands of the other chromosomally-located T4SS regions 1E and 1F contain a single *fic* gene. No *fic* genes were found adjacent to the plasmid-associated T4SS regions 1B, 1C and 1D (Table 2, Fig 2).

Adjacent to the T4SS regions, multiple transfer-associated genes are found, e.g. *parA*, *dnaG*, *trbM* and a *cagT*-like gene (Table 2). Beyond the transfer-associated genes, the T4SS-adjacent regions can contain genes encoding a nickase, an *Eco*RI methyltransferase, a helicase, phage-associated genes and genes encoding hypothetical proteins (Fig 2).

The VirD4 proteins of cluster 1 were similar to the VirD4 proteins of multiple other *Campylobacter* species. Most of the VirD4 proteins encoded by other *Campylobacter* species are of plasmid origin (e.g., *C. upsaliensis* EAL52575.1, *C. peloridis* AJC85493.1, *C. lari* ACM64893.1, *C. coli* and *C. jejuni* VirD4 proteins). The genes encoding the VirD4 proteins of *C. hominis* and *C. ureolyticus* were located on their respective chromosomes; these VirD4 proteins are positioned close to the two plasmid-encoded VirD4 proteins of *C. fetus*, 1C and 1D (Fig 3).

Location and gene content of the cluster 2 T4SS regions

The T4SS regions of cluster 2 are exclusively found in *C. fetus* genomic islands. The GI containing T4SS region 2A is inserted between genes encoding a multidrug resistance ABC transporter (GenBank Accession no. YP_892875.1) and *nhaA2* (GenBank Accession no.

YP_892871.1). The GI of T4SS region 2B is inserted between genes encoding the class II fumarate hydratase FumC (GenBank Accession no. YP_892133.1) and a UvrD/rep helicase (GenBank Accession no. YP_892129.1) (Fig 2).

The T4SS regions 2A and 2B are lacking *virB2, virB3, virB5* and *virB6,* and both genomic islands do not contain *fic* genes. The GI of T4SS region 2B contains only the T4SS genes. The GI of T4SS region 2A contains the transfer-associated genes *parA* and *trbM* (Table 2). Furthermore, the GI of T4SS region 2A encodes a nucleotidyltransferase, a topoisomerase and an *Eco*RI methyltransferase, like the GIs of cluster 1 (Fig 2).

Comparison with the T4SS proteins of other *Campylobacter* species in the NCBI non-redundant database showed that the *C. fetus* VirD4 protein of cluster 2 shared 78% homology with the VirD4 proteins of *C. concisus* (GenBank accession no. WP_021092027.1) and *C. rectus* (Genbank accession no. WP_039888059.1) and 79% homology with the VirD4 protein of *C. showae* (Genbank accession no. WP_002952534.1). The gene contents of the T4SSs of cluster 2 are similar to the *vir* operon of *C. showae* and *C. rectus* [24], consisting of *virB4, virB8, virB9, virB10, virB11* and *virD4*. Furthermore, the GIs of T4SS region 2A contain, similar to the C. *showae* GI T4SS region, nucleotidyltransferase-, topoisomerase- and *Eco*RI methyltransferase- encoding genes (Fig 2: T4SS 2A). The *C. fetus* VirD4 proteins of cluster 2 are chromosomally-encoded, but are phylogenetically positioned close to plasmid-encoded VirD4 proteins of three *C. jejuni* strains (Genbank accession nos. EAQ71774.1, ADC29275.1 and AAN46901.1) (Fig 3).

Gene content of the T4SS regions of cluster 3

The *virD4* genes of cluster 3 were located in a T4SS region encoding *tra* and *trb* conjugative transfer genes. As with the plasmid-encoded *virB/virD* region, the *tra/trb* gene cluster is also found in the octopine-type Ti plasmids of *Agrobacterium tumefaciens* [25]. In *C. fetus*, this T4SS is exclusively located on plasmids and was not identified in the chromosomes of the analysed strains. The *tra/trb* T4SS encoding region is located in the extra-chromosomal element ICE_84112 of Cfv strain 84-112 [17] and on the megaplasmids of strain Cff 04/554 and Cfv 97/608 [19]. The *tra/trb* gene clusters were highly diverse in gene content. Furthermore, in both closed Cfv genomes 84-112 and 97/608, this T4SS region is disrupted by several insertion sequence elements. In strains 04/554 and 97/608, both megaplasmids with the *tra/trb* region contain one adjacent *fic* gene. The ICE of strain Cfv 84-112 contains two *fic* genes, but this ICE contains also a *virB/virD4* T4SS, and it is not known if the *fic* genes are linked to the *tra/trb* T4SS or the *virB/virD4* T4SS of this ICE or to both T4SS regions.

The proteins of the *tra/trb* T4SS share a high sequence identity (>90%) with a conjugal transfer locus in a *C. coli* plasmid [26]. The *C. fetus* VirD4 proteins 3A-3C clustered with a large group of mainly plasmid-encoded VirD4 proteins present in *C. coli* and *C. jejuni* (Fig 3).

Distribution of T4SS regions in C. fetus subspecies and association with pathogenicity, geographic origin and S-layer

Of the 27 analysed *C. fetus* strains, 21 strains contained a T4SS (Table 1). The T4SS regions of cluster 1 were identified in 18 strains, and the T4SS regions of cluster 2 in 15 strains. The majority of strains (n=17) contain a *virD4* gene of the *trb/tra* T4SS region of cluster 3. It was common for strains to harbor multiple T4SSs (Table 1). The identified T4SSs were found in Cff and Cfv strains; Cff strain 98/v445 contains four T4SS regions and a single T4SS region was found in Cff strains 110800-21-2 and 04/554, showing that the T4SS regions are not *C. fetus* subsp. *venerealis* specific. However, the Fisher's exact test showed that the presence of VirD4 proteins is significantly associated with *C. fetus* subspecies *venerealis* ($p = 0.003$).

From the set of 27 strains, nine strains were isolated from bovine abortions (strains 04/554, 03/293, 02/298, 03/596, 92/203, 98/25, Zaf3, 97/608 and CCUG 33900). Most of these strains (7 of the 9 strains; except strains 03/293 and 92/203) contained the *tra/trb* T4SS region 3C and most of these strains (8 of the 9 strains, except strain 04/554) contained one or more of the non-chromosomally-located T4SS regions 1B, 1C and 1D, showing that all *C. fetus* strains isolated from abortions contained at least one T4SS. Since the clinical data of the remaining 18 strains was not available, it was not possible to calculate if there is a significance association between the T4SSs and pathogenicity of the strains.

The strains used in this study were obtained from different countries and the T4SSs were distributed between strains from different countries, showing that the presence of a specific T4SS in *C. fetus* strains does not correlate with geographic origin.

The presence of T4SS regions was compared with the S-layer (Sap*)* serotypes of the *C. fetus* strains (Table 1). Cff strains of both Sap serotypes A and B contain the T4SS region of cluster 1. Cff strain 98/v445 of serotype B contained the T4SS region of cluster 2 as well as two *virD4* genes of the *trb/tra* T4SS region. These *trb/tra* T4SSs were absent in Cff serotype A strains, but are present in many Cfv serotype A strains. The presence of the *virD4* genes is not significantly associated with the serotypes of the *C. fetus* strains ($p = 0.20$).

Discussion

Campylobacter fetus subspecies *fetus* and *C. fetus* subspecies venerealis are genetically highly related, but show a different pathogenicity and host adaptation. *C. fetus* subsp. *venerealis* (including Cfv biovar intermedius) is associated with Bovine Genital Campylobacteriosis and is restricted to the genital tract of cattle. *C. fetus* subsp. *fetus* is associated with sporadic abortions in cattle and has a broader host range possibly because of its ability to survive in the gastro-intestinal tract. What pathogenicity motifs could affect the different pathogenesis during the infection process of Cfv, Cfvi and Cff is unknown. This study demonstrates that multiple T4SSs, which have been suggested to be involved in the pathogenicity of C. fetus strains, are present in both subspecies and that the composition of the T4SS-encoding regions is highly diverse.

We have identified the T4SS-encoding regions by searching the genomes with a Pfam search for virD4 genes, as it has been shown before that conjugative transfer systems can be found just by searching for known sequences, like relaxases, T4CPs (virD4) and virB4 genes [27]. With this approach, we were able to identify three phylogenetically different T4SS regions in C. fetus strains.

i. Subspecies specificity of the T4SSs and association with pathogenicity

The genomic island containing T4SS region 1A has been described as Cfv-specific with a prevalence of 76% in Cfv and complete absence in Cff strains [9]. Our study showed that Cff strains can harbour the complete *virB/virD* T4SS region 1A (Cff strains 110800-21-2 and 98/v445), as well as the plasmid encoded *tra/trb* T4SS region of cluster 3 (Cff strains 04/554 and 98/v445), confirming that the T4SS encoding regions are not Cfv-specific [28].

It has been suggested that genomic island genes could be used as specific targets to detect Cfv [29]. In this study, we demonstrated that T4SS genes are present in strains from both subspecies and that no T4SS class is subspecies-specific, confirming that subspecies identification cannot be accomplished using assays that detect T4SS genes.

Nine strains were isolated from bovine abortions; two phenotypic Cff strains and seven Cfv/Cfvi strains. These strains contained the *tra/trb* T4SS region 3C (except strains 03/293 and 92/203) and contained one or more of the non-chromosomally-located T4SS regions 1B, 1C and 1D (except strain 04/554). This suggests that the T4SSs have a potential role in the pathogenicity of the *C. fetus* strains and this is independent of the *C. fetus* subspecies. In this study, we were not able to calculate the significance of the association between the presence of a specific T4SS region to the pathogenicity of the *C. fetus* strains, because detailed information on the clinic and epidemiology of most of the strains was not

available.

Since the S-layer proteins play an important role in the pathogenesis of *C. fetus* infections [12-15], it was studied if the serotypes of the *C. fetus* strains are associated with a T4SS region. The Fisher's exact test showed no significant association of the serotypes and T4SSs of the strains, but one should take into account that only four strains with serotype B were included in this study, making it not possible to determine the association of serotypes and pathogenicity of the *C. fetus* strains.

ii. Functionality of the T4SSs

C. fetus strains can harbour multiple T4SS-encoding regions. The T4SS region of cluster 1 was present in 18 of the 27 *C. fetus* strains. In 16 of these strains, a plasmid encoding a *tra/trb* T4SS region was also present, suggesting that multiple formats for conjugational transfer are present within *C. fetus*. The finding that Cff strains 82-40, BT 10/98, B0066, B0097 and B0131 lack any T4SS-encoding genes confirms that the T4SS is not essential for the *C. fetus* life cycle outside the bovine genital tract. All strains that were isolated from the bovine genital tract contained at least one T4SS.

Within the analysed T4SS regions, all *virB* and *virD4* genes are oriented in the same direction, suggesting an operon structure. Furthermore, both genomic islands of T4SS region 1A and 2A contain a nucleotidyltransferase, and adjacent to T4SS 1A, 1C and 1D a relaxase is found. The presence of these genes suggests a nucleic acid transport function of the T4SSs, as described for the T4SSs of *C. showae* and *C. curvus* [24].

The T4SS genes and composition of cluster 2 are highly homologous with the T4SSs found in *C. showae* and *C. rectus* [24]. In *C. fetus*, it is not demonstrated that this T4SS is functional [17] and it is unknown if this T4SS mediates conjugative DNA transfer between *C. fetus* strains.

An extensively studied virulence locus in *Helicobacter pylori* is the *cag* pathogenicity island (*cag*PAI) encoding a T4SS. The presence of a *cag*PAI discriminates the highly virulent *cag*PAI-positive *H. pylori* strains from the less virulent *cag*PAI-negative *H. pylori* strains [30]. In *H. pylori,* CagA is translocated into the cytoplasm of an infected cell by the T4SS, where it modulates the host immune system [31]. In *C. fetus*, it was hypothesized that the Fic proteins are translocated by the T4SSs, although the secretion of these proteins could not be proved [10]. The translocation of bacterial Fic proteins to the eukaryotic host affects important pathogen recognition processes in the host cell important for survival and replication [10]. Interestingly however, in this study, most of the Fic motif proteins were not located on a contig containing a T4SS gene cluster because

of contig breaks caused by repetitive sequences. Therefore, we were not able to link the Fic proteins to the T4SS of the analyzed regions. Furthermore, Fic-domain containing proteins might be found in any integrated element [17] and thus their presence or absence is not always related to a T4SS.

iii. Evolution and transfer of T4SSs

An interesting finding was the sequence conservation of the T4SS encoding genes in different *C. fetus* strains. Genomic islands are commonly acquired by horizontal gene transfer, followed by island evolution via genetic rearrangements, gene loss, mutations or acquisition of other mobile genetic elements [25]. The genomic islands containing the T4SS regions are inserted at different chromosomal locations, and gene loss and rearrangements are observed in the GIs of different strains, but the T4SS sequences are conserved in the respective *C. fetus* strains. This indicates an evolutionary relationship of the T4SS sequences, but also that T4SS sequences may be conserved to facilitate a functional conjugation system.

Multiple transfer-associated genes are found adjacent to the T4SS regions of cluster 1 and 2 (Table 2). The presence of these transfer-associated genes suggests a plasmid origin for these regions.

The T4SS regions 1B-1D are non-chromosomally located in megaplasmids or ICEs. These regions contained T4SS proteins that were highly homologous to those present in the chromosomally-located T4SS regions 1A, 1E and 1F. ICEs and plasmids can be transferred between cells using a T4SS [32]. The high homology of the chromosomally- and plasmid located T4SSs could also indicate that *C. fetus* strains contain a gene shuffling mechanism, with which plasmids might pick up either chromosomal genes or integrate sequence modules from foreign plasmids, as described for *H. pylori* [33].

Phylogenetic analysis of the VirD4 proteins showed that the *C. fetus* VirD4 proteins and the VirD4 proteins of different *Campylobacter* species form three clusters (Fig 3). This suggests that these *C. fetus* T4SS regions did not evolve from the same ancestor, but were acquired from different donors either by plasmid transfer or conjugational recombination to the *C. fetus* chromosomes.

Conclusions

Overall, our study showed that *C. fetus* strains contain at least three distinct regions, wherein T4SSs could be located in a genomic island, on plasmids and both chromosomally as well as in extra-chromosomal elements. The presence or absence of T4SS is not related

to the S-layer serotype or to the geographic origin of strains, but it is shown that the presence of *virD4* and *fic* genes is significantly associated with *C. fetus* subsp. *venerealis.* Furthermore, it is suggested that the pathogenicity of the *C. fetus* strains are not congruent with the *C. fetus* subspecies classification. Phylogenetic analysis of T4SS-encoding regions showed that the gene content of these regions is conserved in all the analysed *C. fetus* strains and showed that the T4SSs were most likely not acquired from a single ancestor but from different donors.

Acknowledgements

We like to thank Dr. John Devenish and Dr. Brian Brooks (Canadian Food Inspection Agency) for providing strains. We thank Nathaniel Simon and Mary Chapman for the generation of Illumina MiSeq reads and we thank James Bono for the generation of PacBio RS reads.

Author Contri utions

LG, WM, BD, JW conceived this study. LG, EY and KF performed the experiments. LG, WM, BD, AZ, JW analyzed the data. LG wrote the manuscript. WM, BD, GG, AZ, JW were involved in discussions of the work and manuscript revision. All authors approved the final version.

Supplemental data

S1 Fig. Phylogenetic analysis of Fic-encoding sequences. The scale bar represents the mean number of nucleotide substitutions per site.
Supplemental figure S1 can be found online at https://goo.gl/0QhRyn

S2 Fig. BRIG alignment of chromosomally-located T4SS regions 1A, 1E and 2A, and fic genes. Shown are the locations of the chromosomal T4SS regions 1A, 1E and 2A in the reference genomes, using strain 84-112 or strain 01/165 as reference. Strain 04/554 is used as reference to show the location of a genomic island with four fic-encoding sequences.
Supplemental figure S2 can be found online at https://goo.gl/TTFR3f

S3 Table. The positions of the *fic* genes of the in this study analysed sequences.
Supplemental table S3 can be found online at https://goo.gl/vSbxpR

References

1. Véron, M., Chatelain,R. Taxonomy study of the genus *Campylobacter* Sebald and Verón and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Verón. Int. J. Sys. Bacteriol. 1973;23: 122-134.

2. Fitzgerald C, Tu ZC, Patrick M, Stiles T, Lawson AJ, Santovenia M, et al. *Campylobacter fetus* subsp. *testudinum* subsp. nov., isolated from humans and reptiles. Int. J. Syst. Evol. Microbiol. 2014;64: 2944-2948.

3. Wagenaar JA, van Bergen MA, Blaser MJ, Tauxe RV, Newell DG, van Putten JP. *Campylobacter fetus* infections in humans: exposure and disease. Clin. Infect. Dis. 2014;58: 1579-1586.

4. Gilbert MJ, Kik M, Timmerman AJ, Severs TT, Kusters JG, Duim B, et al. Occurrence, diversity, and host association of intestinal *Campylobacter*, *Arcobacter*, and *Helicobacter* in reptiles. PLoS One. 2014;9: e101599.

5. On SL, Harrington CS. Evaluation of numerical analysis of PFGE-DNA profiles for differentiating *Campylobacter fetus* subspecies by comparison with phenotypic, PCR and 16S rDNA sequencing methods. J. Appl. Microbiol. 2001;90: 285-93.

6. van Bergen MAP, Dingle KE, Maiden MC, Newell DG, van der Graaf-Van Bloois L, van Putten JP, et al. Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. J. Clin. Microbiol. 2005;43: 5888-98.

7. Thompson SA, Blaser MJ. Pathogenesis of *Campylobacter fetus* infections. In: Nachamkin I, Blaser MJ, editors. *Campylobacter*. Washington, DC: ASM Press; 2000: 321-347.

8. Dekeyser J. *Campylobacter* infections in man and animals. J.P. Butzler (ed) CRC Press Inc. Florida. 1984: 181-191.

9. Gorkiewicz G, Kienesberger S, Schober C, Scheicher SR, Gully C, Zechner R, et al. A genomic island defines subspecies-specific virulence features of the host-adapted pathogen *Campylobacter fetus* subsp. *venerealis*. J. Bacteriol. 2010;192: 502-17.

10. Kienesberger S, Schober Trummler C, Fauster A, Lang S, Sprenger H, Gorkiewicz G, et al. Interbacterial macromolecular transfer by the *Campylobacter fetus* subsp. *venerealis* type IV secretion system. J. Bacteriol. 2011;193: 744-758.

11. Worby CA, Mattoo S, Kruger RP, Corbeil LB, Koller A, Mendez JC, et al. The fic domain: regulation of cell signaling by adenylylation. Mol. Cell. 2009;34: 93-103.

12. Blaser MJ, Smith PF, Hopkins JA, Heinzer I, Bryner JH, Wang WL. Pathogenesis of *Campylobacter fetus* infections: serum resistance associated with high-molecular-weight surface proteins. J. Infect. Dis. 1987;155: 696-706.

13. Wesley IV, Bryner JH. Antigenic and restriction enzyme analysis of isolates of *Campylobacter fetus* subsp *venerealis* recovered from persistently infected cattle. Am. J. Vet. Res. 1989;50: 807-813.

14. Pei Z, Blaser MJ. Pathogenesis of *Campylobacter fetus* infections. Role of surface array proteins in virulence in a mouse model. J. Clin. Invest. 1990;85: 1036-1043.

15. Blaser MJ. Role of the S-layer proteins of *Campylobacter fetus* in serum-resistance and antigenic variation: a model of bacterial pathogenesis. Am J Med Sci. 1993;306: 325-329.

16. Tu ZC, Eisner W, Kreiswirth BN, Blaser MJ. Genetic divergence of *Campylobacter fetus* strains of mammal and reptile origins. J. Clin. Microbiol. 2005;43: 3334-40.

17. Kienesberger S, Sprenger H, Wolfgruber S, Halwachs B, Thallinger GG, Perez-Perez GI, et al. Comparative Genome Analysis of *Campylobacter fetus* Subspecies Revealed Horizontally Acquired Genetic Elements Important for Virulence and Niche Specificity. PLoS One. 2014;9: e85491.

18. van der Graaf-van Bloois L, Miller WG, Yee E, Bono JL, Rijnsburger M, Campero C, et al. First Closed Genome Sequence of *Campylobacter fetus* subsp. venerealis bv. intermedius. Genome Announc. 2014;2: 10.1128/genomeA.01246-13.

19. van der Graaf-van Bloois L, Miller WG, Yee E, Rijnsburger M, Wagenaar JA, Duim B. Inconsistency of phenotypic and genomic characteristics of *Campylobacter fetus* subspecies requires re-evaluation of current diagnostics. J. Clin. Microbiol. 2014;52: 4183-4188.

20. Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 2001;29: 2607-2618.

21. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014;42: D222-30.

22. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5: 113.

23. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics. 2011;12: 402-2164-12-402.

24. Warren RL, Freeman DJ, Pleasance S, Watson P, Moore RA, Cochrane K, et al. Co-occurrence of anaerobic bacteria in colorectal carcinomas. Microbiome. 2013;1: 16-2618-1-16.

25. Alt-Morbe J, Stryker JL, Fuqua C, Li PL, Farrand SK, Winans SC. The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. J. Bacteriol. 1996;178: 4248-4257.

26. Pearson BM, Rokney A, Crossman LC, Miller WG, Wain J, van Vliet AH. Complete Genome Sequence of the *Campylobacter coli* Clinical Isolate 15-537360. Genome Announc. 2013;1: 10.1128/genomeA.01056-13.

27. Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EP, de la Cruz F. Mobility of plasmids. Microbiol. Mol. Biol. Rev. 2010;74: 434-452.

28. Abril C, Brodard I, Perreten V. Two novel antibiotic resistance genes, tet(44) and ant(6)-Ib, are located within a transferable pathogenicity island in *Campylobacter fetus* subsp. *fetus*. Antimicrob. Agents Chemother. 2010;54: 3052-5.

29. Moolhuijzen PM, Lew-Tabor AE, Wlodek BM, Aguero FG, Comerci DJ, Ugalde RA, et al. Genomic analysis of *Campylobacter fetus* subspecies: identification of candidate virulence determinants and diagnostic assay targets. BMC Microbiol. 2009;9: 86.

30. Censini S, Lange C, Xiang Z, Crabtree JE, Ghiara P, Borodovsky M, et al. *Cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. Proc. Natl. Acad. Sci. U. S. A. 1996;93: 14648-14653.

31. Tegtmeyer N, Wessler S, Backert S. Role of the cag-pathogenicity island encoded type IV secretion system in *Helicobacter pylori* pathogenesis. FEBS J. 2011;278: 1190-1202.

32. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol. Rev. 2009;33: 376-393.

33. Hofreuter D, Haas R. Characterization of two cryptic *Helicobacter pylori* plasmids: a putative source for horizontal gene transfer and gene shuffling. J. Bacteriol. 2002;184: 2755-2766.

Chapter 7

Comparative genomics for development of *Campylobacter fetus* su species specific PCR assays

Linda van der Graaf – van Bloois[a,b], Aldert L. Zomer[a,b], Birgitta Duim[a,b], Jaap A. Wagenaar[a,b,c]

[a] Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands
[b] WHO Collaborating Center for Campylobacter / OIE Reference Laboratory for Campylobacteriosis
[c] Central Veterinary Institute of Wageningen UR, Lelystad, The Netherlands

In preparation

A stract

Bovine Genital Campylobacteriosis (BGC) is caused by *Campylobacter fetus* subsp. *venerealis* and is a notifiable disease to the OIE (World Organisation for Animal Health). For an effective BGC control program, the reliable differentiation of *Campylobacter fetus* subsp. *venerealis* (Cfv) from the closely related *Campylobacter fetus* subsp. *fetus* (Cff) is required. However, the available molecular *C. fetus* subspecies identification assays lack sensitivity and specificity to differentiate *C. fetus* isolates based on their phenotypic or genotypic differences. Furthermore, the current biochemical subspecies identification is not fully congruent with the genomic differentiation of *C. fetus* strains. In this study, the genome sequences of 41 *C. fetus* strains were analyzed with the large-scale BLAST score ratio (LS-BSR) pipeline to identify Cff and Cfv specific sequences. With this analysis, a gene encoding a putative N-acetyltransferase (NAT) was identified, which contained a 6 bp Cff-specific sequence, which was absent in Cfv strains. This sequence was used for the development of PCR assays to differentiate Cff and Cfv strains. The *C. fetus* subspecies differentiation of the developed NAT PCR assays was in full congruence with the genomic classification of strains and these assays are recommended for molecular differentiation of *C. fetus* subspecies in BGC control programs.

Introduction

Bovine Genital Campylobacteriosis (BGC) is characterized by abortion and infertility in cattle and caused by *Campylobacter fetus* subsp. *venerealis* (Cfv) (1). *Campylobacter fetus* comprises two other subspecies; *Campylobacter fetus* subsp. *testudinum* (Cft) and *Campylobacter fetus* subsp. *fetus* (Cff). The *C. fetus* subspecies show a host and niche association: Cft is reptile-associated (2) and both Cff and Cfv are mammal-associated (1). Cfv is restricted to the genital tract of cattle and includes a biochemical variant, designated Cfv biovar intermedius (Cfvi) (3). Cff can be isolated from a variety of different hosts (1) and cause sporadic abortion in cattle and sheep (4) and infections in humans (5).

BGC is notifiable to the OIE (World Organisation for Animal Health) and an effective BGC control program requires the reliable differentiation of Cfv from Cff when using bacteriological culturing. The original described phenotypic methods to identify Cff, Cfv and Cfvi strains are the 1% glycine tolerance test and H_2S production test; Cff is tolerant to 1% glycine and able to produce H_2S, Cfv is not tolerant to 1% glycine and not able to produce H_2S and Cfvi is not 1% glycine tolerant (like Cfv), but able to produce H_2S (like Cff) (3, 6, 7). Multiple molecular methods for *C. fetus* subspecies identification have been described, but several of these methods are laborious and not suitable for standard routine diagnostic laboratories (8). Furthermore, biochemical phenotyes (8, 9) were not consistent with the genotypes defined by multi locus sequence typing and whole genome sequencing (10, 11) of the *C. fetus* strains.

Concerns on the reliability of biochemical tests have arisen as these assays have shown a poor reproducibility (9) and the phenotypes showed inconsistentcy with the genotypes of the strains (12). This emphasizes the need for a reliable molecular *C. fetus* subspecies identification method, which can be implemented in the BGC control programs.

The recent published *C. fetus* genome sequences (12-14) allow comparative genomics to trace Cff and Cfv specific sequences. In this study, we analyzed the mammal-associated *C. fetus* genomes to find specific sequences for development of Cff and Cfv specific PCR assays. The in this study developed PCR assays were able to identify Cff and Cfv isolates in full congruence with the genomic differentiation of subspecies and are recommended for use in the BGC control programs.

Materials and methods

Genome analysis using LS-BSR pipeline

The genome sequences of 41 *C. fetus* strains were analyzed with the large- scale BLAST score ratio (LS-BSR) pipeline (15) to search for Cff and Cfv specific sequences. Strain information and accession numbers are presented in Supplemental Table 1. The genome sequences are online available from NCBI GenBank and the European Nucleotide Archive (ENA) (16).

Development of PCR assays for C. fetus subspecies identification

With Primer3 (17), Cff and Cfv specific primers and Taqman probes were developed on the sequences of a putative N-acetyltransferase (NAT) (Cff8240_1463) encoding gene. The sequences of the developed primers and probes targeting the NAT gene were checked for specificity with a BLAST against the NCBI GenBank database and with a local BLAST against the available sequence reads of 141 *C. fetus* genomes in project PRJEB8721 (European Nucleotide Archive).

For the conventional NAT PCR assay, two forward primers were designed, one specific for Cff and one specific for Cfv, and a single reverse primer for both subspecies (Table 1). The NAT PCR assay consists of two separate PCR reactions; one for Cff identification with primers Nat_Cff_F and Nat_R and one for Cfv identification with primers Nat_Cfv_F and Nat_R. PCR assays were performed in volumes of 20 µl in an Abi 2720 Thermo Cycler (Applied Biosystems), using the following thermal cycler protocol: 95˚C for 3 min, followed by 30 cycli of 95˚C for 30 sec, 58˚C for 30 sec and 72˚C for 1 min and extended with 72˚C for 5 min. Each PCR reaction contained 10 µl GoTaq Green Master Mix (Promega), 1 µl of 10 pmol/µl forward primer, 1 µl of 10 pmol/µl reverse primer, 2 µl of

a le NAT PCR primers and probes for Cff and Cfv identification

either DNA template or heat-lysed cells, adjusted to 20 µl with nuclease free-water. PCR fragments were visualized on a 2% agarose gel stained with Midori Green Advanced DNA staining (GC Tech) and positive results were indicated by a fragment of 241 bp for Cff and 235 bp for Cfv.

A real-time NAT PCR assay was developed in a multiplex format, with forward primer Nat_RT_F and reverse primer Nat_RT_R, and two probes; Nat_Cff_P specific for Cff and Nat_Cfv_P specific for Cfv, which can be combined in one reaction using different fluorescent labels (Table 1). NAT real-time PCR assays were performed in volumes of 20 µl on a LightCycler 480 System (Roche), using the following thermal cycler protocol: 95˚C for 10 sec, followed by 40 cycli of 95˚C for 10 sec, 54˚C for 10 sec and 72˚C for 20 sec and extended with one step of 40˚C for 10 sec. Each PCR reaction contained 10 µl 2x LightCycler 480 Probes Master Mix (Roche), 1 µl of 12 pmol/µl for each primer, 0.5 µl of 8 pmol/µl for each probe, 2 µl of either purified DNA or heat-lysed cells, and adjusted to 20 µl with nuclease free-water. A color compensation analysis was performed to correct crosstalk between the detection channels for the fluorescent dyes FAM and VIC.

Bacterial strains for PCR validation

The PCR assays were tested with a diverse set of 156 *C. fetus* strains and 15 non-fetus *Campylobacter* spp. strains (Supplemental Table 1). The strains were grown on heart-infusion agar supplemented with 5% sheep blood (Biotrading, Mijdrecht, the Netherlands) for two days under microaerobic conditions (6% O_2, 7% CO_2, 7% H_2, 80% N_2, (Anoxomat, Mart Microbiology, Lichtenvoorde, the Netherlands)) and chromosomal DNA was isolated with the Gentra PureGene DNA isolation Kit (Qiagen). Boiled cell lysates were made by suspending 1 colony in 500 µl molecular grade water and subsequently heating at 95˚C for 10 minutes. The *C. fetus* strains were phenotypically identified with the 1% glycine tolerance test and H_2S production in medium with 0.02% cysteine-HCl test as described before (3). The subspecies of the strains were identified with Multi Locus Sequence Typing (MLST) (9) and Amplified Fragment Length Polymorphism (AFLP) (18).

Results

LS-BRS analysis for Cff and Cfv specific sequences

The LS-BSR analysis of 41 *C. fetus* genomes revealed sequences that could be developed for specific PCR detection of *C. fetus* subspecies. These sequences were located in a gene encoding a putative N-acetyltransferase (NAT) (Cff8240_1463), in a gene encoding a superoxide dismutase (*sodC*) (CFF8240-1330) and in multiple type IV secretion system

Figure Location of NAT PCR primers and probes on Cff and Cfv specific target sequences.

(T4SS) genes.

Multiple T4SS encoding genes were found with the LS-BRS analysis as specifically present in Cfv strains. However, T4SS encoding genes are not Cfv-specific as they can be present in Cff strains (19) and as such are unsuitable as targets for *C. fetus* subspecies identification.

The product of the *sodC* gene, coding for sodium oxide dismutase is involved in the protection of the bacteria against oxidative stress (20). This gene contained a 12 bp Cff-specific sequence, which was absent in the genotypically identified Cfv strains. The effect of this 12 bp specific sequence is unknown. As the gene contained repeat sequences flanking the 12 bp Cff-specific sequence, it was not possible to develop Cff and Cfv specific primers. Furthermore, *C. hyointestinalis lawsonii* strains LMG 14434 and NCTC 12901 contained the 12 bp Cff-specific sequence of the *sodC* gene which would result in false positives. In consequence of these results, we decided to exclude the *sodC* gene from further analysis.

The *C. fetus* NAT gene has a 6 bp insertion/deletion, which was found inserted in all genotypic identified Cff strains and was lacking in Cfv strains. Due to this 6 bp deletion in Cfv strains, there is no shift in the codon reading frame, but the strains are missing two amino acids, aspartate and glutamate, in the final protein. It is unknown if this deletion influences the functionality of the NAT protein.

Development NAT primers and probes

The position of the Cff and Cfv specific primers and probes of the NAT gene are shown in Figure 1. The primer and probe sequences (Table 1) were checked against a large set of 141 online available C. fetus genomes (ENA PRJEB8721). The NAT Cfv-specific primers and probe fitted 100% on all 61 Cfv genomes, the NAT Cff-specific primers fitted 100% on all

80 Cff genomes and the NAT Cff-specific real-time probe fitted on 79 of the 80 Cff strains. Only one Cff strain ERR976359 had one point mutation on position 543bp of the NAT gene, containing a thymine instead of a cytosine, which is located 5bp from the start of the Cff-specific probe. Since only the genome sequence and not the bacterial strain was available, it is unknown if this is a true mismatch or sequence mistake. If it is a true mismatch, this difference would result in a ΔTm of 2˚C and it is unknown if this would affect the PCR results.

NAT PCR-assays

The results of both the conventional and real-time NAT PCR assays with 156 *C. fetus* strains (96 Cff and 60 Cfv) and 15 other *Campylobacter* spp. strains are shown in Supplemental Table 2. The NAT PCR assays, both conventional and real-time, showed 100% specificity to identify *C. fetus* and the identification of Cff and Cfv was fully consistent with the genotypes strains. For all NAT PCR assays, the results using purified DNA were identical to the NAT PCR results using boiled cell lysates (data not shown).

Discussion

To prevent spread of BGC in cattle between countries and improve animal health, strict trade regulations have been prescribed by the OIE (OIE, 2010). Crucial in this approach is the establishment of reliable diagnostics that could be used in BGC control programs, also in resource poor countries. Two biochemical assays, tolerance to 1% glycine and production of H_2S in cysteine-rich medium, are still the prescribed methods to identify Cff, Cfv and Cfvi strains (21). However, the 1% glycine tolerance test has poor reproducibility, and cannot always make a reliable differentiation between Cff and Cfv strains (9). Cff and Cfv strains can be genotypically differentiated with whole genome sequence analysis based on their core genes and core genome SNPs, but these genotypes were inconsistent with the biochemical identification of the strains (12, 22).

In this study, we developed PCR assays, both conventional and real-time, which identify Cff and Cfv strains reliably and can easily be applied in BGC control programs. The developed PCR assays performed perfectly to identify Cff and Cfv in full congruence with the genotypes of the tested strains. Fifteen strains representing the non-fetus *Campylobacter* species that occasionally are isolated from bovine or ovine samples were included to verify specificity of the PCR assays. These 15 strains were all negative in the NAT PCR assays, showing 100% specificity for *C. fetus* identification. The target of the developed PCR assays is a putative N-acetyltransferase, currently annotated as an

aminoglycoside 3-N-acetyltransferase in *C. fetus*. This protein matched with Pfam family antibiotic_NAT (PF02522), suggesting that this protein is involved in antimicrobial resistance. However, the association of this protein with antimicrobial resistance is unlikely, since it displays 57% homology with a *C. jejuni* protein coded by Cj1298, which is involved in protein glycosylation (23). Therefore, we named this protein as a putative N-acetyltransferase in this study.

The *C. fetus* subspecies identification of the PCR assays is congruent with the genotypic identification of Cff and Cfv strains, but it is currently unknown if this differentiation is also congruent with the potential pathogenicity of *C. fetus* strains (19). For the control of *C. fetus*-induced fertility problems, one would prefer to differentiate strains based on their potential pathogenicity. The factors that have been suggested to be associated with *C. fetus* virulence are the lipopolysaccharide determinants of strains (1) or the presence of T4SSs (19). To confirm how pathogenicity of *C. fetus* is regulated and to pin point the genes behind virulence of *C. fetus*, more research is needed. As both Cff and Cfv have been isolated from bovine abortions, it seems conceivable that both Cff and Cfv strains can be pathogenic (19). As the virulence of *C. fetus* is poorly understood, we preferred to differentiate stains based on their genomic differences as we assume a stronger association between WGS genotype and virulence than biochemical phenotype and virulence. Once differences in virulence of *C. fetus* strains have been identified, a diagnostic PCR can be developed to differentiate virulent *C. fetus* strains from non-virulent *C. fetus* strains.

Conclusion

The in this study developed conventional and real-time PCR assays, targeting a putative N-acetyltransferase (NAT), showed a 100% specificity to identify C. fetus and the identify Cff and Cfv was in full congruence with the genomic classification of the strains. It is recommended to use the NAT PCR assays as molecular method for C. fetus subspecies identification in BGC control programs.

Acknowledgements

We like to thank Dr. Brian Brooks and Dr. John Devenish (Canadian Food Inspection Agency) for providing strains.

Supplemental data

Supplemental Table 1. C. fetus strains and accession numbers used with (LS-BSR) pipeline.
Supplemental Table 1 can be found online at https://goo.gl/3Mub9m

Supplemental Table 2. Conventional and real-time NAT PCR results of 156 *C. fetus* strains and 15 non-fetus *Campylobacter* spp. strains.
Supplemental Table 2 can be found online at https://goo.gl/A4LmO0

References

1. Thompson, SA, Blaser, MJ. 2000. Pathogenesis of *Campylobacter fetus* infections, p. 321-347. *In* Nachamkin, I, Szymanski, CM, Blaser, MJ (eds.), *Campylobacter*. ASM Press, Washington, DC.

2. Fitzgerald, C, Tu, ZC, Patrick, M, Stiles, T, Lawson, AJ, Santovenia, M, Gilbert, MJ, van Bergen, M, Joyce, K, Pruckler, J, Stroika, S, Duim, B, Miller, WG, Loparev, V, Sinnige, JC, Fields, PI, Tauxe, RV, Blaser, MJ, Wagenaar, JA. 2014. *Campylobacter fetus* subsp. *testudinum* subsp. nov., isolated from humans and reptiles. Int. J. Syst. Evol. Microbiol. 64:2944-2948.

3. Véron, M., Chatelain,R. 1973. Taxonomy study of the genus *Campylobacter* Sebald and Verón and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Verón. International Journal of Systematic Bacteriology. 23:122-134.

4. Garcia, MM, Eaglesome, MD, Rigby, C. 1983. Campylobacters important in veterinary medicine. Veterinary bulletin. 53:793-818.

5. Wagenaar, JA, van Bergen, MA, Blaser, MJ, Tauxe, RV, Newell, DG, van Putten, JP. 2014. *Campylobacter fetus* infections in humans: exposure and disease. Clin. Infect. Dis. 58:1579-1586.

6. Florent, A. 1959. Les deux vibriosis génitales; la vibriose due à *V. fetus venerealis* et la vibriose d'origine intestinale due à *V. fetus intestinalis*. Mededelingen der Veeartsenijschool van de RijksUniversiteit te Gent. 1-60.

7. Florent, A. 1963. A propos dese vibrions responsables de la vibriose génitale des bovins et des ovins. Bull. Off. Int. Epizoot. 60:1063-1074.

8. van der Graaf-van Bloois, L, van Bergen, MA, van der Wal, FJ, de Boer, AG, Duim, B, Schmidt, T, Wagenaar, JA. 2013. Evaluation of molecular assays for identification *Campylobacter fetus* species and subspecies and development of a *C. fetus* specific real-time PCR assay. J. Microbiol. Methods. 95:93-97.

9. van Bergen, MAP, Dingle, KE, Maiden, MC, Newell, DG, van der Graaf-Van Bloois, L, van Putten, JP, Wagenaar, JA. 2005. Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. J. Clin. Microbiol. 43:5888-98.

10. Iraola, G, Betancor, L, Calleros, L, Gadea, P, Algorta, G, Galeano, S, Muxi, P, Greif, G, Perez, R. 2015. A rural worker infected with a bovine-prevalent genotype of *Campylobacter fetus* subsp. *fetus* supports zoonotic transmission and inconsistency of MLST and whole-genome typing. Eur. J. Clin. Microbiol. Infect. Dis. 34:1593-1596.

11. van Bergen, MAP, Simons, G, van der Graaf-van Bloois, L, van Putten, JP, Rombout, J, Wesley, I, Wagenaar, JA. 2005. Amplified fragment length polymorphism based identification of genetic markers and novel PCR assay for differentiation of *Campylobacter fetus* subspecies. J. Med. Microbiol. 54:1217-24.

12. van der Graaf-van Bloois, L, Miller, WG, Yee, E, Rijnsburger, M, Wagenaar, JA, Duim, B. 2014. Inconsistency of phenotypic and genomic characteristics of *Campylobacter fetus* subspecies requires re-evaluation of current diagnostics. J. Clin. Microbiol. 52:4183-4188.

13. Kienesberger, S, Sprenger, H, Wolfgruber, S, Halwachs, B, Thallinger, GG, Perez-Perez, GI, Blaser, MJ, Zechner, EL, Gorkiewicz, G. 2014. Comparative genome analysis of *Campylobacter fetus* subspecies revealed horizontally acquired genetic elements important for virulence and niche specificity. PLoS One. 9:e85491.

14. van der Graaf-van Bloois, L, Miller, WG, Yee, E, Bono, JL, Rijnsburger, M, Campero, C, Wagenaar, JA, Duim, B. 2014. First Closed Genome Sequence of *Campylobacter fetus* subsp. *venerealis* bv. intermedius. Genome Announc. 2:10.1128/genomeA.01246-13.

15. Sahl, JW, Caporaso, JG, Rasko, DA, Keim, P. 2014. The large-scale blast score ratio (LS-BSR) pipeline: a method to rapidly compare genetic content between bacterial genomes. PeerJ. 2:e332.

16. Leinonen, R, Akhtar, R, Birney, E, Bower, L, Cerdeno-Tarraga, A, Cheng, Y, Cleland, I, Faruque, N, Goodgame, N, Gibson, R, Hoad, G, Jang, M, Pakseresht, N, Plaister, S, Radhakrishnan, R, Reddy, K, Sobhany, S, Ten Hoopen, P, Vaughan, R, Zalunin, V, Cochrane, G. 2011. The European Nucleotide Archive. Nucleic Acids Res. 39:D28-31.

17. Untergasser, A, Cutcutache, I, Koressaar, T, Ye, J, Faircloth, BC, Remm, M, Rozen, SG. 2012. Primer3-new capabilities and interfaces. Nucleic. Acids Res. 40:e115.

18. Wagenaar, JA, van Bergen, MAP, Newell, DG, Grogono-Thomas, R, Duim, B. 2001. Comparative study using amplified fragment length polymorphism fingerprinting, PCR genotyping, and phenotyping to differentiate *Campylobacter fetus* strains isolated from animals. J. Clin. Microbiol. 39:2283-6.

19. van der Graaf-van Bloois, L, Miller, WG, Yee, E, Gorkiewicz, G, Forbes, KJ, Zomer, AL, Wagenaar, JA, Duim, B. 2016. *Campylobacter fetus* subspecies contain conserved type IV secretion systems on multiple genomic islands and plasmids. PLoS One. 11:e0152832.

20. Scott, MD, Meshnick, SR, Eaton, JW. 1987. Superoxide dismutase-rich bacteria. Paradoxical increase in oxidant toxicity. J. Biol. Chem. 262:3640-3645.

21. OIE. 2012. Bovine Genital Campylobacteriosis, p. 652. *In* Anonymous Manual of Diagnostic Tests and Vaccines for Terrestrial Animals (mammal, birds and bees), 7th ed. Office International des Epizooties, Paris.

22. van der Graaf-van Bloois, L, Duim, B, Miller, WG, Forbes, KJ, Wagenaar, JA, Zomer, AL. Whole genome sequencing analysis indicates recent diversification of mammal-associated *Campylobacter fetus* and implicates genetic factor associated with H_2S production. Submitted.

23. Obhi, RK, Creuzenet, C. 2005. Biochemical characterization of the *Campylobacter jejuni* Cj1294, a novel UDP-4-keto-6-deoxy-GlcNAc aminotransferase that generates UDP-4-amino-4,6-dideoxy-GalNAc. J. Biol. Chem. 280:20902-20908.

Chapter 8

Discussion

Discussion

Campylobacter fetus is an important pathogen, associated with abortion in cattle and sheep (1) and can cause disease in humans (2). Three *C. fetus* subspecies have been described; *C. fetus* subsp. *fetus* (Cff), *C. fetus* subsp. *venerealis* (Cfv) and *C. fetus* subsp. *testudinum* (Cft). There is a generally accepted association between the mammal-associated *C. fetus* subspecies (Cff and Cfv) and their specific host and niche specificities, epidemiological characteristics and clinical features. Cff has a broad host range and can be isolated from both the intestines and the genital tract of cattle. The supposed transmission route of Cff is oral, although venereal transmission is possible and Cff may cause abortions in cattle, whereas Cfv is restricted to the genital tract of cattle, is only venereally transmitted and is associated with infertility and abortions in cattle (1, 3, 4). As such is Cfv described as the causative agent of Bovine Genital Campylobacteriosis (BGC). BGC is a syndrome characterized by infertility and abortion in cattle, resulting in high economic losses. The objective of the BGC control programs is the prevention of bovine infections with bacteria causing BGC. A crucial element of an effective control program is the detection and identification of the causative agent of BGC in bovines and bovine products. Nowadays, the BGC control programs focus by definition only on Cfv and not Cff. The assumption that Cff does not cause infertility has led to the statement that Cff isolates have less significance compared to Cfv, although we know that this assumption is questionable as it has been described that Cff isolates can cause abortions in cattle (5). Therefore, we may have to reconsider the currently applied different approaches for handling Cfv and Cff infections in cattle.

Classification of *C. fetus* su sp *fetus* and *C. fetus* su sp*. venerealis*

The origin of the differentiation between Cff and Cfv (formerly *Vibrio fetus*) was the niche specificity of the isolates (6). Florent concluded in 1959 that the venereal isolates (Cfv) were restricted to the genital tract of cows and the intestinal isolates (Cff) were able to colonize the intestines of cows (6). The intestinal isolates had a low capacity to colonize the genital tract of cows compared to the venereal isolates (6). The biochemical tests prescribed to differentiate the two subspecies were 1% glycine tolerance and H_2S production in cysteine-rich medium (4). In 1962, *C. fetus* strains were isolated that could, besides colonize the intestinal tract, also colonize the genital tract of cows, but were phenotypically different from Cff and Cfv; they were negative with the 1% glycine

tolerance test (like Cfv), but were positive with the H_2S production test (like Cff) (7, 8). In consequence, these isolates were designated as an intermediate group: *C. fetus* subsp. *venerealis* biovar intermedius (Cfvi) and were described as biochemically related to the intestinal type Cff, but able to colonize the genital tract of non-gestating cows (7).

Since the first descriptions of Cff and Cfv in 1959 and Cfvi in 1962, all publications about *C. fetus* are referring to these publications and preserve this subspecies differentiation, which was based on the isolation of strains from different niches. But are these assertions on different subspecies true? Florent already argued that studies were lacking to pinpoint the exact origin, ecology, pathogenicity and capacity to cause infertility of the *C. fetus* isolates (6, 7). Furthermore, Florent described that some cases of enzootic sterility and abortions might have been caused by the intestinal isolates (Cff) (7). These comments showed that the assertions in the original studies were not indisputable and not as strong as most often stated in citations. Nowadays we are still maintaining this biochemical differentiation of the *C. fetus* subspecies and use them as fundamental tenet for the BGC control programs, where Cfv has to be eradicated and Cff is not a target for eradication. However, it is known from several studies that Cff strains do have considerable significance in bovine infections as Cff strains have been isolated from bovine abortions (9); for example, in a study from Argentina 16 of the 26 *C. fetus* strains isolated from bovine aborted fetuses were identified as Cff (5), and in a study from the United Kingdom, 22 *C. fetus* strains were isolated from abortions, of which were 6 Cff strains (10). In our studies, we were not able to associate the clinical characteristics of the *C. fetus* strains with their phenotypic subspecies identification (Chapter 4 and 6). The current phenotypic differentiation and association of subspecies with clinical characteristics of the *C. fetus* strains is in our opinion not convincing. More knowledge is needed to support the rationale of the *C. fetus* subspecies differentiation.

A study from 1992 using Pulsed Field Gel Electrophoresis (PFGE) estimated the average genome size of *C. fetus* strains to be 1.5 Mbp (11), but this was an underestimation, since *C. fetus* genomes have an average size of 1.9 Mbp (Chapter 4). Whole genome sequence analysis has a high resolution and provides thereby more insight in bacterial genomes. Using Next Generation Sequencing (NGS), whole genome sequencing becomes easier and accessible for more laboratories. Until 2010, only the genome of Cff strain 82-40 was completely sequenced using the Sanger sequencing method and is online available. Within the last 6 years, more than 200 *C. fetus* genomes generated with NGS have

become available in online databases, including the closed genomes of four *C. fetus* strains (12) (Chapter 3 and 4). Using whole genome analysis, *C. fetus* strains could be classified according to their genomic characteristics: Multi Locus Sequence Type (MLST), S-layer protein (*sap*) type, presence of insertion sequence elements and presence of a type I restriction modification (R-M) system (Chapter 4). The genomic characteristics were consistent with the phylogenetic division using core genome and single nucleotide polymorphisms (SNPs) analysis of the *C. fetus* strains, except for one recently sequenced Cff strain HI-UY, which has the Cfv-associated MLST ST-4 genotype (13) (Chapter 4 and 5). The genomic characteristics were not completely congruent with the phenotypes of strains, since phenotypically identified Cff strains were divided in multiple phylogenetic clusters, including clusters containing Cfv/Cfvi strains (Chapter 4 and 5).

The Cfvi strains were originally described as an intermediate group, biochemically related to the intestinal type (Cff), but like Cfv able to colonize the genital tract of non-gestating cows (7). Surprisingly, whole genome analysis using SNP analysis and core genome comparisons, classified all Cfvi strains in the same clade as Cfv strains; both belong to MLST ST 4, *sap*-type A, and both contain IS elements (Chapter 4 and 5). Based on these genomic characteristics, Cfv and Cfvi strains cannot be differentiated, which raises the question if Cfvi is truly an intermediate group or that the differentiation of Cfv and Cfvi is only based on a biochemical phenotype, which is possibly associated with the partial deletion of a putative cysteine transporter (Chapter 5). The Cfv strains contain an incomplete putative cysteine transporter, probably explaining the negative results of H_2S production in cysteine-rich medium (Chapter 5). The H_2S-negative Cfv strains have a niche restriction to the genital tract of cattle in contrast to the H_2S-positive Cfvi and Cff strains, which are assumed to be able to colonize the intestines of cattle and are not restricted to the genital tract (7). One may speculate that the Cfv strains have a defect causing the incapability to grow outside of the genital tract, but we were not able to associate at a functional level the partial deletion of the putative cysteine transporter with the niche restriction of these strains (Chapter 5).

Understanding the evolution of the mammal-associated *C. fetus* may help to elucidate the diversification of Cff, Cfv and Cfvi strains. In Chapter 5, we suggest in that both Cfv and Cfvi strains diverged recently from a Cff ancestor. This diversification could be associated with an increased demand for beef and dairy products during the 19th and 20th century (16). It is very likely that during that time of improvements in cattle breeding, the

expansion of the *C. fetus* population resulted in a successful niche-specific clone, now designated as Cfv, which was able to spread very fast among the world with the increased international trade and large-scale movements of cattle.

Whole genome analysis of bacterial genomes gives rise to multiple discussions about the current taxonomy of bacteria, for example with *E. coli* and *Shigella spp*. where it is proposed to classify *Shigella* strains as *E. coli* based on similarities in whole-genome phylogenies (17). The core genome and more recently the pan genome have been proposed as the principle genomic unit defining bacterial species (18, 19). With a *C. fetus* core genome analysis, which included Cff, Cfv and Cft strains, the Cff and Cfv genomes shared >99% homology and provided evidence for the taxonomic classification of Cff and Cfv as one *C. fetus* subspecies, with Cft (sharing 97% homology with Cff/Cfv) as different subspecies (15). However, phylogeny based on SNPs in the core genomes of Cff and Cfv strains divided Cff strains in different clades (Chapter 5). Based on the molecular characterization of *C. fetus*, we cannot justify the classification of the clade with Cfv/Cfvi strains as one subspecies and all the other clades with Cff strains as a different subspecies. The core genome phylogeny of the mammal-associated *C. fetus* subspecies suggests that Cfv/Cfvi strains are not a different subspecies, but more a successful niche specific Cff clone (Chapter 5).

Virulence factors of *C. fetus* su sp *fetus* and *C. fetus* su sp *venerealis*

C. fetus genomes hold significant information on differences in virulence factor content among strains. Several potential virulence factors have been identified with whole genome analysis (20, 21), including genes encoding surface layer proteins, a type IV secretion system (T4SS) and genes involved in adhesion, invasion, motility and DNA protection. The functionality of these factors in *C. fetus* has been established to some extend for the surface layer and T4SS, but is unknown for the other putative virulence factors. The first step to study virulence factors can be to identify the distribution of these factors in the *C. fetus* genomes of the different subspecies.

Surface layer

A major virulence factor of *C. fetus* is the surface layer (S-layer). The functionality of this S-layer is shown by the resistance of *C. fetus* cells to phagocytosis (22, 23), the role of these surface array proteins (*sap*) in ovine abortions (24) and the persistence of *C. fetus* in bovine genital tract carriage conducted by antigenic variation of the surface proteins (25).

The antigenic variation and the role of the S-layer in the evasion of the host immune system have been extensively investigated (26-34). The antigenic variation of the S-layer is limited, due to the restricted number of different surface array proteins (*sap)* homologs (28-30), but appears to work for a sufficient period to enable colonization of *C. fetus* in host tissue (1).

Another virulence-associated characteristic of the S-layer of the mammal-associated *C. fetus* strains is its association with serum and acid resistance. Adjacent to the S-layer region in the geomes, Cff serotype B strains contain a putative GDP-mannose 4,6-dehydratase (*wcbK)*; Cff serotype A strains contain a putative UDP-galactopyranose mutase (*glf*) and both Cfv serotype A and Cff serotype B strains contain a putative maltose O-acetyltranferase (*mat1*), which are possibly involved in the capsule polysaccharide biosynthesis (12). *C. fetus* serotype A strains are not killed by human serum, since the binding of complement factor C3 is inhibited by the surface layer (33, 35). Serum resistant *C. jejuni* strains contain an ortholog of *wcbK* (*dmhA*), which is associated with bacteremia cases in human (36) and involved in the formation of lipopolysaccharides (LPS) and serum resistance (37). The *wcbK* containing *C. fetus* serotype B strains are serum sensitive, showing that the function of *wcbK* in *C. jejuni* is different than in *C. fetus*. In our strain collection, *C. fetus* strains isolated from human cases are serotype A and not serotype B, which can likely be explained by the serum resistance of the serotype A strains (12). Bovine isolates are either serotype A or serotype B and both serotypes can be isolated from feces and abortions (Chapter 4), indicating that the different LPS structures affected by *wcbK* do not influence the niche specificity of *C. fetus* in cattle. The human complement system is different from the system in cattle, as well as the complement-binding site may differ and it can be hypothesized that LPS structures may be associated with the colonization of *C. fetus* in humans, but likely not be associated with the colonization in cattle.

The *wcbK* containing serotype B strains are acid resistant, whereas serotype A strains are acid sensitive (9). It is hypothesized that the presence of *wcbK* is associated with the survival of the bacterium in the acidic milieu of the stomach (12), like described for *H. pylori* in humans (38). Cff are supposed to be mainly transmitted orally by feces or contact with aborted fetuses (3) and Cfv strains are venereally transmitted (3). If the acid resistance of the *C. fetus* strains, needed for transfer through the stomach, is caused by the presence of the *wcbK* gene, it is likely that the intestinal *C. fetus* strains are all

serotype B strains, since serotype A strains lack *wcbK* and are acid sensitive (12). However, we show in our strain collection (used for Chapters 4, 5 and 6) that *C. fetus* isolated from feces can be both serotype A and B, suggesting that serotype A strains have another mechanism to survive the bovine gastrointestinal tract or that the used *in vitro* conditions are different than the *in vivo* environment of the bovine gastrointestinal tract. The serum sensitivity and acid resistance of strains carrying *wcbK* genes supports the idea that this gene could be associated with niche specialization, however, the function of *wcbK* during infection remain to be proven experimentally.

Adherence to host cells

The adherence of *Campylobacter* spp to gut epithelial cells is mediated by multiple adhesins, including a major outer membrane protein (MOMP) (gene *cmp*), jejuni lipoprotein A (gene *jlpA*) (39), putative binding component of an ABC transporter (gene *peb1*) (40) and *Campylobacter* adhesion to fibronectin protein (gene *cadF*) (41). The genes encoding *cmp*, *jlpA* and *peb1* are conserved in several *Campylobacter* spp, including *C. jejuni*, *C. lari*, *C. upsaliensis*, *C. coli* (42) and in all *C. fetus* subspecies (15, 21). Fibronectin is known to enhance *C. fetus* attachment to intestinal epithelial cells (43). It is described that *cadF* is present only in Cft and Cff (15, 21), and not in Cfv (21). The genome analysis in Chapter 5 showed that Cfv genomes certainly harbor *cadF* and that this gene contains one SNP compared to Cff strains. It is unknown if this SNP changes the adherence of CadF protein to host cells. Since Cfv colonize the genital tract and not the intestines, it is likely that adherence mechanisms of Cfv may differ from Cff strains, and is it possible that Cfv genomes contain different adhesins that are currently unidentified.

Invasion

Campylobacter spp contain the *Campylobacter* invasion antigen B encoded by gene *ciaB* (44). CiaB has been shown to be involved in invasion of *C. jejuni* strain F38011 into epithelial cells, but inactivation of this gene had no influence on invasion of *C. jejuni* strain 81-176 (44, 45). This suggests that *C. jejuni* invasion mechanisms may differ among strains (45). All *C. fetus* subspecies (Cff, Cfv and Cft) contain an ortholog of the CiaB protein (15, 20) and the CiaB encoding genes of Cff strains differ 1-7 SNPs compared to the CiaB encoding genes of Cfv strains (Chapter 5). However, it is unknown if these SNPs can affect CiaB, resulting in different invasion mechanisms of *C. fetus* strains.

The locus coding for the S-layer of *C. fetus* contains a region encoding an invasion-associated marker ABC-transporter (*iamABC).* This region is also present in *C. jejuni*

genomes and the sequence varies between invasive and non-invasive *C. jejuni* strains (46). Cft strains isolated from human contain a recombination in the *iamABC* locus, and the *iamABC* locus of the human-associated Cft strains showed 99-100% amino acid homology with Cff and Cfv strains, while the reptile-associated Cft strains show 93-97% homology (11). The *iamABC* locus of Cff and Cfv strains differ by 1-5 SNPs as shown in Chapter 5, indicating that no recombination events have occurred in this region, but it is unknown whether these SNPs influence the functionality of the *iam*ABC-transporter.

Cytolethal distending toxin

Cytolethal distending toxin (cdt) is produced by a variety of gram-negative bacteria. In *Campylobacter* species, cdt causes cell distention and death due to blocking of cell division (47). The *cdt* gene cluster consists of three adjacent genes; *cdtA*, *cdtB* and *cdtC* (48). CdtA en CdtC are responsible for binding of *Campylobacter* cells to the host cells and CdtB is responsible for toxin activity, which has DNAse I activity and causes DNA double-strand breaks (47).

The presence of the *cdt* gene cluster could be associated with the pathogenesis of *C. fetus* strains, since *cdt* gene negative strains were found which were less efficient during adherence and invasion *in vitro* (49). In *C. fetus*, CdtB is conserved, but differences in CdtA en CdtC amino acid sequences are found among *C. fetus* strains (49). This suggests that the toxin activity produced by this gene cluster could act similar in *C. fetus* strains, but that cell adhesions can be different. If differences in amino acid sequences of the CdtA and CdtC proteins are associated with host specificity of *C. fetus* strains can be investigated by comparing differences in the *cdt* gene clusters in strains isolated from the intestines compared to strains isolated from the genital tract. To accomplish a sound comparison, the metadata of the strains has to be well documented; the clinical information of our sequenced strains was too limited to perform such a comparison.

Motility

C. fetus strains are characterized by an active motility that is mediated by a single polar flagellum (4). The flagella of *C. jejuni* and *C. coli* are not just for motility and chemotaxis, but are also associated with adhesion, invasion, protein secretion and biofilm formation (50-52). Biosynthesis of flagella requires the involvement of more than 40 structural and regulatory proteins, including a type III secretion system for flagellar assembly (52). In *C. fetus*, multiple flagellar genes have shown to be conserved among Cff en Cfv strains (21). However, Cff strains share a recombination region in gene *flgE2*, encoding a flagellar hook

protein (Chapter 5). This gene is part of the flagellar secretion apparatus in *C. jejuni* and associated with the secretion of virulence proteins (53).

The genome of *C. jejuni* contains O-linked glycosylation, which is associated with flagellar assembly and function (51). Recently, a large region encoding many O-linked glycosylation-related proteins was found in *C. fetus* genomes (15). As shown in Chapter 5, this glycosylation region is conserved in all Cff and Cfv strains and is present in most Cft strains (15). This region contains multiple hypervariable GC-tracts and is possibly involved in O-linked glycosylation of flagellin (15).

It is hypothesized that the Cfv genome may harbor more flagellar and motility genes than Cff (21), but this was not observed in our genome comparisons (Chapter 4 and 5). Furthermore, the association of flagellar motility with virulence modulation of Cff and Cfv strains was not investigated.

Type IV secretion system

Type IV secretion system (T4SS) genes are possibly involved in conjugative plasmid transfer or the secretion of virulence factors (42, 54, 55). The genomes of Cff and Cfv strains can harbor multiple genomic islands containing T4SSs and filamentation induced by cAMP (Fic) domain-encoding genes (Chapter 6). The Fic domain proteins may disrupt host cell processes and are possibly secreted by the T4SS. In *C. fetus*, T4SS has shown to be functional and involved in intra- and interspecies conjugative DNA transfer (56).

In Chapter 6, we showed that T4SSs sequences of regions inserted in the same chromosomal locations were conserved in *C. fetus* and could be divided in three distinct phylogenetic clusters, whereas the sequences of the *fic* genes were highly diverse and divided into multiple clusters (Chapter 6). Furthermore, we observed that the Cfv genomes contain significant more T4SSs and *fic* genes compared to Cff genomes, but the presence of these genes was not subspecies-specific. An interesting observation was that all genomes of strains isolated from bovine abortions contain a region encoding conjugative transfer (Tra) proteins (Chapter 5), suggesting that this region may play a role in pathogenesis. This Tra region is also found in multiple *C. fetus* strains, both Cff and Cfv, isolated from non-abortions (Chapter 6). If strains carrying these genes are involved in pathogenesis of *C. fetus* infections in bovine, must be established with experimental infection *in vivo*.

Virulence factors associated with C. fetus epidemiology

The main difference between the *C. fetus* subspecies is their described niche specificity. Multiple putative virulence factors can be associated with the niche specificity of the strains. The epidemiology of *C. fetus* can be associated with the different transmission routes of the strains, since Cff strains are supposed to be mainly orally and Cfv strains venereally transmitted (3), resulting in the respectively colonization of the intestines (Cff strains) and genital tract (Cfv strains). This may indicate that Cff and Cfv express different factors to colonize the host. However, it is known that Cff strains can also be venereally transmitted and are able to colonize the genital tract of cattle. Therefore, it can be hypothesized that the virulence of *C. fetus* strains does not depend on the subspecies of the strains, but depends primarily on the ability to colonize the genital tract of cattle and the accompanying venereal route of transmission, which may be mediated by specific genes or gene variants.

Functionality of virulence factors

Many putative virulence factors of *C. fetus* have been identified, however, it is still unknown which genomic characteristics of *C. fetus* strains are causing infertility and abortion in cattle. The functionality of the *C. fetus* surface layer associated genes and T4SSs have been shown *in vitro* (1, 56), however how these genomic characteristics mediate bovine infections is still unknown. Multiple virulence factors can be associated with the host specificity of *C. fetus* strains; the surface layer, *cdt* genes, adhesion and invasion factors. Even restriction-modification systems can be involved in the invasion activities of strains, as suggested for *C. jejuni*, where a sheep abortion clone showed unique adenine methylation profiles (57).

The virulence factors involved in the host specificity of strains can possibly be further investigated by using bovine intestinal and placental cell lines cultures. Currently, an animal model to investigate *C. fetus* infections in cattle is not readily available. The role of the surface layer proteins during Cff infection is investigated using an ovine model of abortion (24). To investigate an ovine abortifacient *C. jejuni* clone, a guinea pig animal model was described (58). The use of this guinea pig animal model could be of potential interest to investigate the virulence of *C. fetus* strains. Future *in vivo* experiments in cattle should reveal which virulence factors are involved in the *C. fetus* bovine pathogenesis, however, initial investigations under *in vitro* conditions are required to reveal the functions of the putative virulence factors of the *C. fetus* strains.

Genome sta ility of *C. fetus* su sp *fetus* and *C. fetus* su sp *venerealis*

Mobile genetic elements

The *C. fetus* genomes contain multiple mobile genetic elements, including insertion sequence elements, genomic islands encoding type IV secretion systems, bacteriophages and plasmids (Chapter 4 and 6). Bacteriophages are vehicles for the lateral or horizontal movement of genes that can increase bacterial fitness (59). It has been demonstrated that genes carried by bacteriophages can play a role in many aspects of bacterial virulence (adhesion, invasion, host evasion, and toxin production) (60). In *C. fetus* genomes, the majority of proteins located in the prophage regions are hypothetical proteins with unknown function, but we cannot rule out possible functions that either directly impact virulence or increase the fitness of the host in a particular environment.

Both Cff and Cfv strains contain plasmids, but we observed in our strain collection that Cff contained no more than one plasmid and Cfv strains contained multiple plasmids (Chapter 3 en 4). The plasmids harbored multiple T4SSs (Chapter 6) and contained multiple plasmid-related genes with unknown function. The T4SSs were also found in the *C. fetus* genomes integrated on genomic islands (Chapter 6).

Multiple insertion sequence elements were integrated in the genomes of Cfv, which contributes to genome modification and may affect gene expression. IS elements are classified by the length and sequence of short imperfect terminal inverted repeats (IRs) and direct target DNA repeats (DRs), but primarily by the transposase type. In the Cfv genomes were at least two different types identified, sharing 54% sequence homology (Chapters 2, 3 and 4). In contrast, no IS element was found in Cff genomes (Chapter 4), until we recently isolated the first Cff strain containing a plasmid carrying an IS element (unpublished data). This observation indicates that IS elements are not restricted to the Cfv genomes, but can be present in Cff strains. This IS element has been proposed as PCR target for subspecies identification, but is obviously not suitable to differentiate Cfv and Cff strains (54).

DNA protection systems

Restriction-Modification systems (RM systems) protect the bacteria against the integration of foreign DNA. Most RM systems can be classified in one of four categories, types I-IV. Classification is based on multiple factors, including the number of genes, methylation and restriction activities, co-factors required for activity, recognition sequence, cleavage site, and mechanism of cleavage (61) and these complex mechanisms can be extensively variable among species and strains.

C. fetus genomes harbor two different RM systems; a type I and type II RM system (Chapter 4 and 5). The RM type I system consists of *hsd* genes (host specificity for DNA), which are also present in *C. jejuni* strains and encode an enzyme that is both a restriction-dependent ATPase and a DNA topoisomerase, catalyzing restriction and modification (62). The *hsd* gene locus is complete in Cfv strains, whereas the Cff serotype A strains lack subunit *hsdS2* and Cff serotype B strains were lacking the complete *hsd* gene locus (Chapter 4). The RM type II system consists of one gene *cjeI,* which may play a role in plasmid transformation of *C. jejuni* (63). Gene *cjeI* is found in all Cff and Cfv genomes, but contains 24 SNPs in Cfv strains compared to Cff strains, indicating a recombination region because of the high SNP density (Chapter 5). It is unknown if this recombination influences the functionality or specificity of this type II R-M system in *C. fetus* strains.

Another prokaryotic defense system is a clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR associated genes (CRISPR-*cas*) system. All Cfv and multiple Cff genomes contain degenerated CRISPR-*cas* elements; consisting of only CRISPR repeats and a leader sequence, which are identical in Cff and Cfv genomes (12). Furthermore, the complete CRISPR-*cas* encoding system, which is found in several Cff strains has not been found in epsilonproteobacteria and seems unique to *C. fetus* (unpublished data).

The genomes of Cfv strains contain more foreign DNA compared to Cff strains, like genomic islands (GIs), phages, IS elements and plasmids. This suggests that the Cfv strains are more susceptible to integration of foreign DNA compared to Cff strains, possibly caused by less or not functional RM systems and the lack of a functional CRISPR-*cas* system. The presence of more foreign DNA in Cfv can be due to a “defect” resulting in an uncontrolled DNA uptake, but may also provide the possibility to acquire factors that may enable prolonged colonization or better transmission.

he challenge of *Campylobacter fetus*

Bovine Genital Campylobacteriosis, caused by *C. fetus* subsp. *venerealis*, is a worldwide problem and is responsible for huge economic losses, especially in low and middle income countries (LMIC). The control and eradication of Cfv is associated with the method of breeding; developing countries where natural breeding of cattle is widely practiced show high rates of Cfv infection compared to industrialized countries where natural breeding is less frequently practiced (9). The use of artificial insemination is considered to be the most effective means of controlling Cfv and has greatly reduced the incidence of Cfv infections (64, 65).

Vaccination

Vaccines to promote immunity against *C. fetus* have been developed for cattle and sheep. Commercial vaccines are available consisting of inactivated *C. fetus* cells, for example for bovine Vibrin® (Pfizer) and for ovine a combined vaccine against *C. fetus* and *C. jejuni* (Hygieia®). Both male and female cattle can be vaccinated against *C. fetus*. Vaccination of heifers and cows hastens the elimination of *C. fetus*, but the costs of vaccination of all female cattle in a herd can be high. Vaccination of bulls can control the spread of infection and will be a good option for areas where *C. fetus* infection is endemic, however, vaccination of bulls is not recommended as the sole measurement of *C. fetus* control in infected herds (66) as effects are limited. An experimental vaccine developed in Argentina containing inactivated whole *C. fetus* cells yielded good protection for heifers exposed to *C. fetus*-positive bulls (67), whereas commercial vaccines were shown to be ineffective (67, 68).

Whole genome sequence data can be used to predict potential vaccine candidate proteins and has been performed with *C. fetus* genomes by Ali *et al.* using the Vaxign software (20). With the Vaxign *C. fetus* analysis, potential candidate factors for attenuation and subunit vaccine development were identified; nucleoside diphosphate kinase (Ndk), T4SS, MOMP, substrate binding proteins CjaA and CjaC, surface array proteins and Cdt proteins (20). As described in Chapter 4, T4SS proteins are diverse and not present in all *C. fetus* strains, suggesting that they are not suitable as vaccine proteins. As discussed above, MOMP, surface array proteins and Cdt proteins are present in all *C. fetus* strains, whereas MOMP and CdtB are conserved between the strains, suggesting that these proteins are potential vaccine candidates to protect against *C. fetus* infection. A natural glycoconjugate protein such as the flagellum or the sap protein could be an interesting avenue of approach, since glycoconjugate vaccines have proven to be very effective in preventing disease (69, 70).

The S-layer proteins are used to vaccinate ewes and this vaccine was effectively protecting the ewes from abortion (71). However, the different types of the sap proteins, described in Chapter 4 and 5, would necessitate a multivalent vaccine containing several of the variants to obtain sufficient coverage. The S-layer proteins are probably glycosylated (15) and the attachment of the glycan chains to the S-layer proteins is particularly useful for the construction of vaccines based on these proteins (72). The putative glycosylation of the *C. fetus* S-layer encourage further investigation of the S-layer, which will be advantageous for vaccine development.

Proposal for improvement C. fetus diagnostics

The golden standard method for identification of *C. fetus* subspecies is still phenotypic identification (73). However, the biochemical tests have demonstrated poor reliability and reproducibility (74) and it has been shown that the phenotypes are not consistent with the genomic characteristics of *C. fetus* strains in Chapter 4 and 5. We assume that the supposed differences in clinical and epidemiological features between the subspecies are more prone to be associated with the genomic classification than with the phenotypically differentiation. Therefore it would be preferable to have an fast and reliable molecular method to differentiate Cff and Cfv strains according to the genomic classification. The currently available PCR assays are not fully specific and sensitive for reliable subspecies identification (Chapter 2). Whole genome sequencing can be used for reliable genomic characterization of *C. fetus* strains (Chapter 4 and 5). However, access to NGS is not available for all laboratories, particularly for the ones in developing countries. We performed several genome alignments with core and accessory genomes, to find Cff and Cfv specific targets (Chapter 4, 5, 6 and 7). Cff and Cfv specific SNPs were found (Chapter 5), but the drawbacks of a PCR based on one SNP are that the stability of such SNPs is unknown and that the design of a reliable PCR based on specific SNPs requires a high degree of optimization. Several Cff and Cfv specific genes or sequences were found, but these were not suitable as targets for the PCR, since they were located on mobile elements or contained repeat sequences (Chapter 4, 6 and 7). Finally, we were able to identify a Cff and Cfv specific target sequence, located in a putative N-acetyltransferase encoding gene and developed a PCR assay, which was fully consistent with the phylogenetic WGS genotypes of strains (Chapter 7) and can be used in all laboratories, including the laboratories in developing countries.

The *C. fetus* virulence factors that are associated with BGC are still unknown and it has been shown that both Cfv and Cff strains can be isolated from abortions (Chapter 6). Therefore, BGC control programs should aim to differentiate virulent from non-virulent *C. fetus* strains and further research must be focused on identifying the relevant virulence factors associated with infertility and abortion to implement these as diagnostic tests into the BGC control programs. Until that time, the PCR developed in this thesis can be used to differentiate the two subspecies based on their WGS genotype (Chapter 7). Furthermore, a database with all *C. fetus* WGS data should be built and maintained, including well-documented metadata of the isolates. The database should include the clinical information as well as storage of sequence data of the *C. fetus* strains, possibly the

Bacterial Isolate Genome Sequence Database (BIGSdb) software (75). This database could be used in future to perform extensive genome comparisons to screen the *C. fetus* genomes for relevant virulence factors. Since LMIC have fewer resources for research and diagnostics, it is very likely that most of these countries have no possibility to generate sequence data of their own strains. The enclosure of the WGS data of the *C. fetus* strains from these countries will be the responsibility of the reference laboratories in either the region or in a country outside the region with resources to perform WGS. Global networks of veterinary diagnostic laboratories performing *C. fetus* diagnostics in LMIC should be established and assisted by a laboratory with sequence capacity. In addition, a global public available database should be included to the networks, to share data and allowing comparative genomic analysis of whole genome sequences and the metadata of the *C. fetus* strains.

Conclusion

The main issue of this discussion is the question if it is important to differentiate mammal-associated *C. fetus* subspecies. We showed that the described molecular methods (Chapter 2) as well as the genomic analysis in Chapter 4 and 5 to differentiate strains were not consistent with the phenotypes of strains. An important observation was that the current phenotypic differentiation of Cff and Cfv could not be associated with the virulence of strains (Chapter 4). The genomes of the *C. fetus* strains contained multiple putative virulence factors, like the T4SSs identified in Chapter 6, but we were not able to associate these factors to the clinical characteristics of strains. Our whole genome analysis showed that the Cfv and Cfvi strains diverged relatively recently from a Cff ancestor (Chapter 5), suggesting that Cff, Cfv and Cfvi strains are not different subspecies, but that Cfv and Cfvi appear to be a bovine venereal restricted clone of Cff. This correlates with the hypothesis of this discussion that virulence of *C. fetus* strains possibly depends on the ability of strains to colonize the genital tract of bovine, regardless the subspecies type of the strain. Overall, our results revealed no scientific reason for differentiation of the *C. fetus* strains based on their phenotypes identified with two biochemical tests first described in 1959. We strongly suggest that further research is performed on identifying relevant virulence factors of *C. fetus* strains and to implement these factors as (additional) diagnostic tests into the BGC control programs. Until that time, it is recommended to use the PCR developed in this thesis to differentiate the two *C. fetus* subspecies based on their WGS genotype.

References

1. Thompson, SA, Blaser, MJ. 2000. Pathogenesis of *Campylobacter fetus* infections, p. 321-347. *In* Nachamkin, I, Szymanski, CM, Blaser, MJ (eds.), *Campylobacter*. ASM Press, Washington, DC.

2. Wagenaar, JA, van Bergen, MA, Blaser, MJ, Tauxe, RV, Newell, DG, van Putten, JP. 2014. *Campylobacter fetus* infections in humans: exposure and disease. Clin. Infect. Dis. 58:1579-1586.

3. Garcia, MM, Eaglesome, MD, Rigby, C. 1983. Campylobacters important in veterinary medicine. Veterinary bulletin. 53:793-818.

4. Véron, M., Chatelain,R. 1973. Taxonomy study of the genus *Campylobacter* Sebald and Verón and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Verón. International Journal of Systematic Bacteriology. 23:122-134.

5. Campero, CM, Moore, DP, Odeon, AC, Cipolla, AL, Odriozola, E. 2003. Aetiology of bovine abortion in Argentina. Vet. Res. Commun. 27:359-369.

6. Florent, A. 1959. Les deux vibriosis génitales; la vibriose due à *V. fetus venerealis* et la vibriose d'origine intestinale due à *V. fetus intestinalis*. Mededelingen der Veeartsenijschool van de RijksUniversiteit te Gent. 1-60.

7. Florent, A. 1963. A propos dese vibrions responsables de la vibriose génitale des bovins et des ovins. Bull. Off. Int. Epizoot. 60:1063-1074.

8. Park, RWA, Munro, IB, Melrose, DR, Stewart, DL. 1962. Observations on the ability of two biochemical types of *Vibrio fetus* to proliferate in the genital tract of cattle and their importance with respect to infertility. Br.Vet.J. 118:411.

9. Mshelia, GD, Amin, JD, Woldehiwet, Z, Murray, RD, Egwu, GO. 2010. Epidemiology of bovine venereal campylobacteriosis: geographic distribution and recent advances in molecular diagnostic techniques. Reprod. Domest. Anim. 45:e221-30.

10. Willoughby, K, Nettleton, PF, Quirie, M, Maley, MA, Foster, G, Toszeghy, M, Newell, DG. 2005. A multiplex polymerase chain reaction to detect and differentiate *Campylobacter fetus* subspecies *fetus* and *Campylobacter fetus* -species *venerealis*: use on UK isolates of *C. fetus* and other *Campylobacter spp*. J. Appl. Microbiol. 99:758-66.

11. Salama, SM, Garcia, MM, Taylor, DE. 1992. Differentiation of the subspecies of *Campylobacter fetus* by genomic sizing. Int. J. Syst. Bacteriol. 42:446-50.

12. Kienesberger, S, Sprenger, H, Wolfgruber, S, Halwachs, B, Thallinger, GG, Perez-Perez, GI, Blaser, MJ, Zechner, EL, Gorkiewicz, G. 2014. Comparative genome analysis of *Campylobacter fetus* subspecies revealed horizontally acquired genetic elements important for virulence and niche specificity. PLoS One. 9:e85491.

13. Iraola, G, Betancor, L, Calleros, L, Gadea, P, Algorta, G, Galeano, S, Muxi, P, Greif, G, Perez, R. 2015. A rural worker infected with a bovine-prevalent genotype of *Campylobacter fetus* subsp. *fetus* supports zoonotic transmission and inconsistency of MLST and whole-genome typing. Eur. J. Clin. Microbiol. Infect. Dis. 34:1593-1596.

14. Fitzgerald, C, Tu, ZC, Patrick, M, Stiles, T, Lawson, AJ, Santovenia, M, Gilbert, MJ, van Bergen, M, Joyce, K, Pruckler, J, Stroika, S, Duim, B, Miller, WG, Loparev, V, Sinnige, JC, Fields, PI, Tauxe, RV, Blaser, MJ, Wagenaar, JA. 2014. *Campylobacter fetus* subsp. *testudinum* subsp. nov., isolated from humans and reptiles. Int. J. Syst. Evol. Microbiol. 64:2944-2948.

15. Gilbert, MJ, Miller, WG, Yee, E, Zomer, AL, van der Graaf-van Bloois, L, Fitzgerald, C, Forbes, KJ, Meric, G, Sheppard, SK, Wagenaar, JA, Duim, B. 2016. Comparative genomics of *Campylobacter fetus* from reptiles and mammals reveals divergent evolution in host-associated lineages. Genome biology and evolution. 8:2006--2019.

16. Orozco-terWengel, P, Barbato, M, Nicolazzi, E, Biscarini, F, Milanesi, M, Davies, W, Williams, D, Stella, A, Ajmone-Marsan, P, Bruford, MW. 2015. Revisiting demographic processes in cattle with genome-wide population genetic analysis. Front. Genet. 6:191.

17. Pupo, GM, Lan, R, Reeves, PR. 2000. Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. Proc. Natl. Acad. Sci. U. S. A. 97:10567-10572.

18. Lan, R, Reeves, PR. 2000. Intraspecies variation in bacterial genomes: the need for a species genome concept. Trends Microbiol. 8:396-401.

19. Zhang, Y, Sievert, SM. 2014. Pan-genome analyses identify lineage- and niche-specific markers of evolution and adaptation in *Epsilonproteobacteria*. Front. Microbiol. 5:110.

20. Ali, A, Soares, SC, Santos, AR, Guimaraes, LC, Barbosa, E, Almeida, SS, Abreu, VA, Carneiro, AR, Ramos, RT, Bakhtiar, SM, Hassan, SS, Ussery, DW, On, S, Silva, A, Schneider, MP, Lage, AP, Miyoshi, A, Azevedo, V. 2012. *Campylobacter fetus* subspecies: comparative genomics and prediction of potential virulence targets. Gene. 508:145-156.

21. Moolhuijzen, PM, Lew-Tabor, AE, Wlodek, BM, Aguero, FG, Comerci, DJ, Ugalde, RA, Sanchez, DO, Appels, R, Bellgard, M. 2009. Genomic analysis of *Campylobacter fetus* subspecies: identification of candidate virulence determinants and diagnostic assay targets. BMC Microbiol. 9:86.

22. Corbeil, LB, Schurig, GG, Bier, PJ, Winter, AJ. 1975. Bovine veneral vibriosis: antigenic variation of the bacterium during infection. Infect. Immun. 11:240-244.

23. Corbeil, LB, Corbeil, RR, Winter, AJ. 1975. Bovine venereal vibriosis: activity of inflammatory cells in protective immunity. Am. J. Vet. Res. 36:403-406.

24. Grogono-Thomas, R, Dworkin, J, Blaser, MJ, Newell, DG. 2000. Roles of the surface layer proteins of *Campylobacter fetus* subsp. *fetus* in ovine abortion. Infect. Immun. 68:1687-1691.

25. Schurig, GD, Hall, CE, Burda, K, Corbeil, LB, Duncan, JR, Winter, AJ. 1973. Persistent genital tract infection with *Vibrio fetus intestinalis* associated with serotypic alteration of the infecting strain. Am. J. Vet. Res. 34:1399-1403.

26. Tu, ZC, Wassenaar, TM, Thompson, SA, Blaser, MJ. 2003. Structure and genotypic plasticity of the *Campylobacter fetus sap* locus. Mol. Microbiol. 48:685-698.

27. Tu, ZC, Ray, KC, Thompson, SA, Blaser, MJ. 2001. *Campylobacter fetus* uses multiple loci for DNA inversion within the 5' conserved regions of *sap* homologs. J. Bacteriol. 183:6654-6661.

28. Tu, ZC, Hui, J, Blaser, MJ. 2004. Conservation and diversity of *sap* homologues and their organization among *Campylobacter fetus* isolates. Infect. Immun. 72:1715-1724.

29. Tu, ZC, Gaudreau, C, Blaser, MJ. 2005. Mechanisms underlying *Campylobacter fetus* pathogenesis in humans: surface-layer protein variation in relapsing infections. J. Infect. Dis. 191:2082-9.

30. Tu, ZC, Eisner, W, Kreiswirth, BN, Blaser, MJ. 2005. Genetic divergence of *Campylobacter fetus* strains of mammal and reptile origins. J. Clin. Microbiol. 43:3334-40.

31. Tu, ZC, Dewhirst, FE, Blaser, MJ. 2001. Evidence that the *Campylobacter fetus* sap locus is an ancient genomic constituent with origins before mammals and reptiles diverged. Infect. Immun. 69:2237-2244.

32. Blaser, MJ, Wang, E, Tummuru, MK, Washburn, R, Fujimoto, S, Labigne, A. 1994. High-frequency S-layer protein variation in *Campylobacter fetus* revealed by *sapA* mutagenesis. Mol. Microbiol. 14:453-462.

33. Blaser, MJ, Smith, PF, Repine, JE, Joiner, KA. 1988. Pathogenesis of *Campylobacter fetus* infections. Failure of encapsulated *Campylobacter fetus* to bind C3b explains serum and phagocytosis resistance. J. Clin. Invest. 81:1434-1444.

34. Blaser, MJ. 1993. Role of the S-layer proteins of *Campylobacter fetus* in serum-resistance and antigenic variation: a model of bacterial pathogenesis. Am. J. Med. Sci. 306:325-329.

35. Pei, Z, Blaser, MJ. 1990. Pathogenesis of *Campylobacter fetus* infections. Role of surface array proteins in virulence in a mouse model. J. Clin. Invest. 85:1036-1043.

36. Kivisto, RI, Kovanen, S, Skarp-de Haan, A, Schott, T, Rahkio, M, Rossi, M, Hanninen, ML. 2014. Evolution and comparative genomics of *Campylobacter jejuni* ST-677 clonal complex. Genome Biol. Evol. 6:2424-2438.

37. Blaser, MJ, Perez, GP, Smith, PF, Patton, C, Tenover, FC, Lastovica, AJ, Wang, WI. 1986. Extraintestinal

Campylobacter jejuni and *Campylobacter coli* infections: host factors and strain characteristics. J. Infect. Dis. 153:552-559.

38. McGowan, CC, Necheva, A, Thompson, SA, Cover, TL, Blaser, MJ. 1998. Acid-induced expression of an LPS-associated gene in *Helicobacter pylori*. Mol. Microbiol. 30:19-31.

39. Jin, S, Joe, A, Lynett, J, Hani, EK, Sherman, P, Chan, VL. 2001. JlpA, a novel surface-exposed lipoprotein specific to *Campylobacter jejuni*, mediates adherence to host epithelial cells. Mol. Microbiol. 39:1225-1236.

40. Pei, Z, Burucoa, C, Grignon, B, Baqar, S, Huang, XZ, Kopecko, DJ, Bourgeois, AL, Fauchere, JL, Blaser, MJ. 1998. Mutation in the peb1A locus of *Campylobacter jejuni* reduces interactions with epithelial cells and intestinal colonization of mice. Infect. Immun. 66:938-943.

41. Konkel, ME, Garvis, SG, Tipton, SL, Anderson, DE,Jr, Cieplak, W,Jr. 1997. Identification and molecular cloning of a gene encoding a fibronectin-binding protein (CadF) from *Campylobacter jejuni*. Mol. Microbiol. 24:953-963.

42. 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, Nelson, KE. 2005. Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. PLoS Biol. 3:e15.

43. Graham, LL, Friel, T, Woodman, RL. 2008. Fibronectin enhances *Campylobacter fetus* interaction with extracellular matrix components and INT 407 cells. Can. J. Microbiol. 54:37-47.

44. Konkel, ME, Kim, BJ, Rivera-Amill, V, Garvis, SG. 1999. Bacterial secreted proteins are required for the internaliztion of *Campylobacter jejuni* into cultured mammalian cells. Mol. Microbiol. 32:691-701.

45. Goon, S, Ewing, CP, Lorenzo, M, Pattarini, D, Majam, G, Guerry, P. 2006. A sigma28-regulated nonflagella gene contributes to virulence of *Campylobacter jejuni* 81-176. Infect. Immun. 74:769-772.

46. Carvalho, AC, Ruiz-Palacios, GM, Ramos-Cervantes, P, Cervantes, LE, Jiang, X, Pickering, LK. 2001. Molecular characterization of invasive and noninvasive *Campylobacter jejuni* and *Campylobacter coli* isolates. J. Clin. Microbiol. 39:1353-1359.

47. Yamasakia, S, Asakuraa, M, Tsukamotob, T, Faruquec, S, Reema Debd & T. Ramamurthyd. 1996. Cytolethal distending toxin (*cdt*): genetic diversity, structure and role in diarrheal disease. Toxin review. 25:88.

48. Pickett, CL, Pesci, EC, Cottle, DL, Russell, G, Erdem, AN, Zeytin, H. 1996. Prevalence of cytolethal distending toxin production in *Campylobacter jejuni* and relatedness of *Campylobacter* sp. *cdtB* gene. Infect. Immun. 64:2070-2078.

49. Asakura, M, Samosornsuk, W, Taguchi, M, Kobayashi, K, Misawa, N, Kusumoto, M, Nishimura, K, Matsuhisa, A, Yamasaki, S. 2007. Comparative analysis of cytolethal distending toxin (*cdt*) genes among *Campylobacter jejuni, C. coli* and *C. fetus* strains. Microb. Pathog. 42:174-183.

50. Szymanski, CM, King, M, Haardt, M, Armstrong, GD. 1995. *Campylobacter jejuni* motility and invasion of Caco-2 cells. Infect. Immun. 63:4295-4300.

51. Guerry, P. 2007. *Campylobacter* flagella: not just for motility. Trends Microbiol. 15:456-461.

52. Jagannathan, A, Penn, C. 2005. Motility, p. 331-347. *In* Ketley, JM, Konkel, ME (eds.), *Campylobacter*: molecular and cellular biology.

53. Muller, J, Schulze, F, Muller, W, Hanel, I. 2006. PCR detection of virulence-associated genes in *Campylobacter jejuni* strains with differential ability to invade Caco-2 cells and to colonize the chick gut. Vet. Microbiol. 113:123-129.

54. Abril, C, Vilei, EM, Brodard, I, Burnens, A, Frey, J, Miserez, R. 2007. Discovery of insertion element IS*fe1*: a new tool for *Campylobacter fetus* subspecies differentiation. Clin. Microbiol. Infect. 13:993-1000.

55. Christie, PJ, Atmakuri, K, Krishnamoorthy, V, Jakubowski, S, Cascales, E. 2005. Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu. Rev. Microbiol. 59:451-485.

56. Kienesberger, S, Schober Trummler, C, Fauster, A, Lang, S, Sprenger, H, Gorkiewicz, G, Zechner, EL. 2011. Interbacterial macromolecular transfer by the *Campylobacter fetus* subsp. *venerealis* type IV secretion system. J.

Bacteriol. 193:744-758.

57. Skarp, CP, Akinrinade, O, Nilsson, AJ, Ellstrom, P, Myllykangas, S, Rautelin, H. 2015. Comparative genomics and genome biology of invasive *Campylobacter jejuni*. Sci. Rep. 5:17300.

58. Burrough, ER, Sahin, O, Plummer, PJ, Zhang, Q, Yaeger, MJ. 2009. Pathogenicity of an emergent, ovine abortifacient *Campylobacter jejuni* clone orally inoculated into pregnant guinea pigs. Am. J. Vet. Res. 70:1269-1276.

59. Desiere, F, McShan, WM, van Sinderen, D, Ferretti, JJ, Brussow, H. 2001. Comparative genomics reveals close genetic relationships between phages from dairy bacteria and pathogenic *Streptococci*: evolutionary implications for prophage-host interactions. Virology. 288:325-341.

60. Wagner, PL, Waldor, MK. 2002. Bacteriophage control of bacterial virulence. Infect. Immun. 70:3985-3993.

61. O'Loughlin, JL, Eucker, TP, Chavez, JD, Samuelson, DR, Neal-McKinney, J, Gourley, CR, Bruce, JE, Konkel, ME. 2015. Analysis of the *Campylobacter jejuni* genome by SMRT DNA sequencing identifies restriction-modification motifs. PLoS One. 10:e0118533.

62. Miller, WG, Pearson, BM, Wells, JM, Parker, CT, Kapitonov, VV, Mandrell, RE. 2005. Diversity within the *Campylobacter jejuni* type I restriction-modification loci. Microbiology. 151:337-351.

63. Holt, JP, Grant, AJ, Coward, C, Maskell, DJ, Quinlan, JJ. 2012. Identification of Cj1051c as a major determinant for the restriction barrier of *Campylobacter jejuni* strain NCTC11168. Appl. Environ. Microbiol. 78:7841-7848.

64. McGowan, AC, Murray, RD. 1999. Health status of bulls used for natural breeding on farms in south west Scotland. Zentralbl. Veterinarmed. B. 46:311-321.

65. Irons, PC, Schutte, AP, van der Walt, ML, Bishop, GC. 2004. Genital campylobacteriosis in cattle, p. 1359-1468. *In* Coetzer, JAW, Trustin, RC (eds.), Infectious Diseases of Livestock, 2nd ed.

66. Vasquez, LA, Ball, L, Bennett, BW, Rupp, GP, Ellis, R, Olson, JD, Huffman, MH. 1983. Bovine genital campylobacteriosis (vibriosis): vaccination of experimentally infected bulls. Am. J. Vet. Res. 44:1553-1557.

67. Cobo, ER, Morsella, C, Cano, D, Cipolla, A, Campero, CM. 2004. Immunization in heifers with dual vaccines containing *Tritrichomonas foetus* and *Campylobacter fetus* antigens using systemic and mucosal routes. Theriogenology. 62:1367-1382.

68. Cobo, ER, Cipolla, A, Morsella, C, Cano, D, Campero, C. 2003. Effect of two commercial vaccines to *Campylobacter fetus* subspecies on heifers naturally challenged. J. Vet. Med. B Infect. Dis. Vet. Public Health. 50:75-80.

69. De Gregorio, E, Rappuoli, R. 2014. From empiricism to rational design: a personal perspective of the evolution of vaccine development. Nat. Rev. Immunol. 14:505-514.

70. Nothaft, H, Davis, B, Lock, YY, Perez-Munoz, ME, Vinogradov, E, Walter, J, Coros, C, Szymanski, CM. 2016. Engineering the *Campylobacter jejuni* N-glycan to create an effective chicken vaccine. Sci. Rep. 6:26511.

71. Grogono-Thomas, R, Blaser, MJ, Ahmadi, M, Newell, DG. 2003. Role of S-layer protein antigenic diversity in the immune responses of sheep experimentally challenged with *Campylobacter fetus* subsp. *fetus*. Infect. Immun. 71:147-154.

72. Schaffer, C, Messner, P. 2001. Glycobiology of surface layer proteins. Biochimie. 83:591-599.

73. OIE. 2012. Bovine Genital Campylobacteriosis, p. 652. *In* Anonymous Manual of Diagnostic Tests and Vaccines for Terrestrial Animals (mammal, birds and bees), 7th ed. Office International des Epizooties, Paris.

74. van Bergen, MAP, Dingle, KE, Maiden, MC, Newell, DG, van der Graaf-van Bloois, L, van Putten, JP, Wagenaar, JA. 2005. Clonal nature of *Campylobacter fetus* as defined by multilocus sequence typing. J. Clin. Microbiol. 43:5888-98.

75. Jolley, KA, Maiden, MC. 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics. 11:595-2105-11-595.

Summary

Samenvatting

Summary

Introduction

The bacterium *Campylobacter fetus* (*C. fetus*) can cause disease in humans and animals. In humans, *C. fetus* can cause septicemia and intestinal illness, mainly in immunocompromised individuals. Contaminated food, for example unpasteurized milk, is possibly a source of human *C. fetus* infections. *C. fetus* can be isolated from different animal species, including bovine, ovine, goats, birds and cause abortion and infertility in cattle and sheep.

The species *C. fetus* consists of three subspecies; *C. fetus* subsp. *fetus* (Cff), *C. fetus* subsp. *venerealis* (Cfv) and *C. fetus* subsp. *testudinum* (Cft). Cft is recently described and can be isolated from humans and reptiles. This thesis focuses on the mammal-associated *C. fetus*; Cff and Cfv.

In 1959, Cff and Cfv were described as different subspecies based on their biochemical and clinical characteristics; strains that were only isolated from the genital tract of cattle and caused infertility in cows were designated subspecies *venerealis*, and strains that were isolated from the intestines and caused sporadic abortions were designated subspecies *fetus*. The biochemical tests 1% glycine tolerance and H_2S production can differentiate between the two subspecies; Cff strains are 1% glycine tolerant and produce H_2S in cysteine-rich medium, and Cfv strains are not 1% glycine tolerant and cannot produce H_2S. In 1963, *C. fetus* strains were isolated that were phenotypically different and able to colonize both the genital tract and intestines of cattle. These strains were not 1% glycine tolerant, but were able to produce H_2S, and were designated as a biochemical variant of Cfv; *C. fetus* subsp. *venerealis* biovar intermedius (Cfvi).

Cfv is described as the causative agent of Bovine Genital Campylobacteriosis (BGC), a syndrome characterized by infertility and abortion in cattle. Multiple industrialized countries including the Netherlands have a successful BGC control program. These programs require that bovine semen, embryos and living animals are free of Cfv during import and export. However, BGC is still endemic in many countries, mainly in low and middle income countries. Nowadays, the BGC control programs focus only on Cfv and not Cff, while it is described that Cff isolates can also cause abortions in cattle. It is also known that the biochemical tests used to differentiate Cfv and Cfv are difficult to reproduce and not always reliable.

Aim of the thesis

The aim of this thesis was to study the molecular characteristics and differences of the mammal-associated Cff and Cfv strains. Furthermore, we studied if we could associate the molecular characteristics with the virulence of the *C. fetus* strains.

***C. fetus* su species differentiation**

A crucial element in this study was the reliable differentiation and identification of the *C. fetus* strains. As mentioned above, the biochemical tests are not always reliable and molecular methods are preferable. Multiple molecular methods have been described to differentiate Cff and Cfv strains. We tested all available molecular methods on a set of more than 140 *C. fetus* strains in **chapter 2**. None of the available methods was able to make a reliable differentiation between the *C. fetus* strains. To improve the current diagnostics, we used genome comparisons to develop a novel PCR assay to differentiate Cff from Cfv strains (**chapter 7**). This PCR assay can be used in all diagnostic laboratories in the world, even laboratories in developing countries, since no advanced equipment is needed.

Molecular characteristics of *C. fetus* strains

Only one closed Cff genome was available at the start of this study. During the study, the genomes of more than 35 *C. fetus* strains were sequenced and we were able to complete and close the genomes of a Cfvi, Cfv and Cff strain (**chapter 3 and 4**). We determined the molecular characteristics including several potential virulence factors of the selected *C. fetus* strains with these closed genomes as reference (**chapter 4 and 6**).

We performed a phylogenetic analysis on the core genes of the genomes and this analysis divided the strains in two clusters, which were inconsistent with the biochemical characteristics (phenotypes) of the strains (**chapter 4**). In **chapter 5**, we included more Cff strains and analyzed the single nucleotide polymorphisms (SNPs) in the genomes. This phylogenetic SNP analysis divided the *C. fetus* strains in five clusters; one cluster with all Cfv and Cfvi strains like the core genes analysis of **chapter 4**, and four clusters with Cff strains. This division of *C. fetus* genomes was, like the core genes analysis, not consistent with the phenotypes of the strains (**chapter 4 and 5**).

The sequence data allowed us to perform a time analysis with the *C. fetus* strains, and this analysis revealed that Cfv has recently diversified from a Cff ancestor, possibly during the developments in cattle breeding in the 19th century (**chapter 5**), which may have resulted that the niche-specific Cfv strains could spread very fast among the world.

We showed in chapter 4 and 5 that the biochemical characteristics of the strains were not consistent with the molecular characteristics. However, BGC control programs use the biochemical tests to differentiate between Cff and Cfv, while this differentiation cannot be associated with the virulence of the strains. It is more important for the BGC control programs to differentiate between virulent and non-virulent *C. fetus* strains. At the start of this study it was already known that Cfv strains contain type IV secretion systems (T4SS). Many pathogens harbor T4SSs that translocate DNA or bacterial virulence factors. In **chapter 6,** we studied the T4SS in *C. fetus* genomes and revealed that the *C. fetus* genomes can contain multiple T4SSs. The presence of T4SSs were not Cfv-specific, but were also found in Cff strains.

Conclusion

The results of this thesis showed that the current biochemical tests are not consistent with the molecular characteristics of the *C. fetus* strains. Phylogenetic analysis of the genomes showed that the *C. fetus* strains can be divided in at least five different clusters, and showed that Cfv and Cfvi have recently diversified from Cff. This indicated that Cfv and Cfvi are most likely not different subspecies, but are more likely a Cff clone restricted to the genital tract of cattle. Overall, from the genomes no genes were identified that could explain the differentiation of *C. fetus* strains based on 1% glycine tolerance. However, H_2S production could possibly be linked to a putative cysteine transporter. In the *C.fetus* genomes, several potential virulence factors were identified, including T4SSs, but we could not associate these factors with the clinical characteristics of strains. Further research is necessary to unravel the role of the potential virulence factors in *C. fetus* pathogenesis.

Nederlandse samenvatting

Introductie

De bacterie *Campylobacter fetus* (*C. fetus*) kan ziekte veroorzaken bij mens en dier. Bij mensen kan de bacterie bloedvergiftiging en diarree veroorzaken, vooral bij mensen met een verminderd afweersysteem. Besmet voedsel, bijvoorbeeld ongepasteuriseerde melk, kan een bron zijn voor *C. fetus* infecties bij mensen. *C. fetus* komt voor bij veel verschillende diersoorten, zoals rundvee, schapen, geiten, vogels en kan bij rundvee en schapen abortus en onvruchtbaarheid veroorzaken.

C. fetus wordt onderverdeeld in drie ondersoorten (subspecies); *C. fetus* subsp. *fetus* (Cff), *C. fetus* subsp. *venerealis* (Cfv) en *C. fetus* subsp. *testudinum* (Cft). Cft is recent beschreven en Cft bacteriën (stammen) worden gevonden bij mensen en reptielen. Het onderzoek van dit proefschrift richt zich op de *C. fetus* stammen die bij zoogdieren voorkomen: Cff en Cfv.

In 1959 zijn Cff en Cfv stammen als verschillende subspecies beschreven op basis van hun biochemische eigenschappen en de klinische symptomen die zij veroorzaken in rundvee; stammen die alleen voorkwamen in de voortplantingsorganen van rundvee en bij koeien onvruchtbaarheid veroorzaakten werden subspecies *venerealis* genoemd en de stammen die geïsoleerd werden uit de darm en slechts af en toe abortus veroorzaakten, werden subspecies *fetus* genoemd.

De biochemische testen 1% glycine tolerantie en H_2S productie kunnen het onderscheid maken tussen de twee subspecies; Cff stammen groeien in de aanwezigheid van 1% glycine en Cfv stammen niet en Cff stammen kunnen H_2S produceren in cysteïne-rijk medium en Cfv stammen niet. In 1963 werden *C. fetus* stammen gevonden die zowel de voortplantingsorganen als de darm van rundvee konden koloniseren en afwijkend waren in de biochemische testen. Deze stammen waren niet in staat om te groeien bij 1% glycine, maar wel om H_2S te produceren, en werden daarom beschreven als een biochemische variant van Cfv; *C. fetus* subsp. *venerealis* biovar intermedius (Cfvi).

Momenteel wordt Cfv beschouwd als de veroorzaker van Bovine Genital Campylobacteriosis (BGC), een syndroom dat zich kenmerkt door onvruchtbaarheid en abortus in rundvee. Verschillende geïndustrialiseerde landen, waaronder Nederland, hebben een goed georganiseerd beheersingsprogramma voor BGC, waarbij tijdens de import en export, en op KI-stations levende dieren en hun producten getest worden op de aanwezigheid van BGC veroorzakers. In veel landen is BGC nog een groot probleem in de rundveehouderij. De BGC beheersingsprogramma's richten zich nu alleen op Cfv en niet

op Cff, hoewel ook Cff stammen abortus in rundvee kunnen veroorzaken. Verder is bekend dat de biochemische testen die gebruikt worden om onderscheid te maken tussen Cfv en Cff niet altijd betrouwbaar zijn.

Doel van het onderzoek

Over de bacteriële genomen van de Cff en Cfv stammen was bij de start van het onderzoek nog niet veel bekend. Het doel van het onderzoek beschreven in dit proefschrift was om inzicht te krijgen in de eigenschappen van het DNA (moleculaire eigenschappen) van Cff en Cfv stammen. Verder hebben we ook onderzocht of de specifieke moleculaire eigenschappen van de *C. fetus* stammen samenhangen met het vermogen om ziekte te veroorzaken (virulentie).

Het onderscheiden van de C. fetus subspecies

Een cruciaal element in deze studie was het maken van een betrouwbaar onderscheid en het juist identificeren van de *C. fetus* stammen. Zoals hierboven al is aangeven, zijn de biochemische testen niet altijd betrouwbaar. Moleculaire testen voor identificatie hebben de voorkeur en hiervan zijn er meerdere beschreven om Cff en Cfv stammen te onderscheiden. In **hoofdstuk 2** hebben we de beschikbare moleculaire testen op een grote stammenset van meer dan 140 *C. fetus* stammen getest. Geen van de beschikbare moleculaire testen was in staat om een betrouwbaar onderscheid te maken tussen Cff en Cfv stammen. Ter verbetering van de huidige diagnostiek is in dit onderzoek een moleculaire test ontwikkeld die Cff en Cfv stammen op een snelle en betrouwbare wijze kan onderscheiden (**hoofdstuk 7**). Deze test kan gebruikt worden in diagnostische laboratoria die monsters analyseren van dieren; ook in laboratoria in ontwikkelingslanden, aangezien er geen dure en geavanceerde apparatuur voor nodig is.

Moleculaire eigenschappen van de C. fetus stammen

Aan het begin van het onderzoek was het gesloten genoom van één Cff stam beschikbaar. Tijdens dit onderzoek zijn de genomen van meer dan 35 *C. fetus* stammen gesequenced en waren we in staat om het genoom van een Cfvi, Cff en Cfv stam compleet te maken en te sluiten (**hoofdstuk 3 en 4**). Met deze gesloten genomen als referentie hebben we de moleculaire eigenschappen van de stammen bepaald (**hoofdstuk 4**).

Aan de hand van de sequenties konden we een tijdsanalyse uitvoeren op de *C. fetus* stammen, waarbij we hebben kunnen aantonen dat Cfv recentelijk is afgesplitst van Cff,

mogelijk tijdens de intensivering van de veeteelt in de 19e eeuw (**hoofdstuk 5**). Door de internationale handel in rundvee heeft deze ziekteverwekker zich wereldwijd kunnen verspreiden.

De onderlinge verwantschap van de Cff en Cfv genomen hebben we bepaald met behulp van zogeheten fylogenetische analyses. Eerst is bekeken welke genen in alle stammen aanwezig zijn (core genen) en vervolgens hebben we deze core genen onderling vergeleken (**hoofdstuk 4**). De fylogenetische analyse van de core genen deelde de *C. fetus* stammen in twee verschillende groepen (clusters) in, maar deze clusters kwamen niet overeen met indeling op basis van de biochemische eigenschappen van de *C. fetus* stammen. In **hoofdstuk 5** hebben we meer Cff stammen geanalyseerd en niet alleen naar de core genen gekeken, maar naar alle nucleotiden die verschilden (SNPs) in het core genoom van de stammen. Deze fylogenetische analyse deelde de *C. fetus* stammen in vijf verschillende clusters in; één cluster met alle Cfv en Cfvi stammen, zoals bij de analyse met de core genomen (hoofdstuk 4) en vier clusters met Cff stammen. De indeling van deze clusters kwam, net als de core genoom analyse, niet overeen met de indeling op basis van de biochemische eigenschappen van de stammen (**hoofdstuk 4 en 5**).

Voor de BGC beheersingsprogramma's worden de biochemische testen echter wel gebruikt om *C. fetus* stammen te onderscheiden, terwijl dit onderscheid niet altijd samengaat met het vermogen van de stammen om ziekte te verwekken. Voor de BGC beheersingsprogramma's is het meer van belang om onderscheid te maken tussen *C. fetus* stammen die mogelijk ziekteverwekkend zijn en stammen die dat niet zijn. Aan het begin van ons onderzoek was al beschreven dat Cfv stammen een type IV secretie systeem bevatten. Deze systemen zorgen ervoor dat een bacterie in staat is om DNA of eiwitten uit te scheiden en daardoor mogelijk schadelijk is voor de gastheer. In **hoofdstuk 6** hebben we onderzocht welke *C. fetus* genomen een type IV secretie systeem hebben. Hierbij hebben we aangetoond dat *C. fetus* stammen meerdere en verschillende type IV secretie systemen kunnen bevatten. Verder waren deze systemen niet specifiek voor Cfv stammen, maar vonden we ze ook in Cff stammen. In de toekomst zal met onderzoek, bijvoorbeeld met celkweek of een dierexperiment, moeten worden aangetoond of deze systemen ook functioneel zijn en bijdragen aan het vermogen om ziekte te veroorzaken.

Conclusies

De resultaten in dit onderzoek laten zien dat de indeling van de *C. fetus* stammen op basis van biochemische testen niet overeenkomt met de indeling op basis van moleculaire eigenschappen. De genoomanalyse liet zien dat de *C. fetus* stammen onderverdeeld worden in ten minste vijf verschillende groepen en dat de Cfv en Cfvi stammen recentelijk zijn afgesplitst van Cff. Dit suggereert dat Cfv en Cfvi geen apart subspecies zijn, maar meer een aparte Cff kloon, die beperkt is tot het voorplantingsorgaan van rundvee. In dit onderzoek hebben wij geen moleculaire onderbouwing gevonden voor de biochemische eigenschappen op grond waarvan *C. fetus* onderscheiden worden. Voor de BGC beheersingsprogramma's is het veel meer van belang om een onderscheid te maken tussen virulente en niet-virulente stammen dan stammen te onderscheiden op basis van biochemische eigenschappen. We hebben diverse virulentie factoren gevonden in de *C. fetus* genomen, onder andere type IV secretie systemen, maar we konden deze factoren niet in verband brengen met het vermogen van stammen om ziekte te veroorzaken. Nader onderzoek naar de functionaliteit van de *C. fetus* virulentiefactoren is noodzakelijk.

Curriculum Vitae

Curriculum Vitae

De auteur van dit proefschrift, Linda van der Graaf-van Bloois, werd geboren op 25 oktober 1981 te Rotterdam. Na het behalen van het atheneumdiploma in 2003 (Wartburg College, locatie Revius, Rotterdam) vervolgde zij haar opleiding bij het Hoger Laboratorium Onderwijs aan de Hogeschool Rotterdam. De afstudeerstage werd uitgevoerd binnen het cluster Ketenkwaliteit en Zoönosen van de toenmalige Animal Sciences Group van Wageningen UR te Lelystad. In 2003 haalde zij haar diploma en werd vervolgens assistent onderzoeker binnen diezelfde afdeling van de Animal Sciences Group te Lelystad.

In 2011 startte zij haar promotieonderzoek bij de afdeling Klinische Infectiologie van het departement Infectieziekten en Immunologie van de faculteit Diergeneeskunde, Universiteit Utrecht. Onder begeleiding van prof.dr. Jaap Wagenaar, dr. Birgitta Duim en dr. Aldert Zomer is het promotie onderzoek uitgevoerd naar de moleculaire eigenschappen en verschillen van *Campylobacter fetus* subspecies, waarvan de resultaten in dit proefschrift beschreven zijn.

Sinds 2011 is zij tevens betrokken bij het Foodborne Infections Network (GFN) van de World Health Organization (WHO) waarvoor zij als trainer cursussen heeft gegeven in Kenia, Oezbekistan en Thailand. Zij is tevens werkzaam voor het WHO-Collaborating Center for Campylobacter en het OIE-Reference Laboratory for Campylobacteriosis.

List of pu lications

Pu lications - *Campylobacter fetus*

Van der Graaf-van Bloois L., Duim B., Miller W.G., Forbes K.J., Wagenaar J.A., Zomer A.L. 2016. Whole genome sequence analysis indicates recent diversification of mammal-associated *Campylobacter fetus* and implicates a genetic factor associated with H_2S production. BMC Genomics 17:713.

Van der Graaf-van Bloois L , Miller W.G., Yee E., Gorkiewicz G., Forbes K.J., Zomer A.L., Wagenaar J.A., Duim B. 2016. *Campylobacter fetus* subspecies contain conserved type IV secretion systems on multiple genomic islands and plasmids. PLoS One 11 (4)**:**e0152832.

Van der Graaf-van Bloois L , Miller W.G., Yee E., Rijnsburger M., Wagenaar J.A., Duim B. 2014. Inconsistency of phenotypic and genomic characteristics of *Campylobacter fetus* subspecies requires reevaluation of current diagnostics. J. Clin. Microbiol. 52 (12):4183-4188.

Van der Graaf-van Bloois L , Miller W.G., Yee E., Bono J.L., Rijnsburger M., Campero C., Wagenaar J.A., Duim B. 2014. First closed genome sequence of *Campylobacter fetus* subsp. *venerealis* bv. intermedius. Genome Announc. 2:10.1128/genomeA.01246-13.

Van der Graaf-van Bloois L , van Bergen M.A.P, van der Wal F.J., de Boer A.G., Duim B., Schmidt T., Wagenaar J.A. 2013. Evaluation of molecular assays for identification *Campylobacter fetus* species and subspecies and development of a *C. fetus* specific real-time PCR assay. J. Microbiol. Methods 95 (1):93-7.

Spence R.P., Bruce I.R., McFadden A.M.J., Hill F.I., Tisdall D., Humphrey S., **van der Graaf-van Bloois L ,** van Bergen M.A.P., Wagenaar J.A. 2011. Cross-reaction of a Cfv specific Real-Time PCR. Vet. Record 168:131.

Van Bergen M.A.P., **van der Graaf-van Bloois L** , Visser I.J.R., van Putten J.P.M., Wagenaar J.A. 2006. Molecular epidemiology of *Campylobacter fetus* subsp. *fetus* on bovine artificial insemination stations. Vet. Microbiol. 112:65-71.

Van Bergen M.A.P., Dingle K.E., Maiden M.C.J., Newell D.G.**, van der Graaf-van Bloois L** , van Putten J.P.M., Wagenaar J.A. 2005. Clonal nature of *Campylobacter fetus* as defined by Multilocus Sequence Typing. J. Clin. Microbiol. 43:5888-5898.

Van Bergen M.A.P., Simons G., **van der Graaf-van Bloois L** , van Putten J.P.M., Rombout J., Wesley I., Wagenaar J.A. 2005. Amplified Fragment Length Polymorphism based identification of genetic markers and novel PCR assay for differentiation of *Campylobacter fetus* subspecies. J. Med. Microbiol. 54:1217-1224.

Pu lications – *other*

Louwen R., Horst-Kreft D., de Boer A.G., **van der Graaf-van Bloois L** , de Knegt G., Hamersma M., Heikema A.P., Timmers A.R., Jacobs B.C., Wagenaar J.A., Endtz H.P., van der Oost J., Wells J.M., Nieuwehuis E.E.S., van Vliet A.H.M., Willemsen P.T.J., van Baarlen P., van Belkum A. 2013. A novel link between *Campylobacter jejuni* bactiophage defence, virulence and Guillain-Barré syndrome. Eur. J. Clin. Microbiol. Infect. Dis. 32:207-226.

Felkers F.C., **van der Graaf-van Bloois L** , Wagenaar J.A., Westendorp S.T., van Bergen M.A.P., Dwars R.M., Landman W.J.M. 2011. *Enterococcus hirae*-associated endocarditis outbreaks in broiler flocks: clinical and pathological characteristics and molecular epidemiology. Vet. Quart 31:3-17.

Gaasbeek E.J., van der Wal F.J., van Putten J.P.M., de Boer P., **van der Graaf - van Bloois L** , de Boer A.G., Vermaning B.J., Wagenaar J.A. 2009. Functional Characterization of Excision Repair and RecA-Dependent Recombinational DNA Repair in *Campylobacter jejuni.* J. Bacteriol. 191:3785-93.

Gaasbeek E.J., Wagenaar J.A., Guilhabert M.R., Wösten M.M.S.M., van Putten J.P.M., **van der Graaf - van Bloois L** , Parker C.T., van der Wal F.J. 2009. A DNase encoded by integrated element CJIE1 inhibits natural transformation of *Campylobacter jejuni*. J. Bacteriol. 191 (7):2296-2306.

J-Kaszanyitzky É., Jánosi S., Somogyi P., Dan A., **van de Graaf-van Bloois L** , van Duijkeren E., Wagenaar J.A. 2007. MRSA transmission between cows and humans. Emerg. Infect. Dis. 13:630-632.

Ho T.K., Lipman L.J.A., **van der Graaf- van Bloois L** , van Bergen M.A.P., Gaastra W. 2006. Potential routes of acquisition of *Arcobacter* species by piglets. Vet. Microbiol. 114:123-133