Tesis para obtener el título Doctorado en Ciencias Biológicas Opción Genética

Genómica poblacional aplicada al estudio de *Campylobacter fetus* en humanos y animales de producción



Daniela Costa Duarte 12/11/2024 Tesis para obtener el título: Doctorado en Ciencias Biológicas opción Genética

Genómica poblacional aplicada al estudio de *Campylobacter fetus* en humanos y animales de producción

Daniela Costa Duarte

12 de noviembre de 2024

Presidente del tribunal: Dr. Pablo Smircich

Vocales: Dra. Luisa Berná Dra. Vanesa Amarelle

Dr. Gregorio Iraola Supervisores: Dr. Rúben Pérez Pérez Dra. Lucía Calleros



Facultad de Ciencias Sección Genética Evolutiva PEDECIBA Biología



Institut Pasteur de Montevideo Laboratorio de Genómica Microbiana y Unidad de Bioinformática

Daniela Costa Duarte

Genómica poblacional aplicada al estudio de Campylobacter fetus en humanos y animales de producción

Tesis para obtener el título de Doctorado en Ciencias Biológicas opción Genética, 12 de noviembre de 2024

Tribunal: Dr. Pablo Smircich, Dra. Luisa Berná, Dra. Vanesa Amarelle Tutor: Dr. Gregorio Iraola; Cotutores: Dr. Rúben Pérez Pérez, Dra. Lucía Calleros

Universidad de la República

Doctorado en Ciencias Biológicas Opción Genética PEDECIBA Biología Facultad de Ciencias

Índice general

1	Justificación y antecedentes			1
	1.1	Antecedentes		
	1.2	Hipóte	esis y objetivos	4
	1.3	Estruc	etura de este trabajo	4
2	Introducción		ón	6
	2.1	2.1 Sobre <i>Campylobacter</i>		6
		2.1.1	Taxonomía	6
		2.1.2	Principales características	8
		2.1.3	Un poco de historia	0
		2.1.4	Especies de Campylobacter	1
		2.1.5	Genómica aplicada al estudio de Campylobacter 12	2
	2.2	2 Campylobacter fetus		3
		2.2.1	Posición taxonómica 14	4
		2.2.2	Subespecies 1	5
	2.3 La Campylobacteriosis Genital Bovina (CGB)		mpylobacteriosis Genital Bovina (CGB) 1	7
2.3.1 Características		2.3.1	Características	8
		2.3.2	Diagnóstico	9
		2.3.3	Control y prevención de la CGB 2	7
		2.3.4	Incidencia en nuestro país 28	8
		2.3.5	Situación sanitaria y manejo de la producción ganadera	
			en Uruguay	9

		2.3.6	Revisión: Pathogenomics of emergent Campylobacter	
			species	31
3	Obte	ención	de genomas de Campylobacter fetus aisladas en Uru-	
	gua	У		54
	3.1	Introd	ucción	54
		3.1.1	De <i>reads</i> a genomas	55
		3.1.2	De genomas a genes	65
	3.2	Hipóte	esis y objetivos	67
		3.2.1	Hipótesis	67
		3.2.2	Objetivos	68
	3.3	Metod	lología	69
		3.3.1	Antecedentes	69
		3.3.2	Procesamiento de datos crudos de secuenciación masiva	70
		3.3.3	Ensamblaje de los genomas	73
		3.3.4	Post-procesamiento de genomas ensamblados	74
	3.4 Resultados		76	
		3.4.1	Origen de los aislamientos	76
		3.4.2	Métricas de los datos de secuenciación masiva	77
		3.4.3	Preprocesamiento de <i>reads</i> cortos	78
		3.4.4	Ensamblado y control de calidad	79
		3.4.5	Procesamiento de los ensamblados híbridos	81
	3.5	Discus	sión y conclusiones	88
	3.6	Artícu	lo: Complete Genome Sequence of Campylobacter fetus	
		Isolate	ed from a Sheep	94
4	Estu	idio co	mparativo de genomas de <i>Campylobacter fetus subsp</i> .	
	fetu	s y Can	npylobacter fetus subsp. venerealis	96
	4.1	Introd	ucción	96

		4.1.1	Impacto de la disponibilidad de secuencias genómicas	
			en la microbiología	6
		4.1.2	C. fetus: del diagnóstico fenotípico a la genómica com-	
			parativa	8
	4.2	Hipóte	esis y objetivos)1
		4.2.1	Hipótesis)1
		4.2.2	Objetivos)1
	4.3	Metod	lología	3
		4.3.1	Resumen de la estrategia utilizada10	3
		4.3.2	Obtención del Dataset	5
		4.3.3	Caracterización de los genomas	5
		4.3.4	Obtención del genoma <i>core</i>	6
		4.3.5	Marco clonal, recombinaciones y filogenias 10	6
		4.3.6	Anotación de los genomas	17
		4.3.7	Pangenomas	17
		4.3.8	Diseño de cebadores específicos de linaje 10	8
	4.4	Result	ados	0
		4.4.1	Dataset: distribución geográfica de los aislamientos,	
			fuentes, hospederos	0
		4.4.2	Caracterización de los genomas	3
		4.4.3	Estructura poblacional	8
		4.4.4	Distancias pareadas	3
		4.4.5	Recombinación	5
		4.4.6	Búsqueda de blancos moleculares para tipificación12	8
	4.5	Discus	sión	.3
	4.6	Conclu	usiones	0
5	Gen	ómica	poblacional aplicada al estudio de aislamientos de $C_{\rm c}$	
-	fetu	s bovin	10s 15	51
	5.1	Introd	lucción	51

	5.2	2 Hipótesis y Objetivos		160
		5.2.1	Hipótesis1	160
		5.2.2	Objetivos1	160
	5.3	Metod	lología	161
		5.3.1	Set de datos	161
		5.3.2	Obtención del alineamiento del marco clonal y estruc-	
			tura poblacional	161
		5.3.3	Determinación de la antigüedad de los linajes 1	162
		5.3.4	Construcción y análisis del pangenoma	164
		5.3.5	Identificación de factores de virulencia y otros elemen-	
			tos de asociados a la patogenicidad 1	165
		5.3.6	Búsqueda de elementos móviles	166
	5.4	Result	ados	170
		5.4.1	Dataset	170
		5.4.2	Estructura poblacional y distribución geográfica de C.	
			<i>fetus</i> ST4	172
		5.4.3	Rasgos genéticos	179
		5.4.4	Análisis de las regiones recombinantes1	188
		5.4.5	Identificación de factores de virulencia1	190
		5.4.6	Elementos móviles	202
	5.5	Discus	sión	215
		5.5.1	Rasgos genómicos diferenciales entre sublinajes sugie-	
			ren procesos de adaptación	217
	5.6	Conclu	usiones	226
6	Estu	dio de	e una serie de casos clínicos de Campylobacter fetus	
	ocui	ridos	en Montevideo entre 2013 y 2018 2	227
	6.1	Introd	lucción	227
	6.2	Antece	edentes	228
		6.2.1	Descripción de los casos	228

	6.3	Metodología			
		6.3.1	Obtención de secuencias genómicas	230	
		6.3.2	Análisis bioinformáticos	230	
	6.4 Resultados			232	
		6.4.1	Análisis de MLST de los aislamientos	232	
		6.4.2	Predicción del genoma <i>core</i> y el marco clonal	232	
		6.4.3	Filogenias y recombinación	233	
		6.4.4	Distancias entre los aislamientos	233	
	6.5	Discus	sión	236	
	6.6	6.6 Conclusiones			
	6.7	.7 Artículo: Polyclonal Campylobacter fetus infections among un-			
		related	d patients Montevideo Uruguay 2013–2018	239	
7	Consideraciones finales 243			243	
8	Apéndices 250			250	
Bi	Bibliografía 266				

1

Justificación y antecedentes

1.1. Antecedentes

La producción ganadera es una de las actividades económicas más rentables de nuestro país, y dada la marcada orientación de esta actividad hacia la exportación, Uruguay debe cumplir las exigencias del mercado internacional para garantizar la competencia de sus productos. La situación sanitaria de Uruguay en relación a la producción ganadera se considera muy buena, y ha abierto las puertas hacia los mercados más exigentes. Sin embargo, las tasas de procreo han sido históricamente bajas, situándose en el entorno de 63 % en el período 1981 a 2012, lo que se traduce a pérdidas económicas millonarias. Si bien este problema es multifactorial, se considera importante el impacto de las enfermedades infecciosas. Se han registrado casos de éxito en los que la adopción de acciones oportunas en un rodeo, como las descritas en Subsección 2.3.5, que pueden aumentar significativamente la eficiencia reproductiva, con aumentos en índices de procreo superiores al 20 % [1]. Pero en ausencia de datos epidemiológicos actualizados, la adopción de medidas sanitarias a nivel nacional es prácticamente imposible.

En 2015 surge el proyecto "Desarrollo y validación de metodologías para el diagnóstico y control de la campylobacteriosis genital bovina", financiado por el Fondo Sectorial de Salud Animal de la ANII, cuyos responsables fueron

los Dres. Lucía Calleros, Ruben Pérez y Gregorio Iraola. Este trabajo deriva de la ejecución de dicho proyecto.

Uno de los objetivos propuestos en este proyecto consistió en aislar cepas de *Campylobacter. fetus* bovinas pertenecientes a diferentes establecimientos rurales, y secuenciar sus genomas. Fruto de este trabajo, se obtuvieron 11 aislamientos de *C. fetus* de un total de 520 muestras de raspado prepucial bovino analizadas, provenientes de 98 establecimientos rurales, equivalente a una prevalencia de 2,6 % [2, 3].

Para conocer el *status* sanitario del Uruguay es necesario contar con herramientas de diagnóstico fáciles de implementar, sensibles y específicas. La Organización Mundial de Salud Animal (OMSA) recomienda el aislamiento y cultivo de *C. fetus* para el diagnóstico de la Campylobacteriosis genital bovina (CGB), el cual requiere medios de transporte adecuados y una plataforma tecnológica que permita el crecimiento de esta especie en condiciones de microaerofilia, además de personal altamente calificado y específicamente entrenado. Si bien existen varias herramientas de diagnóstico molecular, la baja sensibilidad y/o especificidad reportadas limitan su uso como técnicas de diagnóstico definitivo.

En este escenario resulta interesante estudiar exhaustivamente la diversidad genética de *C. fetus*, partiendo de una base de datos genómicos confiables y diversos, con el fin de identificar potenciales blancos moleculares de diagnóstico. En este sentido, con el fin de incorporar datos genómicos de calidad provenientes de cepas de circulación local, la siguiente etapa del proyecto consistió en obtener los genomas, tanto de los aislamientos recuperados en la fase anterior, como de otras cepas aisladas de bovinos, ovinos y humanos en Uruguay, gracias a la colaboración de diferentes instituciones, tanto públicas como privadas, que cedieron sus aislamientos.

El trabajo plasmado en esta tesis abarca el análisis de los genomas obtenidos en el marco del proyecto antes mencionado.

1.2. Hipótesis y objetivos

En Uruguay, la presencia de *C. fetus* ha sido registrada tanto en el ámbito veterinario como en el clínico, con significativas implicaciones económicas y sanitarias. Desde su detección hace 60 años, ha afectado la producción ganadera y causado pérdidas económicas por la disminución de la fertilidad y los tratamientos veterinarios. Además, la infección en humanos puede llevar a complicaciones graves, aumentando la carga sobre el sistema de salud.

Este trabajo surge bajo la hipótesis de que reconocer la diversidad genética de la especie es esencial para entender la enfermedad y diseñar estrategias de diagnóstico rutinario en la práctica clínica y veterinaria. En base a esta hipótesis se trazan los siguientes objetivos específicos:

- **Generar** una base de datos genómicos de cepas de *C. fetus* asociados a mamíferos que contenga secuencias de diversos orígenes, incluyendo las que circulan en nuestro país.
- **Describir** los genomas recolectados e inferir la estructura poblacional, y su relación con el origen de los aislamientos.
- **Identificar** características genéticas relevantes para la ecología de *C. fetus*, que permitan ser utilizadas para inferir el potencial patogénico y determinar el riesgo clínico o veterinario de una cepa.

1.3. Estructura de este trabajo

El punto de partida del trabajo plasmado en esta tesis fueron los datos de secuenciación de segunda y tercera generación, obtenidos a partir de los aislamientos mencionados en el Capítulo 1. La obtención de estos genomas se describe en el Capítulo 3, titulado "Obtención de genomas de *Campylobacter fetus* aisladas en Uruguay".

En el Capítulo 4, "Estudio comparativo de genomas de *Campylobacter fetus subsp. fetus* y *Campylobacter fetus subsp. venerealis*", se recogen datos disponibles en bases de datos públicas y se analizan en conjunto con los genomas obtenidos en el Capítulo 3 para analizar de forma global la diversidad genética en las cepas de *C. fetus* secuenciadas, con el fin de identificar posibles blancos moleculares de diagnóstico de la CGB. Allí se describe y caracteriza la estructura poblacional de *C. fetus* asociados a mamíferos y su relación en la preferencia a hospedero, subespecie y origen geográfico. Por último se describe el desarrollo de una herramienta basada en PCR con potencial diagnóstico.

La CGB, como su nombre lo indica, es la manifestación clínica de la colonización del tracto genital de *C. fetus* en bovinos, razón por la cual resulta interesante explorar en detalle los genomas de *C. fetus* asociados a estos animales en forma exclusiva. Este trabajo es plasmado en el Capítulo 5, titulado "Genómica poblacional aplicada al estudio de aislamientos de *C. fetus* bovinos" donde se describe la estructura poblacional derivada de estos genomas y su relación con el origen geográfico. Además se caracteriza cada grupo derivado de la estructura poblacional obtenida y se identifican diferenciales relacionados a la movilidad genética y la interacción hospedero-patógeno.

En el Capítulo 6 "Estudio de una serie de casos clínicos de *Campylobacter fetus* ocurridos en Montevideo entre 2013 y 2018" se caracteriza una serie de casos clínicos de *C. fetus* ocurridos en Montevideo.

Por último, el Capítulo 7 provee una mirada global a los principales resultados obtenidos en este trabajo.

Introducción

2.1. Sobre Campylobacter

Campylobacter es un género de bacterias muy grande y diverso, cuyas especies colonizan tanto humanos como otros mamíferos, aves y reptiles. Si bien la especie tipo del género es *C. fetus*, las especies más estudiadas son *C. coli* y *C. jejuni*, debido a que son responsables de la mayor parte de las enteritis en humanos a nivel mundial [4]. Más aún, hasta el 90% de los aislamientos clínicos de *Campylobacter* corresponden a la especie *C. jejuni* [5]. Sin embargo, otras especies diferentes a *coli-jejuni* han sido asociadas a diferentes patologías en humanos, tanto entéricas como sistémicas, pero su prevalencia es subestimada debido al sesgo en el diagnóstico hacia las especies más comunes. Así es que el potencial patogénico de estas especies emergentes de *Campylobacter* es desconocido.

A continuación se describe la composición actual del género *Campylobacter* y su taxonomía, además de sus principales características.

2.1.1. Taxonomía

Campylobacter es un género perteneciente a la clase *Epsilonproteobacteria*, clase *Campylobacterales*, familia *Campylobacteraceae*, siendo *Campylobacter fetus* su especie tipo.

En la actualidad, según el *Listado de Nombres Procariotas con Posición en Nomenclatura* (LPSN por sus siglas en inglés), el género *Campylobacter* está formado por 43 especies, 13 de ellas definidas en los últimos 3 años [6]. La Tabla 2.1 muestra las especies validadas y publicadas que componen actualmente el género *Campylobacter*. La Figura 2.1 muestra el árbol filogenético del género, obtenido de https://tygs.dsmz.de/ [7, 8].

Nombre	Año y referencia
Campylobacter anatolicus	2021 [9]
Campylobacter armoricus	2019 [10]
Campylobacter aviculae	2020 [11]
Campylobacter avium	2009 [12]
Campylobacter bilis	2022 [13]
Campylobacter blaseri	2018 [14]
Campylobacter canadensis	2007 [15]
Campylobacter coli	1948 [16], 1973 [17]
Campylobacter concisus	1981 [tanner_wolinella_1981]
Campylobacter corcagiensis	2014 [18]
Campylobacter cuniculorum	2009 [19]
Campylobacter curvus	1984 [tanner_wolinella_1984], 1991 [20]
Campylobacter estrildidarum	2021 [11]
Campylobacter fetus	1919 [21], 1963 [22]
Campylobacter geochelonis	2016 [23]
Campylobacter gracilis	1981 [tanner_wolinella_1981], 1995 [24]
Campylobacter helveticus	1993 [25]
Campylobacter hepaticus	2016 [26]
Campylobacter hominis	2001 [27]
Campylobacter hyointestinalis	1985 [28]
Campylobacter iguaniorum	2015 [29]
Campylobacter insulaenigrae	2004 [30]
Campylobacter jejuni	1931 [31], 1973 [17]
Campylobacter lanienae	2000 [32]

Tabla 2.1: Especies pertenecientes al género Campylobacter presentes en LSPD

Nombre	Año y referencia
Campylobacter lari	1984 [33]
Campylobacter majalis	2022 [34]
Campylobacter massiliensis	2021 [35]
Campylobacter mucosalis	1981 [36], 1985 [37]
Campylobacter novaezeelandiae	2020 [38]
Campylobacter ornithocola	2017 [39]
Campylobacter peloridis	2009 [40]
Campylobacter pinnipediorum	2017 [41]
Campylobacter portucalensis	2021 [42]
Campylobacter rectus	1981 [tanner_wolinella_1981], 2015 [43]
Campylobacter showae	1993 [44]
Campylobacter sputorum	1940 [45], 1973 [17]
Campylobacter subantarcticus	2010 [46]
Campylobacter suis	2022 [34]
Campylobacter taeniopygiae	2021 [11]
Campylobacter upsaliensis	1991 [47]
Campylobacter ureolyticus	1978 [48], 2010 [49]
Campylobacter volucris	2010 [46]
Campylobacter vulpis	2021 [50]

Tabla 2.1 continuación

2.1.2. Principales características

El nombre *Campylobacter* proviene de la palabra en griego *kampylos*, que significa curvo, y *bacter*, refiriéndose a su típica forma de bastón. Se trata de bacterias de tinción gram negativa, de 0,2 a 0,8 μ m de ancho por 0,5 a 5 μ m de longitud, con forma de bastón curvo o espiralado. La mayoría de las especies son móviles y presentan un movimiento característico como

de sacacorchos, gracias a la presencia de un flagelo polar, en uno o en ambos extremos, cuya longitud puede superar 2 o 3 veces la longitud de la bacteria.



de/ [7, 8]

Son bacterias microaerofílicas [51], que obtienen energía principalmente a partir de compuestos orgánicos como carbohidratos, aminoácidos y ácidos orgánicos.

Desde el punto de vista epidemiológico, *Campylobacter* es un género asociado a enfermedades gastrointestinales transmitidas por alimentos. Según describe Backert en su libro "Fighting Campylobacter Infections: Towards a One Health Approach" el típico paciente de campylobacteriosos es un niño o joven que vive en un área rural, bebe leche cruda, le gusta comer pollo y nadar en aguas superficiales naturales [52].

2.1.3. Un poco de historia

Si bien el género *Campylobacter* es descrito por primera vez en 1963 [22], los registros más antiguos de la asociación de vibrios similares a *Campylobacter* en humanos datan de 80 años antes: Escherich describe en 1886 la presencia de bacterias espiraladas en el colon de niños muertos a causa de lo que llamó "cólera infantil", y en niños con enfermedades entéricas, pero no logró aislar el microorganismo en medio sólido. Sin embargo, sus hallazgos no fueron reconocidos hasta 1984 [53].

En 1909, los veterinarios McFadyean y Stockman reportan la presencia de una bacteria desconocida, similar a vibrios, aislada frecuentemente en fetos abortados ovinos. Tres años después Smith y Taylor reportan la presencia de esta bacteria, a la que llamaron *Vibrio fetus*, en fetos bovinos [21]. Desde entonces diversos estudios reportaron la presencia de *Vibrio fetus* o bacterias similares en casos de abortos y diarreas en humanos, cerdos y bovinos [54]. No fue sino hasta 1963 cuando Sebald y Veron describen por primera vez el género *Campylobacter* [22]: Se propone que las especies *Vibrio bulbulus* y *V. fetus*, cuyos contenidos GC se aproximan a 30 y 34 % respectivamente, sean parte del género *Campylobacter* (de las palabras griegas para curvo y bastón).

Se propone la siguiente definición del género *Campylobacter g nov*: Bacterias Gram-negativas que ocurren como bastones curvos y finos, que pueden tomar formas esféricas, son móviles debido a la presencia de un flagelo polar mono o multítrico, no esporulantes, anaerobias facultativas u obligadas, que reducen nitrito a nitrato, no acidifican medios que contienen azúcar, son no proteolíticas, y sapróficas o a veces patógenas para el humano y otros animales. El contenido GC de *Campylobacter* es 30 a 34%. La especie tipo es *Campylobacter fetus*.

En esta revisión del género se incluían las especies *C. fetus* (especie tipo), *C. coli* [16], aislada a partir de heces de cerdos; *C. jejuni* [31, 55], aislada de ganado o humanos con enfermedades gastrointestinales y fetos bovinos abortados; y por último, *C. sputorum* [45], aislada en humanos y ganado.

Subsiguientemente se fueron agregando otras especies al género, gracias a mejoras en las técnicas de aislamiento, y con el comienzo de la implementación de técnicas de tipificación molecular más sofisticadas. Esto permitió que se confirmaran especies de *Campylobacter*, como es el caso de *Bacteroides gracilis* y *B. ureolyticus* [43] o bien se trasladaran a otros géneros, como sucedió con las especies *C. pylori* y *C. mustelae*, que fueron incorporadas al género *Helicobacter* [56].

2.1.4. Especies de Campylobacter

Puede distinguirse en el género *Campylobacter* un grupo de especies termotolerantes, capaces de crecer a 41,5 °C en condiciones de microaerobiosis. En este grupo encontramos a *C. coli, C. jejuni, C. lari, C. upsaliensis, C. helveticus*, entre otras. Las especies *C. jejuni* y *C. coli* han sido ampliamente estudiadas, ya que representan casi el 90 % de los casos clínicos de campylobacteriosis y son responsables de más del 80 % de las enteritis a nivel mundial, superando incluso a especies como *E. coli* y *Salmonella* [57].

En los últimos años ha aumentado el número de especies diferentes a *C. coli y C. jejuni* asociados a infección en humanos o animales, siendo reconocidas como patógenos emergentes. El reconocimiento de nuevas especies de *Campylobacter* ha sido posible en gran medida gracias a mejoras en la implementación de métodos moleculares y de cultivo rutinarios empleados para el diagnóstico, tanto en la práctica clínica como en la veterinaria. Sin embargo, el rol de la secuenciación genómica y análisis basados en genómica comparativa ha sido fundamental para conocer las bases evolutivas y epidemiológicas de las especies emergentes de *Campylobacter*.

2.1.5. Genómica aplicada al estudio de *Campylobacter*

Con el advenimiento de técnicas de secuenciación masiva y el creciente interés en el estudio de *Campylobacter* se han secuenciado miles de genomas de cepas extraídas de diferentes hospederos y ambientes, siendo la secuenciación de los genomas una herramienta fundamental para la confirmación de las nuevas especies. Un ejemplo de esto lo constituye la especie *C. hepaticus*, aislada de pollos de uso comercial portadores de la enfermedad del hígado manchado o SLD por sus siglas en inglés (*S*potty *L*iver *D*isease) [26].

El número de genomas disponibles en bases de datos públicas de especies de *Campylobacter* viene en aumento y se acelera en la década del 2000, probablemente debido a mejoras en métodos de diagnóstico y la comercialización

de tecnologías de secuenciación masiva. En la Fig.2 de la revisión "Pathogenomics of Emeging *Campylobacter* species" (pág. 17) [58], se muestra la evolución en el número de taxa (A) y genomas disponibles en NCBI (B) en el tiempo.

Los resultados de una búsqueda de genomas ensamblados utilizando el término "Campylobacter" en el sitio web del NCBI denota la disponibilidad de más de 83.000 genomas secuenciados, de los cuales más del 96 % pertenecen a las especies *coli* y *jejuni*. La baja disponibilidad de genomas de la mayoría de las especies emergentes impide la realización de estudios poblacionales y de conocer la evolución y el potencial patogénico de las mismas. En el artículo mencionado anteriormente [58], mostrado al final del capítulo se revisan los principales estudios genómicos realizados hasta su publicación en 2019.

2.2. Campylobacter fetus

Como se mencionó anteriormente (Ver pág. 10), en 1963 Sebald y Veron describen el género *Campylobacter*, con *C. fetus* como su especie tipo.

Campylobacter fetus es una especie históricamente relacionada a la producción de fallas reproductivas en ganado vacuno y en ovinos. Pero también supone un importante patógeno humano, al ser causante de la mayoría de las bacteremias causadas por *Campylobacter*, afectando principalmente a individuos inmunocomprometidos o añosos [59], presentándose a menudo con desenlaces fatales.

Su primer registro data de 1909, en asociación con la ocurrencia de abortos en ovinos. Luego, se detecta su presencia en fetos bovinos abortados y debido a su semejanza con *Vibrio colerae*, Smith y Taylor le adjudican el nombre *Vibrio fetus*.

No fue sino hasta 1947 que se asocia la entonces llamada *Vibrio fetus* a casos de aborto y bacteremia en humanos [60]. En 1959 Florent reporta la presencia de dos variantes de *V. fetus* en base a observaciones fenotípicas:

- *V. fetus venerealis* Esta denominación reúne a las cepas de *V. fetus* en las que se observó transmisión venérea, con tropismo por el tracto genital de bovinos, observándose en estos animales la ocurrencia de abortos recurrentes e infertilidad.
- *V. fetus intestinalis* Las cepas agrupadas bajo la subespecie *intestinalis* fueron caracterizadas por su capacidad de colonizar el tracto intestinal de ovinos, bovinos, humanos, cerdos y aves. Sin embargo, se describe su aislamiento a partir del tracto genital de ovinos y bovinos, en los que produce abortos esporádicos.

En el mismo trabajo, Florent propone a su vez la utilización de pruebas bioquímicas para diferenciar entre *V. fetus venerealis* y *V. fetus intestinalis*, al observar que *V. fetus intestinalis* a diferencia de la subespecie *venerealis*, producía H_2S y era capaz de crecer en un medio suplementado con glicina 1 % [61]. Este método de diagnóstico es aceptado hasta la actualidad [62].

2.2.1. Posición taxonómica

Campylobacter fetus pertenece al grupo de especies de *Campylobacter* no termotolerantes, debido a que, en general, a diferencia de las especies termotolerantes como *C. jejuni* y *C. coli* no son capaces de crecer a 42 °C. La especie hermana a *C. fetus* es *C. hyointestinalis*, especie descrita en 1985 [28], formada por dos subespecies con gran asociación a hospedero [63]. *C. iguaniorum* y *C. lanienae* son las otras dos especies cercanas a *C. fetus*, que a diferencia de *C. fetus* y *C. hyointestinalis* no han sido reportadas como causantes de infección en humanos o animales (Ver Fig.1 en pág. 3 de "Pathogenomics of emergent *Campylobacter* species" [58]).

2.2.2. Subespecies

Campylobacter fetus reúne tres subespecies: *C. fetus testudinum* (Cft), *C. fetus fetus* (Cff) y *C. fetus venerealis* (Cfv). A continuación se describe brevemente la subespecie Cft, asociada principalmente a reptiles, para luego profundizar en Cff y Cfv, subespecies que tienen a diferentes mamíferos como sus principales hospederos. (Ver Figura 2.2).



Campylobacter fetus testudinum (Cft)

Esta subespecie es la más reciente [64], ya que se propuso su creación para agrupar aislamientos de *C. fetus* obtenidos de humanos y reptiles que mantenían diferencias genéticas significativas con las cepas típicamente aisladas de mamíferos [65]. En 2001, Tu et al. reportan que las cepas de *C. fetus* aisladas de reptiles se diferencian de las "clásicas" aisladas de mamíferos en sus secuencias de los genes 16S, *recA* y *sapD*, y pueden ser discriminadas mediante la aplicación de técnicas como RAPD, PFGE e hibridación ADN-ADN, proponiendo además la separación de las cepas reptilianas en un nuevo taxo, definido a nivel de una nueva especie o subespecie [66].

El primer aislamiento de una cepa de *C. fetus* de un hospedero reptil fue reportado en 1985 [67], durante una investigación clínica de un posible caso de salmonelosis en un niño. Desde entonces se han reportado diversos casos clínicos, sobre todo en pacientes añosos o inmunocomprometidos. Cft se ha aislado a partir de heces de reptiles sanos y sintomáticos y en humanos que adquirieron la enfermedad mediante transmisión zoonótica o a partir de alimentos [68, 69].

Campylobacter fetus fetus (Cff)

Inicialmente descrita como una cepa de *C. fetus* capaz de colonizar el tracto intestinal de ovinos, bovinos, y humanos, entre otros mamíferos, Cff es reconocida inicialmente con el nombre *V. intestinalis*, hasta la creación del género *Campylobacter* en 1972 [17].

Cff se encuentra comúnmente en el tracto intestinal de bovinos y ovinos, transmitiéndose principalmente por vía oral-fecal. En vacas preñadas la bacteria puede traslocarse hacia las membranas placentarias, produciendo inflamación y abortos esporádicos [70, 71].

En humanos, Cff se relaciona principalmente con la enfermedad gastrointestinal, aunque en algunos casos la bacteria puede diseminarse a otros órganos y causar infecciones sistémicas [5], como bacteremia, meningitis, osteomielitis y endocarditis [57]. La infección puede ocurrir a partir del consumo de alimentos contaminados, especialmente carne de animales infectados, y también puede transmitirse de persona a persona [72] o de animales a humanos [70, 73].

Campylobacter fetus venerealis (Cfv)

Esta subespecie es la más antiguamente reportada. Desde el punto de vista genómico, está íntimamente relacionada con la subespecie *fetus*, sin embargo, difiere en cuanto a sus características ecológicas y relevancia clínica, ya que Cfv coloniza el tracto genital de los animales expuestos y presenta transmisión venérea. Las vacas infectadas presentan infertilidad temporal, alteraciones en el ciclo estral y abortos, causando la enfermedad conocida como Campylobacteriosis Genital Bovina (CGB).

2.3. La Campylobacteriosis Genital Bovina (CGB)

La CGB es una enfermedad venérea frecuente en bovinos, que se caracteriza por producir problemas reproductivos y abortos. A continuación se describen sus principales características.

2.3.1. Características

Se considera a la subespecie Cfv y al biovar intermedio (Cfvi) como los agentes causantes de la CGB. Los toros actúan como reservorios ya que son portadores asintomáticos, que presentan los epitelios del pene y el glande colonizados, generalmente sin una respuesta inflamatoria o infección ascendente asociada. La ausencia de síntomas clínicos y el hecho de que la calidad del semen de toros portadores de Cfv no resulta afectada, dificultan el diagnóstico de la CGB. Generalmente la enfermedad es detectada cuando se observan cambios en la capacidad reproductiva de un rebaño, ya que ante la presencia de un animal infectado, Cfv se disemina rápidamente de forma venérea. Si bien la monta natural es especialmente importante en la diseminación de la enfermedad, en ausencia de controles, el semen o instrumental contaminados también son vehículos de transmisión en sistemas en los que se utiliza inseminación artificial [74].

A diferencia de los toros, en vacas la enfermedad tiene una presentación clínica característica. A partir de la exposición, Cfv puede colonizar todo el tracto genital y generar una reacción inflamatoria que provoca vaginitis, cervicitis, endometritis y salpingitis, entre otras manifestaciones [75]. Cfv no tiene un efecto negativo sobre la fecundación, sin embargo, coloniza las membranas placentarias en hembras preñadas, lo cual produce sufrimiento fetal debido a interferencias en la circulación, y finalmente el aborto [76], que, generalmente, ocurre entre los 30 y 70 días de gestación. Si bien pueden ocurrir abortos en fases más tardías, estos son poco frecuentes y se producirían a partir de reacciones asociadas a endotoxinas termoestables secretadas por Cfv [77]. En general, las hembras son portadoras de Cfv temporalmente. Se ha demostrado la persistencia de Cfv en vacas infectadas experimentalmente, que llegó a superar los 90 días post-inoculación [78]. Las vacas pueden seguir siendo portadoras de Cfv de una temporada de servicio

a la siguiente, con lo que puede seguir transmitiéndose aún tiempo después de la infección inicial.

2.3.2. Diagnóstico

Se han desarrollado diferentes técnicas de diagnóstico de CGB, en general enfocadas a identificar el agente etiológico de la enfermedad, ya sea mediante métodos directos o indirectos. Debido a que los toros son portadores asintomáticos, el diagnóstico se realiza preferentemente a partir de ellos, siendo la mucosa prepucial la muestra más habitual, que puede ser extraída mediante raspaje, aspiración o lavado. En casos de aborto también se pueden utilizar muestras derivadas de placenta y órganos o fluidos del feto para el diagnóstico. Sin embargo las muestras de *mucus* vaginal no se consideran las más apropiadas debido a la naturaleza temporal de la infección en vacas.

Para garantizar el correcto diagnóstico, las muestras deben ser trasladadas cuidadosamente, en recipientes herméticamente cerrados y refrigerados y ser procesadas inmediatamente.

Identificación de C. fetus

Se han propuesto diferentes metodologías para la detección de *C. fetus* a nivel de especie y de subespecie con resultados variables. Estos pueden clasificarse en tres categorías: bacteriológicos, inmunológicos y moleculares.

Los métodos bacteriológicos dependen del aislamiento y cultivo de *C*. *fetus*. La naturaleza fastidiosa de *Campylobacter* en general y de la especie *fetus* en particular, para la cual se requieren condiciones de microaerobiosis y medios selectivos específicos, dificulta el aislamiento de la bacteria y por lo tanto su diagnóstico. *C. fetus* se detecta en medios de cultivo selectivos, que contienen antibióticos que inhiben el crecimiento de otras especies de *Campylobacter*. Las características que determinan la presencia de *C. fetus* en cultivo son:

- **Crecimiento:** Se ha descrito a *C. fetus* como una especie no termotolerante, a diferencia de especies como *C. coli* y *C. jejuni*, que son capaces de crecer a 42°C. Sin embargo se ha descrito la tolerancia de aislamientos de *C. fetus* a esta temperatura [79].
- **Colonias:** La morfología de colonia es suficiente para discriminar a *C. fetus* de otras *Campylobacters*, pero no de diferenciar entre subespecies. *C. fetus* se desarrolla en colonias que aparecen típicamente de 2 a 5 días post-inoculación, de color rosa-grisáseo, redondas, convexas, lisas y brillantes y con borde regular, de 1 a 3 mm de diámetro, al ser crecidas en el medio TTE a 37°C bajo una atmósfera con entre 5 y 10% de oxígeno.
- **Pruebas bioquímicas:** El manual terrestre de la Organización Mundial de Sanidad Animal (OMSA) [62] recomienda el uso de las pruebas de tolerancia a glicina 1 % y de producción de H_2S para distinguir las subespecies de *C. fetus*. La capacidad discriminatoria de estas pruebas fue propuesta por Florent en 1959 [61] (Ver página 14), definiendo como Cff a aquellos aislamientos productores de H_2S en un medio suplementado con cisteína, que son además capaces de crecer en presencia de glicina 1 %, mientras que un aislamiento de Cfv resultaría negativa a ambas pruebas. En 1973 Veron y Chatelain describen una cepa de *C. fetus* con un fenotipo intermedio, es decir, al igual que Cff produce H_2S , sin embargo es intolerante a la glicina, tal como Cfv. Así es que se define *C. fetus* subsp. *venerealis* bv. *intermedius* para englobar cepas intermedias.

Métodos inmunológicos: Se han aplicado diferentes técnicas inmunológicas para la detección de antígenos específicos de *C. fetus* directamente sobre muestras de lavado prepucial de toros. La utilización de técnicas basadas en inmunofluorescencia permite detectar hasta 10² unidades formadoras de colonias (UFC) directamente a partir de muestras de lavado prepucial de toros, con sensibilidad y especificidad en el entorno del 93 y 89 % respectivamente [80]. Otra técnica basada en la detección de antígenos de *C. fetus* fue descrita por Campero et al., en un trabajo en el que se realizan procedimientos de inmunohistoquímica sobre diferentes tejidos de fetos bovinos y ovinos abortados [81], con alta especificidad y sensibilidad.

Otras técnicas inmunológicas se basan en la detección de anticuerpos en muestras de lavado prepucial, *mucus* vaginal o tejidos de fetos abortados [82]. El test de aglutinación se ha utilizado extensamente como método de *screening*, pero debido a la alta tasa de falsos positivos y negativos no se recomienda su uso. También se ha usado ELISA para detectar la presencia de anticuerpos IgA específicos de antígenos de *C. fetus* en *mucus* vaginal, con el fin de determinar la etiología del causante de abortos en rebaños bovinos con problemas reproductivos, con una especificidad de 98,5 %. Sin embargo esta técnica no distingue respuestas inmunológicas a diferentes subespecies [62].

Métodos moleculares: Se han desarrollado diversas metodologías moleculares para la identificación y tipificación de *C. fetus*. A continuación se describen algunas de ellas:

Secuenciación del gen 16S: Fue propuesta como una herramienta eficaz para la identificación de *C. fetus* a nivel de especie, pero la baja variabilidad presente en este gen impide una mayor resolución [83].

- Polimorfismo de longitud de fragmentos amplificados (AFLP): Esta técnica de perfilado a nivel genómico fue utilizada por Duim et al. para determinar su utilidad en la tipificación de diferentes especies de *Campylobacter*. Se logró distinguir a *C. fetus* a nivel de especie y subespecie en las 10 cepas analizadas [84]. A pesar de que se consideró como una metodología válida para la identificación de las subespecies de *C. fetus*, su complejidad dificulta su utilización en el diagnóstico [85].
- Electroforesis en gel de campo pulsado (PFGE): Esta técnica, desarrollada por Salama et al., se ideó inicialmente para discernir las variaciones en la longitud de los genomas entre distintas subespecies. En este método, el ADN se somete a la acción de enzimas de restricción como SmaI y SaiI, y los fragmentos resultantes se separan mediante electroforesis en un gel sometido a campos eléctricos variables [86]. Sin embargo, existen discrepancias entre los resultados obtenidos mediante esta técnica y los datos provenientes del análisis bioquímico de las cepas. Estas discrepancias llevaron a la conclusión de que la longitud de los genomas no constituye un criterio suficiente para diferenciar entre las subespecies. Por consiguiente, Hum et al. emplearon PFGE para explorar los patrones de bandas resultantes y así distinguir las diferencias entre las subespecies [87]. A pesar de ello, se evidenció que el análisis visual de los patrones de bandas obtenidos no era un método confiable para la tipificación de un gran número de cepas. La identificación de la subespecie se basaba en la comparación visual entre el patrón de bandas de una muestra y el de una cepa, lo que hacía que este proceso estuviera sujeto a la interpretación subjetiva del investigador [88]. Por esto y debido a su complejidad no es considerado un método de diagnóstico de rutina.

- *Multi Locus Sequence Typing* (MLST): Esta técnica fue desarrollada inicialmente para tipificar aislamientos de *Neisseria meningitidis* provenientes de portadores sanos y de infecciones invasivas [89]. Se basa en la variabilidad de secuencia de fragmentos de genes *housekeeping*: a cada secuencia única de cada *locus* se le asigna un número, para luego combinar los alelos para obtener un ST o *sequence type*. Existen diferentes esquemas de MLST que se aplican a diferentes organismos, y se encuentran depositados en una base de datos especializada. Para implementar MLST en *C. fetus* Van Bergen et al. se basaron en el esquema disponible para la tipificación de las especies *C. coli-C. jejuni* [91], compuesto por loci en los genes *aspA*, *glnA*, *gltA*, *glyA*, *pgm*, *tkt*, y *uncA*. Actualmente la base de datos https://pubmlst.org/ [92] cuenta con 78 registros de ST para *C. fetus*.
- **Reacción en Cadena de la Polimerasa (PCR):** Las metodologías fundamentadas en PCR representan una alternativa más eficaz para la identificación de *C. fetus* y sus subespecies en laboratorios de diagnóstico. Se han creado diversas técnicas de PCR, tanto convencionales como en tiempo real, que se basan en distintas secuencias blanco. A continuación, se mencionan algunas de estas metodologías:
- Blom et al. (1995): La primera metodología que empleó PCR, utilizando el gen 16S como objetivo, fue descrita por Blom et al. en 1995. Esta herramienta utiliza cebadores específicos para amplificar un fragmento de aproximadamente 600 pares de bases (pb) en el gen 16S, mientras que una sonda nucleotídica específica permite la identificación de *C. fetus* [93].
- Oyarzabal et al. (1997): Esta herramienta, también basada en la amplificación de un fragmento del gen 16S, no requiere de la aplicación de otra técnica molecular para identificar *C. fetus* [94] y probó ser específica al

emplearse en más de 300 aislamientos pertenecientes a los géneros *Campylobacter, Arcobacter, Helicobacter*, entre otros.

- Hum et al. (1997): La metodología de diagnóstico desarrollada por Hum et al. es ampliamente reconocida y utilizada en la actualidad [3]. Este enfoque se basa en una técnica de PCR *multiplex*, que implica el uso de dos pares de cebadores específicos: el par MG3F/R para amplificar un fragmento de 960 pb específico de la especie *C. fetus*, correspondiente al gen *cstA*, y el par VenSF/R para amplificar selectivamente un fragmento de 142 pb del gen *parA*, que se mostró específico de la subespecie Cfv. En 2011 Spence et al. demostraron que la especie *C. hyointestinalis*, encontrada comúnmente en bovinos, puede producir falsos positivos para el par de cebadores VenSF/R [95].
- Wang et al. (2002): Otra metodología basada en PCR *multiplex* fue desarrollada por Wang et al. con el propósito de identificar simultáneamente las principales especies patógenas de *Campylobacter* mediante PCR de colonias. Esta técnica utiliza un par de cebadores para la amplificación de un fragmento del gen de ARNr 23S, dando resultados positivos para cepas de los géneros *Campylobacter*, *Arcobacter* y *Helicobacter*. Por otro lado, la identificación de Cff se realiza mediante la amplificación de una región específica del gen *sapB2*. Esta PCR *multiplex* se sometió a prueba utilizando un total de 137 cepas, de las cuales 12 eran de *C. fetus*, amplificándose el fragmento específico esperado en las 6 cepas de Cff, mientras que se obtuvieron resultados negativos para las 6 cepas de Cfv analizadas [96].
- Van Bergen et al. (2005): La técnica basada en PCR propuesta por Van Bergen et al. en 2005 permite la amplificación específica de un fragmento exclusivo de

Cfv, identificado mediante AFLP [97]. Los autores comprobaron resultados 100 % consistentes al perfilado mediante AFLP, pero se demostraron inconsistencias con la metodología desarrollada por Hum et al. en 8 de 65 cepas analizadas, y una baja correspondencia con los resultados de Wang et al. [96].

- McMillen et al. (2006): El par de cebadores diseñado por Hum et al. para la amplificación específica de un fragmento del gen *parA* de Cfv fue empleado por McMillen et al. para desarrollar una metodología en la que se utiliza la actividad de exonucleasa 5'-3' de la polimerasa Taq para clivar una sonda marcada durante la amplificación mediante qPCR, con el fin de mejorar la sensibilidad de la PCR convencional [98]. Los autores pusieron a prueba esta estrategia en 219 muestras de lavado prepucial y 120 de mucus vaginal bovino, demostrando su utilidad al poder aplicarse directamente sobre las muestras sin necesidad de un paso previo de extracción de ADN. Sin embargo, debido a la reactividad cruzada observada en la PCR desarrollada por Hum et al., como demostró Spence et al., y a la prevalencia de *C. hyointestinalis* en bovinos, la especificidad de la técnica no puede ser completamente garantizada.
 - Abril et al. (2007): En 2007, un estudio realizado por Abril et al. informó sobre la presencia de un elemento de inserción denominado ISCfe1 en 25 de 26 cepas de Cfv, mientras que no se pudo detectar en ninguna de las 27 cepas de Cff evaluadas [99]. A raíz de este hallazgo, se desarrollaron los cebadores CVEN-L y CVEN-R2, diseñados específicamente para amplificar un fragmento de 233 pb del primer ORF (*Open Reading Frame*) del elemento de inserción, que codifica la trasposasa TnpA.

- Iraola et al. (2012): La metodología de PCR *multiplex* desarrollada por Iraola et al. en 2012 combina la especificidad de los cebadores diseñados por Hum et al. para identificar *C. fetus* a nivel de especie, con la amplificación de un fragmento del gen *virB11* de Cfv, identificado por Moolhuijzen et al. en base a la comparación de los genomas de las cepas Cff 82-40 y Cfv AZUL-94 [101, 100]. Para esta metodología, testeada para identificar 20 cepas de *C. fetus* (14 Cfv, 1 Cfvi y 4 Cff) y 5 cepas de otras especies de *Campylobacter*, se lograron tasas de sensibilidad y especificidad de 100 % [100].
- Graaf-Van Bloois et al. (2013): En un artículo publicado por Graaf-Van Bloois et al. en 2013 se comparan diferentes técnicas basadas en PCR que permitían el diagnóstico de *C. fetus* a nivel de especie y de subespecie [85], y proponen una PCR *multiplex* en tiempo real que combina los blancos moleculares reportados por Abril et al.: *nahE* para la identificación a nivel de especie y el elemento de inserción ISCfe1 como secuencia específica para Cfv. Si bien la identificación a nivel de especie tuvo 100% de especificidad y sensibilidad, la identificación de Cfv falló en 2 de 60 cepas de Cfv, a lo que los autores atribuyen a variación interna del elemento de inserción ISCfe1 en la región abarcada por los cebadores. Al considerar otra región de ISCfe1 más conservada, si bien se logró el 100% de sensibilidad, se registró reacción cruzada con *C. hyointestinalis*, razón por la cual los autores no recomiendan la utilización de ISCfe1 como blanco de diagnóstico de Cfv.
 - Iraola et al. (2016): Esta técnica consiste en la amplificación mediante qPCR de una región de 78 pb del gen 16S, que incluye una segmento hipervariable de 19 pb que permite distinguir a *C. fetus* de otras especies, identificado mediante el alineamiento de 1907 secuencias del gen

16S de *Campylobacter* [102]. La metodología fue testeada sobre 76 aislamientos, dando resultados negativos en todos los pertenecientes a *Campylobacter* no *fetus* (10/10) y positivos en los 66 restantes (1 Cft, 10 Cfvi, 20 Cfv, 25 Cff y 4 *Campylobacter spp*).

2.3.3. Control y prevención de la CGB

La Organización Mundial de Salud Animal (OMSA) incluye a la CGB en su lista de enfermedades de notificación obligatoria, lo que significa que los casos registrados deben notificarse a esta organización. Esta Organización actúa como contralor internacional, estableciendo normas y recomendaciones con el fin de mejorar la sanidad y el bienestar de los animales, proporcionando además guías metodológicas para el diagnóstico, prevención y tratamiento de las principales enfermedades que los afectan. La OMSA recomienda, en el Capítulo 11.3 del Código Sanitario para los Animales Terrestres, la exigencia de un certificado veterinario internacional que acredite la ausencia del agente causante de la CGB en animales o semen a importar, justificando la inexistencia de situaciones que posibiliten la enfermedad, o mediante las técnicas de diagnóstico recomendadas en el Manual de Animales Terrestres [103].

A pesar de las imposiciones destinadas a limitar la dispersión de la CGB en el comercio internacional, en nuestro país no existe un programa oficial para el control de la enfermedad, y no se realizan estudios sistemáticos para evaluar su prevalencia. En un informe presentado por el MGAP en 2017¹ el Estado reporta la presencia de CGB en Uruguay, pero no detalla datos epidemiológicos.

¹Informe sanitario presentado ante la OEA correspondiente al año 2016.

Recomendaciones de la OMSA

El Manual de las Pruebas de Diagnóstico y de las Vacunas para los Animales Terrestres, elaborado por la OMSA [62], describe a la subespecie Cfv como la causante de la CGB y recomienda su diagnóstico mediante la caracterización del patógeno a través de cultivo y pruebas fenotípicas de tolerancia a la glicina y producción de H_2S . Se desaconseja el uso de PCR para el diagnóstico debido a la falta de sensibilidad y/o especificidad en las herramientas propuestas, así como el análisis genómico por su falta de concordancia con los ensayos fenotípicos.

2.3.4. Incidencia en nuestro país

La presencia de CGB en Uruguay se detectó en la década de 1960 asociada a ganado lechero. En ganado cárnico se identificó en la década de 1980.

Los reportes en los que se estima la prevalencia de la CGB en Uruguay son escasos. El más importante, realizado por Repiso et al., en el que se analizaron 1.754 muestras de raspaje prepucial de toros mediante aislamiento e inmunofluorescencia directa (IFD) [104]. En este trabajo se estimó una prevalencia predial de 37 %, mientras que la prevalencia individual fue de 28 % de toros positivos mediante IFD y 2,7 % mediante cultivo. Otro reporte, realizado por un grupo de investigadores del INIA y el DILAVE en el que analizan 104 muestras de toros para carne obtenidas durante los años 2013 y 2014 en el norte de Uruguay, reveló una prevalencia total de 55 % mediante IFD y PCR [1].
2.3.5. Situación sanitaria y manejo de la producción ganadera en Uruguay

La ganadería es una de las actividades económicas más importantes de nuestro país, representando entre 6 y 7 % del PIB en el período comprendido entre 2015 y 2019 [105]. Con un *stock* de más de 11,5 millones de cabezas de ganado bovino, Uruguay se sitúa como el mayor productor de bovinos por habitante del mundo y con la mayor producción de ganado bovino por kilómetro cuadrado en la región. Esta industria está claramente dirigida a la exportación: Según datos del anuario elaborado por el Instituto Nacional de Carnes (INaC), Uruguay se ubica entre los primeros 10 exportadores mundiales de carne vacuna con más de 500.000 toneladas de peso carcasa exportados en 2022, destinando ese año al consumo doméstico unas 94.200 toneladas de carne, lo que se traduce en un 15 % del total producido, mientras que la actividad lechera, según datos del Instituto Nacional de la Leche (INaLe), ubica al país en el 9no puesto de exportadores a nivel mundial, destinando el 70 % de la producción a la exportación, que en el último año móvil ascendió a casi 212.000 toneladas métricas.

La actividad ganadera en nuestro país se desarrolla extensivamente, y a pesar de su importancia para la economía del país se aplica con bajos niveles tecnológicos en la mayoría de los establecimientos. La estrategia de control para mitigar el impacto de la CGB y otras enfermedades venéreas que inciden sobre la producción ganadera consiste en cortar la cadena de transmisión mediante la implementación de estrategias de manejo optimizadas. Entre las medidas recomendadas para este fin de mencionan [1]:

Inseminación artificial: Se recomienda la adopción de esta tecnología para la cría de ganado, en contraposición a la monta natural. Sin embargo, es necesario realizar controles sobre las dosis de semen distribuidas para descartar la presencia de Cfv. Esta medida es adoptada en mayor medida por los establecimientos más grandes, en los que la atención veterinaria es continua, sin embargo, la práctica de repaso con toros es extendida. En 2015 el alcance de la inseminación artificial fue de 83,4 % en predios lecheros, y un 36 % de estos aseguró realizar repaso con toros [106].

- Vacunas: Se ha demostrado la eficiencia de las vacunas para el tratamiento de rebaños con signos de infección, pero esta práctica es poco extendida, ya que sólo el 26 % de los establecimientos lecheros manifestó realizar vacunación contra CGB.
- **Tenencia y movilidad de toros:** Los toros son los diseminadores de la CGB, y como tales, su manejo debe ser controlado. Se recomienda que al comprar toros, estos sean vírgenes para minimizar el riesgo de CGB, o bien realizar controles que descarten la enfermedad, así como también limitar su movilidad entre predios. Sin embargo, el préstamo de toros para servicio es habitual, y se realiza sin control de enfermedades venéreas.
- Acciones ante un brote: Ante el diagnóstico de un caso de CGB en un rodeo, todo el rebaño debe ser muestreado. Los casos positivos deben apartarse para evitar una mayor diseminación, y realizar tratamiento mediante lavados y el uso de antibióticos tanto local como por vía sistémica, y al menos tres controles sucesivos que descarten la presencia de CGB, o bien, el descarte de los animales infectados. También se recomienda la vacunación de todo el rodeo.





REVIEW

Pathogenomics of Emerging Campylobacter Species

Daniela Costa,^{a,b} © Gregorio Iraola^{a,c,d}

^aMicrobial Genomics Laboratory, Institut Pasteur de Montevideo, Montevideo, Uruguay
 ^bSección Genética Evolutiva, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay
 ^cCenter for Integrative Biology, Universidad Mayor, Santiago de Chile, Chile
 ^dWellcome Sanger Institute, Hinxton, United Kingdom

SUMMART	1
INTRODUCTION	2
CURRENT TAXONOMY AND AVAILABLE GENOMIC DATA	2
Campylobacter concisus	4
Campylobacter fetus	7
Mammal-associated C. fetus strains	7
Reptile-associated C. fetus strains	8
Campylobacter hepaticus	9
Campylobacter hyointestinalis	10
Campylobacter lanienae	11
Campylobacter lari and Related Taxa	11
Campylobacter ureolyticus	12
Campylobacter upsaliensis: a Salient Example of a Not (Enough) Sequenced	
Pathogen	12
EVALUTIONARY MECHANICME IN EMERCINE CAMOVI ORACTER CRECKED	
EVOLUTIONARY MECHANISMS IN EMERGING CAMPYLOBACTER SPECIES	13
Host-Associated Population Structure and Adaptation	13 13
Host-Associated Population Structure and Adaptation	13 13 14
Host-Associated Population Structure and Adaptation	13 13 14 15
Host-Associated Population Structure and Adaptation	13 13 14 15 15
Host-Associated Population Structure and Adaptation	13 13 14 15 15 16
Host-Associated Population Structure and Adaptation Barriers to Homologous Recombination Horizontal Gene Transfer Role in the acquisition of antimicrobial resistance Role in the acquisition of virulence mechanisms Genome Reduction	13 13 14 15 15 16 16
Host-Associated Population Structure and Adaptation Barriers to Homologous Recombination Horizontal Gene Transfer Role in the acquisition of antimicrobial resistance Role in the acquisition of virulence mechanisms Genome Reduction. FUTURE DIRECTIONS	13 14 15 15 16 16 17
EVOLUTIONARY MECHANISMS IN EMERGING CAMPTLOBACTER SPECIES Host-Associated Population Structure and Adaptation Barriers to Homologous Recombination Horizontal Gene Transfer Role in the acquisition of antimicrobial resistance Role in the acquisition of virulence mechanisms Genome Reduction FUTURE DIRECTIONS CONCLUSIONS	13 14 15 15 16 16 17 18
EVOLUTIONARY MECHANISMS IN EMERGING CAMPTLOBACTER SPECIES Host-Associated Population Structure and Adaptation Barriers to Homologous Recombination Horizontal Gene Transfer Role in the acquisition of antimicrobial resistance Role in the acquisition of virulence mechanisms Genome Reduction FUTURE DIRECTIONS CONCLUSIONS REFERENCES	13 14 15 15 16 16 17 18 18

SUMMARY Campylobacter is among the four main causes of gastroenteritis worldwide and has increased in both developed and developing countries over the last 10 years. The vast majority of reported Campylobacter infections are caused by Campylobacter jejuni and, to a lesser extent, C. coli; however, the increasing recognition of other emerging Campylobacter pathogens is urgently demanding a better understanding of how these underestimated species cause disease, transmit, and evolve. In parallel to the enhanced clinical awareness of campylobacteriosis due to improved diagnostic protocols, the application of high-throughput sequencing has increased the number of whole-genome sequences available to dozens of strains of many emerging campylobacters. This has allowed for comprehensive comparative pathogenomic analyses for several species, such as C. fetus and C. concisus. These studies have started to reveal the evolutionary forces shaping their genomes and have brought to light many genomic features related to pathogenicity in these neglected species, promoting the development of new tools and approaches relevant for clinical microbiology. Despite the need for additional characterization of genomic diversity in emerging campylobacters, the increasing body of literature describing pathogenomic studies on these species deserves to be discussed from an integrative perspective. This review compiles the current knowledge and highlights future work toward deepening our understanding about genome dynamics and the mechanisms

Clinical Microbiology Reviews

Citation Costa D, Iraola G. 2019. Pathogenomics of emerging *Campylobacter* species. Clin Microbiol Rev 32:e00072-18. https://doi.org/10.1128/CMR.00072-18. Copyright © 2019 American Society for Microbiology. All Rights Reserved. Address correspondence to Gregorio Iraola, giraola@pasteur.edu.uy. Published 3 July 2019

cmr.asm.org 1

31

governing the evolution of pathogenicity in emerging *Campylobacter* species, which is urgently needed to develop strategies to prevent or control the spread of these pathogens.

KEYWORDS *Campylobacter*, emerging pathogens, genome evolution, pathogenomics, whole-genome sequencing

INTRODUCTION

he first recognized Campylobacter infection was reported in 1913 by McFaydean and Stockman, as they found a curved-shaped microorganism causing abortion in sheep and cattle (1). This bacterium remained unnamed until 1919, when Smith and Taylor isolated the same microorganism from bovine fetal fluids and named it Vibrio fetus (2). Then, in 1973 Véron and Chatelain proposed the genus Campylobacter by reclassifying V. fetus to Campylobacter fetus (3). In addition to this long and recognized importance as a veterinary pathogen since the beginning of the 20th century, C. fetus was subsequently identified as the causative agent of bloodstream infections in humans (4, 5). However, the major relevance of campylobacters as a main cause of human disease was just uncovered in the early 1980s, after the development and widespread implementation of selective media for the isolation of Campylobacter from stool samples. Today, the most relevant species within the genus is C. jejuni, a leading cause of bacterial gastroenteritis in humans whose worldwide incidence is even higher than that of very well-known pathogens that cause acute gastrointestinal infections, such as Escherichia coli, Shigella, or Salmonella. A close relative to C. jejuni is C. coli, which causes 1 to 25% of all Campylobacter-related diarrheal diseases (6). The remaining species of the genus have been much less studied, but the enhanced ability to detect campylobacters caused by the routine implementation of molecular techniques and the improvement of culture media and growth conditions allowed for the description and identification of a growing diversity of Campylobacter species distinct from C. jejuni and C. coli as relevant pathogens for humans and other animals.

The clinical awareness of many of these emerging Campylobacter species coincided with the advent of high-throughput sequencing as a popular tool for studying the microbial world, which fed the interest in applying whole-genome sequencing and comparative genomics to elucidate how emerging campylobacters cause disease, transmit, and evolve. The first complete genome sequence of C. jejuni was published almost 20 years ago (7), and today, several thousands of C. jejuni and C. coli genomes can be accessed through public databases. Accordingly, the increasing number of whole-genome sequences has allowed a transition in comparative genomics studies that initially included only a few genomes and now comprise hundreds to thousands of them. However, the availability of genomic data for emerging Campylobacter species is still lagging and more fragmented, hindering the improvement of our understanding of the biology of nonclassical Campylobacter pathogens. In this review, we summarize the state-of-the-art literature about emerging campylobacters in light of comparative genomics, discuss how these data are helping to uncover basic aspects of Campylobacter pathobiology and its applications in clinical microbiology, and highlight upcoming challenges in the field, including future work needed to mitigate sequencing bias in favor of well-known species. This review constitutes a comprehensive resource for researchers working on Campylobacter genomics and emerging pathogens, aiming to integrate the knowledge and future challenges in the field of emerging Campylobacter pathogens.

CURRENT TAXONOMY AND AVAILABLE GENOMIC DATA

To date, the genus *Campylobacter* consists of 32 officially described species and 9 subspecies, namely, *C. avium* (8), *C. blaseri* (9), *C. canadensis* (10), *C. coli*, *C. concisus* (11), *C. corcagiensis* (12), *C. cuniculorum* (13), *C. curvus* (14) *C. fetus* subsp. fetus (3), *C. fetus* subsp. venerealis (3), *C. fetus* subsp. testudinum (15), *C. geochelonis* (16), *C. gracilis* (17), *C. helveticus* (18), *C. hepaticus* (19), *C. hominis* (20), *C. hyointestinalis* subsp. hyointestinalis

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews



FIG 1 Phylogenetic relationships between described Campylobacter species. A phylogenetic tree of Campylobacter species dividing the genus into five distinct groups, namely, the C. fetus group, C. jejuni group, C. lari group, C. concisus group, and C. ureolyticus group, is shown. Names were assigned by considering the most clinically relevant species within each group. Tip labels are colored in red for species documented to cause infections.

(21), C. hyointestinalis subsp. lawsonii (22), C. iguaniorum (23), C. insulaenigrae (24), C. jejuni subsp. jejuni (3), C. jejuni subsp. doylei (25), C. lanienae (26), C. lari subsp. lari (27), C. lari subsp. concheus (27), C. mucosalis (28), C. ornithocola (29), C. peloridis (27), C. pinnipediorum subsp. pinnipediorum (30), C. pinnipediorum subsp. caledonicus (30), C. rectus (31), C. showae (32), C. sputorum (33), C. subantarcticus (34), C. troglodytis (35), C. upsaliensis (36), C. ureolyticus (37), and C. volucris (38). These species cluster in five discrete phylogenetic groups, which all contain pathogenic microorganisms (Fig. 1), highlighting the clinical relevance of the whole genus. Despite this scenario clearly reflecting the taxonomic diversity and the widespread presence of pathogenic lineages in the genus Campylobacter, not a single genome is available for some species, like C. canadensis, C. troglodytis, and C. mucosalis. Also, for many others, including C. volucris, C. peloridis, C. rectus, C. insulaenigrae, C. hominis, C. helveticus, C. cuniculorum, C. corcagiensis, C. ornithocola, and C. avium (31% of the genus), only a single representative genome per species is available (Table 1). Importantly, when we exclude C. jejuni and C. coli, 13 out of the remaining 30 species (43%) have been at least sporadically reported to be the causative agent of infections in humans and/or other animals, and many of them are frequently associated with diverse clinical presentations, such as invasive blood infections, periodontal infections, abscesses, meningitis, diarrhea, or gastroenteritis (Table 2). The lack of sufficient genomic information on the causative agents of these infections prevents the exploration of intraspecific genetic variability

October 2019 Volume 32 Issue 4 e00072-18

TABLE 1 Reported	hosts and available	genomic information	for members of the o	enus Campylobacter

		No. of genomes at:	
Species or subspecies	Reported host(s) (reference)	PATRIC	NCBI
C. avium	Chicken, turkey (8)	3	3
C. blaseri	Seal (9)	2	1
C. canadensis	Whooping crane (10)	0	0
C. coli	Cattle (120), chicken (121), dog (122), duck (123), goat (124), monkey (125), pig (126), seagull (127), sheep (128), human (129)	1,571	981
C. concisus	Cat (126), dog (122), human (130, 131)	168	163
C. corcagiensis	Lion-tailed macaque (12)	1	2
C. cuniculorum	Rabbit (13)	2	3
C. curvus	Dog (122), human (14, 132)	8	3
C. fetus subsp. fetus	Cattle (3), sheep (133), human (65)	28	24
C. fetus subsp. venerealis	Cattle (3), human (134)	44	26
C. fetus subsp. testudinum	Reptiles (15), human (68), monkey (64)	33	23
C. geochelonis	Hermann's tortoise (16)	3	3
C. gracilis	Dog (122), human (40, 130)	3	3
C. helveticus	Dog (18), cat (18), human (135)	3	3
C. hepaticus	Chicken (19)	13	16
C. hominis	Human (20)	2	2
C. hyointestinalis subsp. hyointestinalis	Cattle (136, 137), human (138, 139), deer and reindeer (140), swine (137), sheep (137), hamster (136), doq (122)	21	18
C. hyointestinalis subsp. lawsonii	Swine (22), cattle (81)	10	10
C. iguaniorum	Reptiles (23), alpaca (141)	3	3
C. insulaenigrae	Pinnipeds (24, 142, 143), cetaceans (24), human (86)	1	2
C. jejuni subsp. jejuni	Cattle (120), chicken (121)	2,595	1,602
C. jejuni subsp. doylei	Human (25, 144, 145)	6	6
C. lanienae	Cattle (83), swine (83), sheep (137), human (26)	27	27
C. lari subsp. lari		13	13
C. lari subsp. concheus	Shellfish (27)	1	1
C. mucosalis	Dog (122), pig (28), human (146)	1	1
C. ornithocola	Wild birds (29)	1	1
C. peloridis	Shellfish (27)	1	1
C. pinnipediorum subsp. pinnipediorum	Sea lion (30)	5	8
C. pinnipediorum subsp. caledonicus	Seal (30)	1	3
C. rectus	Dog (122), human (147)	1	2
C. showae	Dog (122), human (32)	10	10
C. sputorum	Cattle (137, 148), sheep (33), swine (149), dog (122, 135), human (33)	7	7
C. subantarcticus	Wild birds (34)	2	2
C. troglodytis	Chimpanzee (35), human (150)	0	0
C. upsaliensis	Cat (151, 152), dog (97, 151), human (96)	6	6
C. ureolyticus	Cattle (153), horse (154), human (37)	7	7
C. volucris	Black-headed gull (38), human (87)	1	1

and patterns of genomic evolution. Consequently, relevant information about how a vast number of emerging *Campylobacter* species cause disease and transmit between hosts is currently unavailable. However, several groups have made a considerable effort to generate whole-genome sequences for some emerging campylobacters whose relevance for public health is frequently underestimated, uncovering genomic features that represent valuable contributions to understanding the disease biology and epidemiology of these microorganisms. Thus, these cases are discussed for each individual species that have deserved attention from the field of comparative genomics.

Campylobacter concisus

Campylobacter concisus was originally reported in 1981 from periodontal lesions (11); however, its role as an oral pathogen has remained uncertain since healthy individuals have been found to carry this species in the saliva (39). Additionally, *C. concisus* has been detected in fecal samples from diarrheic patients but also in healthy

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

TABLE 2 Emerging Campylobacter species with reported infections in humans or other animals and current record of pathogenomic studies

			nor or genome
Species or subspecies	Human disease(s) [reference(s)]	Animal disease	studies [reference(s)]
C. concisus	Gastroenteritis (131, 155), brain abscess (156), arthritis, Crohn's disease and ulcerative colitis (48, 157), Barrett's esophagitis (158)	Not reported	5 (44, 47, 52, 53, 159)
C. curvus	Preterm birth (160), empyema (161), alveolar abscess (126), liver abscess (162), gastroenteritis (14)	Not reported	Not available
C. fetus subsp. fetus	Gastroenteritis, bacteremia, cellulitis, neurological infections (meningitis, meningoencephalitis, subdural empyema, brain abscess), perinatal infections (uterus infection, abortion, placentitis), vascular infections (endocarditis, vasculitis, thrombophlebitis, pericarditis) (6, 65)	Sporadic abortion (sheep and cow) (163)	6 (56, 58, 61–64)
C. fetus subsp. venerealis	Vaginosis (134)	Infertility and abortion (cow) (164)	7 (55–58, 60, 61, 68)
C. fetus subsp. testudinum	Bacteremia, subdural hematoma (68)	Not reported	4 (69–72)
C. gracilis	Bacteremia (165), empyema (166), brain abscess (156), head infection (167), Crohn's disease (40), ulcerative colitis (168), periodontitis (130)	Dog (diarrhea) (122)	1 (169)
C. helveticus	Diarrhea (135)	Cat (diarrhea) (18), dog (diarrhea) (122)	Not available
C. hepaticus	Not reported	Chicken (spotty liver disease) (19)	2 (19, 74)
C. hyointestinalis subsp. hyointestinalis	Gastroenteritis (138, 139)	Proliferative ileitis (pig) (21)	(77, 78, 82)
C. insulaenigrae	Gastroenteritis (135), septicemia (86)	Not reported	Not available
C. lari subsp. lari	Gastroenteritis (153, 170), bacteremia (88, 89, 171)	Not reported	2 (85, 86)
C. mucosalis	Gastroenteritis (50, 135)	Diarrhea (dog) (122)	Not available
C. pinnipediorum subsp. pinnipediorum	Not reported	Pinniped (abscess) (30)	1 (30)
C. pinnipediorum subsp. caledonicus	Not reported	Pinniped (abscess) (30)	1 (30)
C. rectus	Gastroenteritis (172), Crohn's disease (40, 42), ulcerative colitis (168), periodontal disease (147, 173), bacteremia (174), oral abscess (175), bone abscess (156), empyema thoracis (176)	Not reported	Not available
C. showae	Crohn's disease (40, 42), ulcerative colitis (168), abscess (156)	Diarrhea (dog) (122)	Not available
C. sputorum	Gastroenteritis (177, 178), abscess (179), bacteremia (180)	Diarrhea (dog) (122), sheep (abortion) (33)	2 (116, 181)
C. upsaliensis	Gastroenteritis (96, 182–184), abortion (185), bacteremia (186), breast abscess (187)	Not reported	Not available
C. ureolyticus	Gastroenteritis (92, 153), Crohn's disease (40, 168), ulcerative colitis (168), oral and perianal abscesses (167)	Not reported	3 (94)
C. volucris	Bacteremia (87)	Not reported	Not available

individuals, questioning its role in diarrheic disease. More recently, the prevalence of *C. concisus* has been found to be increased in both children and adult patients with inflammatory bowel disease (IBD) (40–42), suggesting that *C. concisus* may be implicated in its development and progression. Together, these studies show that the role of this species as a human pathogen is still unclear. This has motivated the development of several whole-genome sequencing projects aiming to uncover the genetic variability of this species with greater resolution and its relationship with disease.

First attempts to characterize the genetic diversity of *C. concisus* and its relationship with pathogenicity started with pulsed-field gel electrophoresis (PFGE), DNA-DNA hybridization, and ribosomal gene analyses (43). These results revealed a complex intraspecific taxonomy with high genetic heterogeneity that led to the description of genomospecies, defined as groups of genetically divergent strains without an apparent

October 2019 Volume 32 Issue 4 e00072-18

phenotypic distinction. In 2011, two papers came out describing the whole-genome sequencing and comparison of two C. concisus strains: BAA-1457 (also referred as 13826), isolated from a patient with acute gastroenteritis, and UNSWCD, isolated from a biopsy specimen from a child with Crohn's disease (CD) (44, 45). These initial genomic comparisons confirmed previous findings based on nongenomic approaches and concluded that the two strains presented enough genetic diversity to be classified as distinct species. Also, these studies identified novel genetic features, like the presence of potentially secreted proteins exclusive to the species C. concisus, that could be used as disease markers or diagnostic targets. The assessment of strain-specific genomic regions that were potentially associated with virulence uncovered the presence of type VI secretion system (T6SS) genes in strain BAA-1475. This macromolecular system is implied in host-pathogen interactions and virulence and has been identified in many well-known bacterial pathogens, like Salmonella, Pseudomonas, Yersinia, or Vibrio (46). Another important difference between C. concisus strains was the zonula occludens toxin (zot) gene found in BAA-1475, which was inserted within a prophage. Zot increases intestinal permeability by affecting the tight junctions, a phenotype that is characteristic of IBD. Hence, a defect in the primary intestinal barrier caused by C. concisus Zot could be a mechanism by which this species may be related to the development of IBD. Also, UNSWCD and BAA-1457 were dissimilar in their flagellin glycosylation pathways, suggesting that genomic variability may also determine differential immune responses against genetically distinct C. concisus strains. Together, these preliminary differences found between these two strains, mainly in their repertory of virulence-associated genes, may explain the distinct pathogenic phenotypes found among C. concisus strains and set the basis for future studies involving pathogenomic analyses in this species.

A more comprehensive genomic comparison that involved 36 C. concisus strains from patients with gastroenteritis (isolated from intestinal biopsy specimens) or IBD (isolated from the oral cavity) deepened the characterization of virulence factor repertories and provided some clues about the relationship between intraspecific diversity and pathogenicity (47). This study analyzed for the first time multiple C. concisus genomes from the two main genomospecies, suggesting that the phylogenetic structure (genomospecies) was not linked to oral or intestinal origin, hence supporting a previous hypothesis that proposed that oral strains were the causative agents of gastroenteritis after translocating to the intestinal tract (39, 48). The same study reported several genetic differences between the genomospecies. For example, phosphate transport genes pstS, pstA, and pstC were specific to genomospecies 1, suggesting that different C. concisus genotypes may differ in their phosphate transport capacity. Also, genomospecies 2 encoded an aquaporin Z gene that functions to maintain intracellular osmotic pressure and that may be involved in C. concisus adaptation to environments with fluctuating osmolarity. Another important difference was the uneven distribution of CRISPR/Cas genes that were exclusively found in genomospecies 2. Even though CRISPR/Cas systems prevent the incorporation of foreign DNA, like plasmids and phages (49), no correlation was found between the presence/absence of CRISPR/Cas and the prophage that contains the zonula occludens toxin gene. Indeed, zot was detected in C. concisus strains from both genomospecies.

Another important discovery in this study was the presence of two genomic islands coding for type IV secretion systems (T4SS) and protein effectors similar to those found in pathogens like *Legionella pneumophila* and *Helicobacter pylori* (50). These islands were differentially prevalent in oral or intestinal strains, and even though the data suggested that they may preferably integrate into enteric *C. concisus* strains, analysis of a higher number of strains would be necessary to determine if these differences are statistically significant. These kinds of genomic features that have been identified can be used to guide phenotypic assays that could shed light on the pathogenicity potential between oral cavity- and intestinal biopsy specimen-derived strains or be used as genotypic markers for source tracking. Accordingly, a follow-up study from the same group focused on studying potential genetic factors discriminating commensal

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

from IBD-associated *C. concisus* strains by examining 86 genomes. This study found a novel gene in *C. concisus* that codes for the protein Csep1, which is homologous to enterotoxin B in *Staphylococcus aureus*. Enterotoxin B is involved in *S. aureus* pathogenesis by inducing diarrhea and activating T cells which produce large amounts of proinflammatory cytokines (51). Interestingly, the gene coding for Csep1 in *C. concisus* presented a 6-bp insertion (*csep1-6bpi*) in most strains isolated from CD patients in comparison with the gene from healthy controls (52). Based on this result, the authors suggested the use of *csep1-6bpi* as a molecular marker for CD-associated *C. concisus* strains. Beyond its potential application in the development of molecular methods to detect pathogenic CD-associated *C. concisus* strains, the fact that Csep1 is a secreted protein opens new opportunities to elucidate the molecular mechanisms by which *C. concisus* is implied in CD pathogenesis.

The last pathogenomic analysis of *C. concisus* constitutes the largest sequencing effort so far, which produced 104 *C. concisus* genomes from strains isolated from saliva, feces, and intestinal biopsy specimens. This analysis found no association between the genomospecies and IBD, diarrhea, or healthy controls, since strains were unevenly distributed in both phylogenetic lineages. However, when assessing the anatomical site of collection, genomospecies 2 was predominant in gut mucosal samples, while genomospecies 1 was predominant in oral samples. These differences were also reflected in the pangenomes of both genomospecies, given that genomospecies 2 harbored a bigger accessory genome than genomospecies 1, so this extensive genomic variation between *C. concisus* genomospecies could be related to functional variation and adaptation to different sites within the host. Indeed, several genes whose prevalence increases or decreases in association with the anatomical descent from the oral cavity to mucosal biopsy specimens to feces were identified, supporting the suggestion that the genetic heterogeneity of *C. concisus* is related to the source of isolation more than to the clinical phenotype (53).

Despite recent efforts that have significantly enlarged the number of available genomes from different sources and clinical conditions, the role of *C. concisus* in human disease remains elusive, with several studies arriving at contradictory results. Furthermore, most of these studies lack experimental evidence associated with the observed genomic variation which could help to infer mechanistic aspects of *C. concisus* pathogenicity. A couple of recent works have advanced the understanding of phenotypic variation by identifying signals of differential adaptation of genomospecies to the intestinal tract (54) and differential growth in response to specific carbon sources (55). This kind of phenotypic information, coupled with future global-scale sequencing surveys of *C. concisus*, can provide the conditions to apply genomewide association studies, an increasingly useful approach to discover genotype-phenotype associations in bacterial populations.

Campylobacter fetus

Campylobacter fetus is the type species of the genus and has been historically recognized as a livestock pathogen causing reproductive problems, mainly in cattle, and more recently as an increasingly reported opportunistic pathogen in humans. Currently, *C. fetus* is divided into three subspecies: *C. fetus* subsp. *fetus* and *C. fetus* subsp. *venerealis* are primarily isolated from humans and cattle, respectively, and have traditionally been defined on the basis of two biochemical tests (growth in 1% glycine and H₂S production). *C. fetus* subsp. *venerealis* by. *intermedius* is also described as a biochemical variant at the intrasubspecific level. *C. fetus* subsp. *testudinum* is genetically divergent from the others and is mostly isolated from reptiles but is also isolated from ill humans.

Mammal-associated C. fetus strains. The first C. fetus genome was sequenced with the Sanger method and belongs to strain C. fetus subsp. fetus 82-40, which was isolated from an immunocompromised human patient in the United States. Three years after the release of this genome in 2006, the first C. fetus subsp. venerealis genome was reported from strain Azul-94, isolated from a bovine abortion in Argentina (56). An

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

initial comparison of these two genomes allowed identification of markers that were subsequently used for the molecular characterization of *C. fetus* subsp. *venerealis* and *C. fetus* subsp. *fetus* (57) and to propose candidate virulence determinants that could potentially have explained clinical differences between these two subspecies. In particular, a genomic island harboring a type IV secretion system and putative plasmid genes was detected in the bovine strain *C. fetus* subsp. *venerealis* Azul-94 and was absent in the human-associated strain *C. fetus* subsp. *fetus* 82-40. However, further studies evidenced that this and other similar genomic islands coding type IV secretion systems were frequent in the chromosomes and plasmids of both subspecies (58). Whole-genome sequencing of individual *C. fetus* subsp. *venerealis* bv. *intermedius* strains isolated from cattle could not explain the subtle biochemical phenotype that distinguishes *C. fetus* subsp. *venerealis* bv. *intermedius* from *C. fetus* subsp. *venerealis* (59, 60), and no further effort has been dedicated to elucidate this peculiarity.

The first comprehensive analysis of C. fetus genomes involved the comparison of 24 strains, revealing a remarkable inconsistency between population structure (genomic characteristics) and the biochemical tests (phenotypic characteristics) traditionally applied to differentiate C. fetus subsp. fetus from C. fetus subsp. venerealis (61). This study showed the presence of two phylogenetic lineages based on the core genome; one of them exclusively conformed to phenotypically determined C. fetus subsp. fetus, and the other conformed to both phenotypically determined C. fetus subsp. fetus and C. fetus subsp. venerealis, supporting this lack of correlation between phenotypic and genomic information, which raises questions about the clinical relevance of C. fetus subspecies typing by phenotypic assays. Additionally, further inconsistencies were revealed even between multilocus sequence typing (MLST) and core genome phylogenies, indicating that some alleles can undergo homoplasy, confounding epidemiological conclusions based on traditional genotyping approaches (62). Afterwards, an extended study that involved the analysis of 42 strains provided a more accurate description of the population structure of mammal-associated C. fetus strains, concluding that C. fetus subsp. venerealis has derived from a C. fetus subsp. fetus ancestor and suggesting the deletion of a putative cysteine transporter as the reason for the H₂S-negative phenotype in C. fetus subsp. venerealis strains. Nevertheless, the overall conclusion of this study reinforces the notion of inconsistency between biochemical tests and genomics (63). Together, these results point to whole-genome analysis as the current standard approach for typing C. fetus strains from mammal origin. Accordingly, a recent study that analyzed the genomes of 182 C. fetus strains mainly isolated from cattle and humans revealed the presence of 8 discrete lineages that adapted as livestock pathogens or human intestinal pathobionts, as they were found in the gut metagenomes of healthy individuals (64). This study provided the basis to further investigate the evolution and transmission of C. fetus, since its presence in the human gut microbiota could facilitate alternative ways of contagion. Indeed, zoonotic transmission is currently the most widely accepted and documented route (65), but the recent identification of human-to-human transmission between men who have sex with men (66) reinforces the hypothesis that C. fetus behaves both as a zoonotic pathogen and as a pathobiont resident of the intestinal microbiota. In this sense, the upcoming challenge to completely understand the evolutionary landscape and epidemiology of mammal-associated C. fetus should be focused on the analysis of whole genomes from strains isolated from healthy individuals.

Reptile-associated *C. fetus* **strains.** Reptile-associated *C. fetus* isolates are genetically distant from mammal-associated strains, as originally observed by MLST, which evidenced two discrete clusters (sharing 90% of identity) enclosing reptile- and mammal-associated strains (67). These reptile-associated *C. fetus* strains have been found to cause infections in humans (68), and their genetic distance from mammal-associated strains motivated the whole-genome sequencing of its type strain, *C. fetus* subsp. *testudinum* 03-427, originally isolated from a human (69). This genome was released months before the official taxonomic revision that defined genetically distant, reptile-associated *C. fetus* subsp. *testudinum* (15). The

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

subsequent whole-genome sequencing of strain *C. fetus* subsp. *testudinum* Pet-3, isolated from a reptile, revealed high genomic homogeneity between these two isolates from different host species (70).

The first comprehensive comparative study including reptile- and mammal-associated C. fetus strains (that analyzed 61 C. fetus subsp. fetus/C. fetus subsp. venerealis genomes and 18 C. fetus subsp. testudinum genomes) identified a recombinant locus that differentiates reptile- from human-derived C. fetus subsp. testudinum strains and that could be related to the invasive phenotype of human infections. These genomic comparisons also confirmed that mammal- and reptile-associated strains cluster in different phylogenetic lineages which are genetically isolated due to the existence of a barrier to lateral gene transfer or interlineage recombination. Consequently, several host-associated genomic features were found, including a tricarballylate catabolism pathway present in C. fetus subsp. testudinum but absent in C. fetus subsp. fetus/C. fetus subsp. venerealis which might explain adaptation to reptilian hosts (71). However, the recent identification of a C. fetus lineage that was isolated from reptiles but that was genetically closer to mammal strains and that showed strong signals of recombination with C. fetus subsp. testudinum evidenced that barriers to homologous recombination between divergent lineages of C. fetus occurring within the same host are not absolute (72). As the reptile gut seems to be a frequent niche for C. fetus, future work should couple metagenomics and Campylobacter selective media to uncover novel related taxa whose whole-genome sequencing and comparison could improve our understanding about the ecology and pathogenicity of this species.

Campylobacter hepaticus

Spotty liver disease (SLD) is an emerging infectious disease prevalent in Europe and Australia characterized by multifocal liver lesions with high mortality rates that particularly affects free-range and floor-raised chicken flocks (73). This infection is caused by *Campylobacter hepaticus*, a recently described species that was originally isolated from poultry with SLD and whose zoonotic risk and potential for transmission to other animals are yet unknown (19).

The whole genomes of four C. hepaticus strains were originally sequenced to support the description of C. hepaticus as a novel species (19), including its type strain, HV10. However, the authors did not detail the virulence genes or pathogenicity mechanisms encoded by the C. hepaticus genome. Subsequently, a more comprehensive study focused on the genomic characterization of this species sequenced 10 strains isolated from SLD cases in the United Kingdom (74). This work concluded that C. hepaticus represents a distinct genomic lineage, with the pathogenic C. jejuni/C. coli strains being its closest phylogenetic relatives. Also, this study revealed that C. hepaticus has experienced reductive genome evolution, as evidenced by a lower GC content and an average genome size reduction of \sim 140 kb with respect to those of C. jejuni. This reduction involved the loss of iron acquisition systems and the lack of many well-known Campylobacter virulence determinants, like adhesion factors and capsular polysaccharide biosynthesis or cytolethal distending toxin (CDT) genes. As in many other bacterial pathogens, genome reduction is typically associated with niche specialization, like the chicken liver, so the potential of this emerging pathogen to transmit and cause disease in other hosts (like humans) seems remote. However, the identification of a C. jejuni plasmid encoding tetracycline resistance present in the genomes of C. hepaticus strains evidenced that horizontal transfer between this species and other campylobacters can mediate the acquisition of antimicrobial resistance and pathogenicityrelated determinants.

The availability of a larger, global *C. hepaticus* isolate collection is needed to explore the pangenome variability in this species and to determine its population structure, which will be helpful to better understand the impact of genome reduction and horizontal gene transfer in the evolution and epidemiological dynamics of this emerging pathogen. As well, uncovering which genotypes are circulating globally will provide

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

information that may be used to develop specific typing and detection tools for this emerging pathogen.

Campylobacter hyointestinalis

Campylobacter hyointestinalis was originally isolated from swine with proliferative enteritis and was described as a novel species in 1980 (21). Since then, it has been recovered mostly from healthy animals (like sheep, deer, hamsters, dogs, and cattle) but has also been sporadically isolated from livestock and human infections, pointing to this species as an emerging zoonotic pathogen (6). On the basis of genetic and phenotypic traits, *C. hyointestinalis* is currently divided into two subspecies: *C. hyointestinalis* subsp. *lawsonii* is mainly restricted to pigs, and *C. hyointestinalis* subsp. *hyointestinalis* of *C. hyointestinalis* were based on determining genetic and protein profiles, evidencing considerable intraspecies variability (76).

The whole genome of the species type strain, C. hyointestinalis subsp. hyointestinalis DSM 19053, originally isolated in 1985 from the intestine of a pig, was sequenced and released in 2014. However, no further analyses of this genome were performed. Recently, two closed genomes belonging to the human strain C. hyointestinalis subsp. hyointestinalis LMG 9260 and the porcine strain C. hyointestinalis subsp. lawsonii LMG 15993 were released (77). Comparison of these genomes revealed an average nucleotide identity between the two subspecies lower than the standard threshold (95%) used for bacterial species delimitation, indicating that C. hyointestinalis subsp. hyointestinalis and C. hyointestinalis subsp. lawsonii could be reclassified as separate species within the genus Campylobacter. Indeed, a more recent study that performed whole-genome sequencing of 18 strains isolated from cattle, sheep, and deer from New Zealand confirmed the previously suggested genomewide plasticity of C. hyointestinalis (78). This work revealed high rates of gene gain/loss across C. hyointestinalis lineages, probably accounting for the effects of horizontal gene transfer. The presence of strains with an unusually high number of genes, multiple insertions of genomic islands, and a significant proportion of recombinant sites suggests that genomic introgression has been frequent along the evolutionary history of C. hyointestinalis subspecies. Indeed, pangenome estimations revealed that 67% of C. hyointestinalis genes belong to the core genome but only 5% of them correspond to the species clonal frame. Further variability has been reported in virulence-associated genes, like the gene for cytolethal distending toxin (CDT), for which new variants have been described in C. hyointestinalis (79, 80).

Comprehensive comparative genomic analyses including both subspecies have been limited because of the unavailability of C. hyointestinalis subsp. lawsonii genomes. However, the release of a set of nine C. hyointestinalis subsp. lawsonii genomes (81) has recently allowed proper comparison of the genetic diversity and evolutionary patterns distinguishing C. hyointestinalis subsp. hyointestinalis and C. hyointestinalis subsp. lawsonii. This study confirmed the phylogenetic separation of C. hyointestinalis subsp. hyointestinalis and C. hyointestinalis subsp. lawsonii using genomewide information and identified that the two subspecies have separately evolved with null or extremely limited gene flow between lineages. This genetic isolation is probably driven by adaptation to distinct ecological niches, which has determined the fixation of genomic signatures in both subspecies. For example, the generalist C. hyointestinalis subsp. hyointestinalis enclosed a bigger and more diverse accessory genome than the specialist C. hyointestinalis subsp. lawsonii. This increased diversity is probably driven by a stronger incidence of genomewide recombination events in C. hyointestinalis subsp. lawsonii than in C. hyointestinalis subsp. hyointestinalis. Accordingly, the genomes of both subspecies encode distinct repertories of CRISPR/Cas and restriction-modification systems that play an important role in DNA recombination, repair, and integration (49). This has probably influenced genome plasticity and led to the observed differences in the accessory genomes of both subspecies (82).

October 2019 Volume 32 Issue 4 e00072-18

Together, the observed genomic variability of *C. hyointestinalis* may indicate a great potential to adapt to different ecological niches, underpinning its capacity to colonize a great variety of mammal species both as a commensal and as a disease-causing agent (78). Further comparisons between commensal and disease-associated strains may provide new insight into the genetic mechanisms underlying *C. hyointestinalis* pathogenicity.

Campylobacter lanienae

Campylobacter lanienae was first described in 2000 from the feces of healthy individuals during a hygiene survey of abattoir workers (26). Subsequently, it has been recovered from healthy cattle, sheep, and swine (83). In a single report, *C. lanienae* was isolated from symptomatic infections in lab chinchillas with gastric ulcer; however, these results were not fully conclusive about the pathogenic role of this species (84). The fact that *C. lanienae* has not been reported to cause symptomatic infections either in humans or in other animals indicates that this species could have limited pathogenic potential or be a nonpathogenic member of the genus *Campylobacter*.

In a recent study, the whole genomes of 26 C. lanienae strains and 50 strains from three putative novel C. lanienae-related taxa isolated from diverse hosts (including humans, swine, sheep, and goat) were sequenced (85). In this work, the authors compared these genomes with those from other sister species, such as C. fetus, C. iguaniorum, and C. hyointestinalis, evidencing that C. lanienae and its related taxa present a reduced gene content and distinct CRISPR/Cas loci. Additionally, the C. lanienae lineage presented a higher diversity of flagellin genes than other Campylobacter species and the absence of genes involved in selenium metabolism. Beyond the unknown consequences of these genetic distinctions in the biology of C. lanienae, a more in-depth (not yet reported) comparison of these genomes could shed light on the apparent nonpathogenic phenotype of C. lanienae strains, as they could be screened for genes that are well-known virulence determinants in major Campylobacter pathogens. Anyway, this work represents a good example of how the sequencing of species that are not relevant for human or animal health can improve our understanding of host adaptation and virulence evolution in the pathogenic members of the genus Campylobacter.

Campylobacter lari and Related Taxa

Campylobacter lari was originally isolated from gulls as nalidixic acid-resistant, thermophilic strains (NARTC). Subsequently, the urease-producing thermophilic group (UPTC), the nalidixic acid-susceptible group (NASC), and the urease-positive NASC were identified as phenotypic variants of the originally described *C. lari* strains. A more comprehensive inspection of these variants using molecular typing methods resulted in the taxonomic revision and reclassification of several strains as novel taxa, such as *C. peloridis* (27) and *C. volucris* (38). Additionally, other *C. lari*-like species have recently been described, including *C. insulaenigrae* (24), *C. subantarcticus* (34), and *C. ornithocola* (29). Currently, *C. lari* is divided into two subspecies: *C. lari* subsp. *lari* and *C. lari* group. These bacteria are typically isolated from coastal regions, marine environments, molluscs, and aquatic birds and mammals. However, the sporadic isolation of *C. lari* and related taxa from human infections (86–89) highlights their potential as emerging pathogens.

Indeed, the first member of this group whose whole-genome sequence became available was *C. lari* subsp. *lari* strain RM2100, isolated from a girl with watery diarrhea. The analysis of this genome allowed for the determination that many virulence and antibiotic resistance mechanisms present in the major pathogen *C. jejuni* were also conserved in *C. lari*. Additionally, this isolate harbored a megaplasmid similar to the conjugative plasmid pTet found in *C. jejuni*, coding for type IV secretion systems, invasins, and adhesins that may contribute to its pathogenic potential (90).

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

A subsequent comparative analysis aiming to expand the characterization of the *C. lari* group released the whole-genome sequences for several UPTC strains, *C. lari* subsp. *concheus*, and other related taxa, such as *C. peloridis*, *C. subantarcticus*, *C. volucris*, and *C. insulaenigrae*. This study revealed that the *C. lari* group is very homogeneous, with more than 70% of the genes identified in the previously sequenced *C. lari* subsp. *lari* RM2100 being conserved among its members. However, an important conclusion of this work is the absence of genes or pathways potentially implicated in the association of the *C. lari* group with marine environments and aquatic animals (91).

Future work focused on elucidating this and other ecological aspects of this group, like its potential pathogenicity and zoonotic risk, should aim to generate comprehensive sets of whole-genome sequences for many species that are currently represented by just a single sequenced strain (most except *C. lari*). This would allow for the application of pangenome analyses providing a more comprehensive insight into the intraspecific genomic variation of the *C. lari* group. This could result in the development of lineage-specific molecular characterization tools useful to improve the screening of species belonging to the *C. lari* group in diverse environments and hosts, including the identification of clinical strains causing emerging infections in humans.

Campylobacter ureolyticus

In 2010, a polyphasic analysis was applied over a diverse collection of 26 *Bacteroides ureolyticus* strains to reassess the taxonomic position of this species. This study demonstrated that *B. ureolyticus* should be more suitably allocated within the genus *Campylobacter*; hence, it was renamed *C. ureolyticus* (37). Subsequently, a retrospective study that screened more than 7,000 patients with diarrhea evidenced the presence of *C. ureolyticus* in 23.8% of *Campylobacter*-positive samples, representing the first report of *C. ureolyticus* in the feces of patients with gastroenteritis and suggesting the role of this species as an emerging enteric pathogen (92).

The first whole-genome analysis of C. ureolyticus strains was published in 2013 and was based on the comparison of two genomes. The species type strain, DSM 20703 (originally isolated in 1978 from amniotic fluid), was sequenced in that study, and strain ACS-301-Sch-V-3b (isolated from the vaginal tract of a woman) had been previously sequenced as part of the Human Microbiome Project (93). This work uncovered the virulence gene repertories of C. *ureolyticus*, which resembled those present in other Campylobacter pathogens, including genes for adhesion and colonization (cadF, PEB1, icmF, and flpA), invasion (ciaB, type IV secretion systems), and toxin production (S layer, RTX, and Zot). Additionally, the study revealed that the two strains shared only 83% of their genes, suggesting considerable intraspecific heterogeneity within C. ureolyticus. Since then, only two additional genomes have been sequenced: that of strain CIT007, which was originally isolated from stools from an elderly woman presenting with diarrheal illness and end-stage chronic renal disease (94), and that of strain RIGS9880, which was isolated from an immunocompromised patient with diarrhea (95). Both strains presented very similar virulence repertories in comparison with the previously sequenced genomes, proposing that these genes are conserved features of C. ureolyticus that may define its pathogenic potential.

Considering the increasing clinical awareness of *C. ureolyticus* as an emerging pathogen causing human gastrointestinal disease, future research should focus on the generation of extensive whole-genome sequencing data from a representative collection of strains. This will allow the population structure, accessory gene dynamics, and selective pressures shaping the genomes of this pathogen to be uncovered. This information can be useful to seek genotype-phenotype associations, dissect its epidemiological behavior, and identify transmission patterns, which are largely unknown.

Campylobacter upsaliensis: a Salient Example of a Not (Enough) Sequenced Pathogen

For many nonclassical *Campylobacter* species reported to be causative agents of infections in humans and other animals, sequencing efforts have been extremely

October 2019 Volume 32 Issue 4 e00072-18

limited, often including just single genomes of type strains. This reflects the existing bias toward *Campylobacter* species that are more frequently or more easily isolated. Despite this general bias in bacterial genomics, the effort of sequencing nonclassical, clinically, or economically nonrelevant species deserves to be claimed as a way to improve our understanding of the mechanisms directing the evolution of bacterial pathogenicity.

C. upsaliensis represents a salient example of this situation. This species belongs to the thermophilic campylobacters and is phylogenetically close to C. jejuni. Despite C. upsaliensis being highlighted as an important emerging gastrointestinal pathogen more than 2 decades ago, its clinical underestimation has been mainly explained by the fact that it is sensitive to the antibiotics routinely used in selective media for the isolation of C. *jejuni* (96). Nevertheless, many studies carried out in different geographic areas have increasingly reported high prevalences of this species in human infections (6). Also, companion animals, such as dogs and cats, have been identified to be possible reservoirs of C. upsaliensis (97). Indeed, a pioneering study that applied amplified fragment length polymorphism (AFLP) to characterize C. upsaliensis strains isolated from humans and dogs evidenced two main genomic clusters: one exclusively composed of human strains and the other comprising both human and dog strains (98). This work revealed that, despite dogs being a possible source for human infections, other routes of transmission would explain most cases of C. upsaliensis in humans. Undoubtedly, whole-genome sequencing of C. upsaliensis populations isolated from humans, dogs, and other animals could provide an enhanced understanding of its epidemiology, host-associated evolution, and virulence mechanisms.

EVOLUTIONARY MECHANISMS IN EMERGING CAMPYLOBACTER SPECIES

The evolution of bacterial populations is directed by the incidence of two main mechanisms: DNA damage or replication errors which generate deletions, rearrangements, and point mutations and the horizontal exchange of genetic material, through which genes are externally acquired and eventually incorporated through recombination. The impact of these diversification mechanisms in the structuring of bacterial populations can be explained under entirely neutral evolutionary models, where microorganisms do not differ in their fitness and all members of the population can be tracked back to the most recent common ancestor using the coalescent framework. Though neutral models can explain the population structure, bacteria are probably under selection pressures for adaptive traits which influences their fitness, specifically, in host-adapted species which are exposed to particular conditions within hosts. Some Campylobacter species can survive in the environment, but they are mainly found in association with vertebrate hosts; hence, they are presumably subjected to selective pressures whose traces can be observed as genomic signatures. Here, we summarize and discuss how pathogenomic studies have contributed to unveil the main evolutionary mechanisms in emerging Campylobacter species and their clinical relevance as drivers of new pathogenic phenotypes.

Host-Associated Population Structure and Adaptation

Emerging *Campylobacter* species are mainly host adapted and colonize a wide variety of niches within birds, mammalian, and reptilian hosts. Accordingly, the observation of coexisting lineages associated with different hosts suggests that natural selection acts to maintain that given population structure, for example, human-associated and cattle-associated *C. fetus* strains that represent different phylogenetic lineages that exhibit host-specific core gene repertories under positive selection (64). In particular, strong positive selection signals in the *flgD* gene (coding for the flagellar hook cap protein FlgD) were found in human-adapted lineages. Interestingly, diversifying alleles of this gene have been found to be a defining feature of hyperinvasive *C. jejuni* strains (99). In cattle-associated lineages, the enterobactin uptake receptor *cfrA* has been identified to be the most diversifying gene. The expression of the CfrA protein is induced under iron-restricted conditions and plays an important role in iron scavenging and colonization in *C. jejuni* (100), suggesting that selection acting on *cfrA* could

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

Clinical Microbiology Reviews

be associated with niche adaptation and virulence in *C. fetus*. Another example of host adaptation is observed in mammal- and reptile-associated *C. fetus* strains that represent phylogenetically distinct lineages. Remarkably, the *tcuRABC* operon, which directs the catabolism of tricarballylate, allowing its utilization as a carbon and energy source, is present in the reptile-associated lineage consisting of *C. fetus* subsp. *testudinum* strains but absent in the mammal-associated lineage consisting of *C. fetus* subsp. *fetus* and *C. fetus* subsp. *venerealis* (71). This genomic signature reflects functional adaptation to different ecological niches within mammal and reptile hosts.

Similarly, other emerging species possess a structured population. An example is *C. hyointestinalis*, which is subdivided into *C. hyointestinalis* subsp. *hyointestinalis* (a generalist colonizing several mammalian species) and *C. hyointestinalis* subsp. *lawsonii* (a specialist adapted to pigs), which represent clearly divergent phylogenetic lineages (82). The genomes of these host-adapted subspecies are characterized by distinct accessory gene patterns which reflect dissimilarities in their functional repertories. Specifically, genes involved in DNA replication, recombination, and repair, such as those coding for CRISPR/Cas and restriction-modification systems, are unevenly distributed among subspecies. These genes may modulate the generation of sequence diversity that is the source for natural selection and the subsequent adaptation of lineages to ecological niches, like different host species. These observations are reminiscent of the major pathogen *C. jejuni*, where the differential host tropism of lineages has been well documented. For example, differences in the vitamin B₅ biosynthesis pathway have been identified between cattle-associated and chicken-associated clonal complexes, suggesting adaptation to the host diet (101).

As hosts represent a complex combination of selective pressures given by distinct immune responses along tissues or fluctuating concentrations of metabolites and cell by-products, different Campylobacter genotypes could adapt to distinct ecological subniches within the same host organism and eventually exhibit distinct virulent phenotypes. Indeed, this kind of adaptation to subniches within the same host has been proposed in C. jejuni, where up to 10 different clonal complexes have been reported coexisting in a single chicken flock (102). Interestingly, this has been reported in emerging species, such as C. fetus subsp. testudinum, which also shows signatures of genomic adaptation to different subniches within the same host. Specifically, a recombination event in the *iamA* gene has been identified among strains causing invasive disease in humans (recovered from blood, bile, hematoma, or pleural fluid) but absent in strains isolated from stool samples. The iamA gene belongs to an ABC transporter system that is considered a virulence factor associated with invasion in C. jejuni (103). These examples evidence that genomic traces in host-adapted lineages within structured populations can be identified at different levels of complexity and could be related to the pathogenic potential of emerging Campylobacter species.

Barriers to Homologous Recombination

Recombination occurs between individual microorganisms, so it can be detected when comparative analyses are performed at the population level through the identification of genomic mosaicisms. This evidences gene flow occurring between physically close cells, hence, the absence of an ecological barrier between them. However, different types of barriers to homologous recombination can be implicated in a maintaining population structure, particularly in host-associated lineages, where strains colonizing a certain host are typically not in contact with those colonizing a different host. For example, the *C. hyointestinalis* subspecies found in different mammalian species are separated by a strong recombination barrier (82). Interestingly, these subspecies present a borderline average nucleotide identity indicating an underlying speciation process driven by genetic and ecological isolation. Also, in *C. fetus* a barrier to homologous recombination between mammal-adapted and reptile-adapted strains reflects that these lineages have been evolving separately for a long time in association with different host species (71).

This phenomenon can have adaptive explanations, since bacteria sharing similar niches will require certain combinations of genes that confer a fitness advantage in that

October 2019 Volume 32 Issue 4 e00072-18

environment. Recombination can be the underlying mechanism by which adaptive genetic variants can be incorporated and selected. A clear example of recombination barriers and differential gene flow between closely related *Campylobacter* lineages can be observed in *C. coli*. Barriers to homologous recombination have been identified in *C. coli* clade 1 strains, which mainly represent clinical and farm animal isolates, with respect to clade 2 and 3 strains, which are more abundant in watercourses and riparian environments. On the contrary, substantial genome introgression from *C. jejuni* has been detected in *C. coli* clade 1 after a long period of independent evolution (104).

These examples evidence that limited gene flow between *Campylobacter* populations may be a mechanism driving genomic diversification and possible adaptation to new hosts in emerging species. Some similarities can be observed in emerging and major *Campylobacter* pathogens, where recombination barriers and gene flow have shaped their genome dynamics. From the clinical perspective, understanding how genomic admixture underpins the adaptation of pathogens to new environments and/or hosts is relevant, since this information can help to provide an understanding of phenotypic variation associated with pathogenicity potential, identify new reservoirs for zoonotic species, and predict potential host jumps that can lead to emerging infections.

Horizontal Gene Transfer

Horizontal gene transfer is a main driver of bacterial evolution, and its role in genome plasticity is being understood more precisely as new genomic information becomes available for different bacterial lineages. The impact of horizontal gene transfer in *Campylobacter* evolution has been well-documented in *C. jejuni*, where it has played an important role in the acquisition of antimicrobial resistance and virulence. Additionally, horizontal gene transfer, including the incorporation of plasmids and the integration of genomic islands, has been described in several emerging *Campylobacter* species. Here, we summarize the most salient examples of horizontal gene transfer events in emerging campylobacters and discuss how this can impact antimicrobial resistance and virulence.

Role in the acquisition of antimicrobial resistance. Antimicrobial resistance conferred by horizontal gene transfer has been largely documented mainly in C. ieiuni and C. coli (105). Indeed, as Campylobacter species possess genetic mechanisms for natural transformation and conjugation, antimicrobial resistance genes could be rapidly transferred between strains (106). For example, C. jejuni isolates carrying a plasmid coding for a cfr(C) gene that confers multidrug resistance were recently described. Interestingly, this genetic mechanism was also found in C. coli primary isolates recovered from cattle (107). Considering the widespread distribution of several emerging Campylobacter species in the farm environment, this constitutes a risk for the appearance of new Campylobacter lineages that could incorporate this plasmid. The fact that this plasmid can also be successfully transferred in vitro between C. jejuni and C. coli suggests that it could disseminate to other species coexisting in the same niche. Indeed, the analysis of C. hepaticus and C. lari genomes, among other emerging Campylobacter genomes, has recently revealed the presence of plasmids coding for the tetracycline resistance gene tetO and other antimicrobial resistance determinants which are widespread in mobile elements found in C. jejuni and C. coli (74, 90). Additionally, other mechanisms of horizontal gene transfer seem to be important in emerging Campylobacter species, as evidenced by the detection of tetracycline and aminoglycoside resistance genes within a transferable genomic island found in C. fetus (108). This suggests that the intraand interspecies transfer of mobile elements coding for antimicrobial resistance is possible between Campylobacter species. Importantly, as most Campylobacter infections in humans are caused by the ingestion of contaminated animal products or contact with animals, the dissemination and fixation of antimicrobial resistance mechanisms in emerging campylobacters that are mainly adapted to the farm environment constitute a possible vehicle for the appearance of emerging pathogens with extended antimicrobial resistance repertories. This is particularly relevant, since the same classes

October 2019 Volume 32 Issue 4 e00072-18

Clinical Microbiology Reviews

of antimicrobials are basically being used in food-producing animals and in human medicine, generating similar selective pressures in both environments.

Role in the acquisition of virulence mechanisms. Mobile elements are important for virulence in major pathogens like C. jejuni. Particularly, this is mainly caused by the presence of plasmids containing genes homologous to the genes for type IV secretion systems (T4SS), which are macromolecular machineries used to exchange DNA between bacteria but also to inject protein effectors into host cells, which can lead to functional impairment (109). Indeed, experimental mutations to inactivate T4SS genes caused reduced adherence and invasion in C. jejuni (110). Interestingly, these systems have subsequently been described in other Campylobacter species, like C. fetus, where diverse T4SS clusters are present in their genomes and are contained within pathogenicity islands that can be found both integrated into the chromosome and in plasmids. Even though the effector proteins that could be delivered by these T4SS and cause damage to the host cell have not been experimentally determined, T4SS-containing pathogenicity islands in C. fetus also code for filamentation induced by cAMP (FIC) domain proteins. These proteins could be potential effectors for virulence since in other bacteria they have critical roles in cellular processes, including disruption of host cell signaling pathways leading to cytotoxic effects (111).

Phylogenetic characterization of T4SS genes in *C. fetus* revealed multiple evolutionary origins and acquisition from different *Campylobacter* donors. These genes are also conserved in other emerging species, like *C. ureolyticus*, *C. upsaliensis*, and *C. lari*, indicating that horizontal gene transfer of plasmids and other genetic elements harboring T4SS has been an important evolutionary force shaping the repertory of virulence genes in *Campylobacter* species. The exploration of T4SS diversity in other *Campylobacter* genomes, together with the development of experimental approaches in emerging species, can provide important information about the role of these secretion systems in the virulence of *Campylobacter* species. Importantly, the identification of genes coding for effector proteins that modulate virulence could lead to the development of straightforward molecular typing tools targeting these genes, as they have been developed, for example, to characterize *cagA*-positive *Helicobacter pylori* strains, which are known to increase the risk of developing gastric cancer (112, 113).

Genome Reduction

Genome reduction has been documented in diverse bacterial lineages and is typically associated with functional specialization, host association, and/or increased pathogenicity in some species (114, 115), since niche adaptation requires selection for traits that optimize pathogen fitness in the new environment. Reduced genomes are also characterized by diminutive gene sets with a loss of metabolic functions and a low genomic GC content. The genome size in the genus Campylobacter varies from \sim 1.4 Mb to \sim 2.5 Mb, and the GC content ranges from very low values of about 28% up to 45% (116). The most relevant species for human and animal health, C. jejuni and C. coli, have genomes smaller than those of most of the emerging species and are typically associated with the gastrointestinal niche. This highlights the relevance of genome reduction during the evolution of pathogenicity in the genus Campylobacter. However, the most salient example of genome reduction is the recently described species C. hepaticus, which presents one of the lowest GC values within the genus (~28%) and a smaller genome than its closest relative, C. jejuni. The genes lost in this species include genes for pathways for iron acquisition and metabolism, which is consistent with adaptation to an iron-rich environment, such as the chicken liver, which constitutes its reservoir. Additionally, C. hepaticus genomes code for a very reduced repertory of virulence-associated genes in comparison to that in C. jejuni, with the C. hepaticus genome lacking many well-known genetic factors important during Campylobacter infection, like those coding for the cytolethal distending toxin (CDT) and capsular and extracellular polysaccharides (74). This reduced set of virulence genes may be the consequence of the evolution of attenuated virulence in C. hepaticus, which has been documented in other bacteria and which occurs as a result of immune evasion within the host (117). Beyond being beneficial for the pathogen due to the establishment of a

October 2019 Volume 32 Issue 4 e00072-18



FIG 2 Campylobacter taxa and available genomes through time. (A) Graph showing the number of described Campylobacter species through time since the official description of the genus in 1973. (B) Graph showing the number of available whole-genome sequences through time for C. jejuni/C. coli and the rest of the Campylobacter members, including those emerging species discussed in this review.

long-term chronic infection, this may constitute a limited capacity for transmission within hosts or survival in the environment. Consequently, further work is needed to determine the impact of genome reduction in *C. hepaticus* and how this correlates with virulent phenotypes, which would constitute a cornerstone for studying this evolutionary mechanism in this and other emerging *Campylobacter* species.

FUTURE DIRECTIONS

Beyond the fact that C. jejuni and C. coli still remain the main causes of bacterial gastroenteritis worldwide, we now identify most Campylobacter species to be clinically relevant pathogens. Among them, C. fetus, C. concisus, C. ureolyticus, and C. upsaliensis are being systematically detected from humans, although the increasingly frequent report of infections caused by other related species still has not motivated a unified effort to understand the genomic variability of emerging Campylobacter species. On the contrary, these efforts have been mostly isolated, causing some important species, such as C. upsaliensis, to remain neglected. Accordingly, our understanding about host adaptation, transmission, and pathogenicity evolution in Campylobacter species could be enhanced by coordinating sequencing efforts among international groups, working as a consortium to fill the gap in genomic information for emerging campylobacters that exists today. As for many major pathogens, it is clear that the availability of genomic data and the information derived from subsequent comparative analyses constitutes a major opportunity for an improved understanding of the mechanisms driving Campylobacter evolution. In particular, this could facilitate the development of high-resolution tools for typing emerging species. For example, the development of core genome multilocus sequence typing (cgMLST) schemes like those currently available for C. jejuni and C. coli requires comprehensive genomic data sets that represent global isolate collections. Once cgMLST schemes, among other tools, become available for species like C. fetus, C. ureolyticus, C. concisus, or C. upsaliensis, understanding the genetic variation that underpins ecological and epidemiological patterns will be straightforward, enabling the rapid identification and more efficient tracking of emerging campylobacters, which will eventually result in more effective interventions. In parallel, the identification of genetic features associated with virulence traits will guide the development of new therapeutic procedures to prevent or control infections caused by species different from C. jejuni and C. coli, contributing to limiting the emergence and spread of new clinically relevant genetic variants.

The description of novel *Campylobacter* species has accelerated since 2009 (Fig. 2), mainly due to improved culturing conditions for *Campylobacter* and the exploration of new hosts and environments. Also, the recent application of culture-free methods, like 16S rRNA

October 2019 Volume 32 Issue 4 e00072-18

cmr.asm.org 17

47

amplicon sequencing, to study the microbiome of wild animals has allowed the identification of novel species, like C. pinnipediorum (30). This approach seems promising to discover new species but also to monitor the presence of emerging pathogens without the need for bacterial isolation, which can provide useful epidemiological information to guide more comprehensive sequencing efforts. Also, metagenomics can be applied to have a general overview of the human microbiome composition at the population level by analyzing samples from urban environments (118). For example, as the urban sewage microbiome has been shown to recapitulate the human gut microbiome, metagenomic sequencing of wastewater samples can provide information about the circulation of pathogens and antimicrobial-resistant genotypes in the population (119), including unappreciated Campylobacter species. Furthermore, as many emerging campylobacters (like C. concisus) have been found in healthy individuals, the reanalysis of thousands of public gut metagenomic data from healthy humans can help to recover genomic information from Campylobacter lineages that are being carried asymptomatically. This approach has proved useful to identify and quantify C. fetus lineages in the intestinal microbiota of healthy humans (64) and could be easily extended to analyze other species. Together, the application of integrative approaches that consider improved culture conditions, whole-genome sequencing, and microbiome analysis can set the basis for future research toward the elucidation of the epidemiological behavior and real clinical impact of emerging Campylobacter species.

CONCLUSIONS

Not every infectious disease is subject to epidemiological research. Hence, the resultant lack of information can lead to the unappreciated spread of pathogens which constitute public health emergencies. Indeed, this originally happened with C. jejuni, whose importance in human gastroenteritis was unseen until methodological improvements for isolation and identification became available. In this sense, emerging Campylobacter species that have been reported at a high frequency in certain geographic regions, like C. upsaliensis, are known to be underestimated because they are susceptible to some antibiotic combinations used in selective media for the isolation of C. jejuni and C. coli. This represents an important challenge, since the application of high-resolution tools to characterize these infections, like those based in comparative genomics, require the availability of bacterial isolate collections and clinical information. However, we envisage that constant improvements in clinical microbiology will gradually uncover the real burden of many emerging Campylobacter species. Indeed, this has already happened for some species, like C. fetus, which is currently recognized as the most frequent *Campylobacter* species recovered from human blood infections, after being largely underestimated and considered an infrequent pathogen. Thus, this underpinned the recent development of pathogenomic analyses that resulted in a better understanding of its epidemiology, evolution, and transmission.

Upcoming efforts focused on understanding the biology of emerging campylobacters will require the extensive use of pathogenomic approaches to comprehensively address a set of key unpostponable aspects, including the genomic screening of global collections of main emerging species, like *C. upsaliensis, C. ureolyticus*, or *C. concisus*, and the incorporation of microbiome data analysis to recover asymptomatically carried lineages or to evaluate the incidence of *Campylobacter* infection in the host microbiota composition. The outcomes of these studies will assist with the rational selection of genomic markers for the design of novel diagnostic assays, the development of tools for molecular epidemiology, and the identification of phenotype-genotype associations that can at last have an impact on improving the surveillance, control, and treatment of infections caused by emerging *Campylobacter* species.

REFERENCES

1. Smibert RM. 1978. The genus Campylobacter. Annu Rev Microbiol 32:673–709. https://doi.org/10.1146/annurev.mi.32.100178.003325.

October 2019 Volume 32 Issue 4 e00072-18

2. Smith T, Taylor MS. 1919. Some morphological and biological characters of the *Spirilla (Vibrio Fetus*, n. sp.) associated with disease of the

cmr.asm.org 18

Clinical Microbiology Reviews

fetal membranes in cattle. J Exp Med 30:299-311. https://doi.org/10 .1084/jem.30.4.299.

- Véron M, Chatelain R. 1973. Taxonomic study of the genus *Campylobacter* Sebald and Véron and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Véron. Int J Syst Evol Microbiol 23:122–134. https://doi.org/10.1099/00207713 -23-2-122.
- Guerrant RL, Lahita RG, Winn WC, Roberts RB. 1978. Campylobacteriosis in man: pathogenic mechanisms and review of 91 bloodstream infections. Am J Med 65:584–592. https://doi.org/10.1016/0002-9343(78)90845-8.
- Kahler RL, Sheldon H. 1960. Vibrio fetus infection in man. N Engl J Med 262:1218–1222. https://doi.org/10.1056/NEJM196006162622404.
- Man SM. 2011. The clinical importance of emerging Campylobacter species. Nat Rev Gastroenterol Hepatol 8:669–685. https://doi.org/10 .1038/nrgastro.2011.191.
- Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream M-A, Rutherford KM, van Vliet AHM, Whitehead S, Barrell BG. 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable seguences. Nature 403:665–668. https://doi.org/10.1038/35001088.
- Rossi M, Debruyne L, Zanoni RG, Manfreda G, Revez J, Vandamme P. 2009. Campylobacter avium sp. nov., a hippurate-positive species isolated from poultry. Int J Syst Evol Microbiol 59:2364–2369. https://doi .org/10.1099/ijs.0.007419-0.
- Gilbert MJ, Zomer AL, Timmerman AJ, Spaninks MP, Rubio-García A, Rossen JW, Duim B, Wagenaar JA. 2018. *Campylobacter blaseri* sp. nov., isolated from common seals (*Phoca vitulina*). Int J Syst Evol Microbiol 68:1787–1794. https://doi.org/10.1099/ijsem.0.002742.
- Inglis GD, Hoar BM, Whiteside DP, Morck DW. 2007. Campylobacter canadensis sp. nov., from captive whooping cranes in Canada. Int J Syst Evol Microbiol 57:2636–2644. https://doi.org/10.1099/ijs.0.65061-0.
- 11. Tanner ACR, Badger S, Lai C-H, Listgarten MA, Visconti RA, Socranksy SS. 1981. Wolinella gen. nov., Wolinella succinogenes (Vibrio succinogenes Wolin et al.) comb. nov., and description of Bacteroides gracilis sp. nov., Wolinella recta sp. nov., Campylobacter concisus sp. nov., and Eikenella corrodens from humans with periodontal disease. Int J Syst Evol Microbiol 31:432–445. https://doi.org/10.1099/00207713-31-4-432.
- Koziel M, O'Doherty P, Vandamme P, Corcoran GD, Sleator RD, Lucey B. 2014. Campylobacter corcagiensis sp. nov., isolated from faeces of captive lion-tailed macaques (*Macaca silenus*). Int J Syst Evol Microbiol 64: 2878–2883. https://doi.org/10.1099/ijs.0.663867-0.
- Zanoni RG, Debruyne L, Rossi M, Revez J, Vandamme P. 2009. Campylobacter cuniculorum sp. nov., from rabbits. Int J Syst Evol Microbiol 59:1666–1671. https://doi.org/10.1099/ijs.0.007286-0.
- Abbott SL, Waddington M, Lindquist D, Ware J, Cheung W, Ely J, Janda JM. 2005. Description of *Campylobacter curvus* and *C. curvus*-like strains associated with sporadic episodes of bloody gastroenteritis and Brainerd's diarrhea. J Clin Microbiol 43:585–588. https://doi.org/10.1128/JCM .43.2.585-588.2005.
- Fitzgerald C, Tu Z, 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. https:// doi.org/10.1099/ijs.0.057778-0.
- Piccirillo A, Niero G, Calleros L, Pérez R, Naya H, Iraola G. 2016. Campylobacter geochelonis sp. nov. isolated from the western Hermann's tortoise (*Testudo hermanni hermanni*). Int J Syst Evol Microbiol 66: 3468–3476. https://doi.org/10.1099/ijsem.0.001219.
- Vandamme P, Daneshvar MI, Dewhirst FE, Paster BJ, Kersters K, Goossens H, Moss CW. 1995. Chemotaxonomic analyses of *Bacteroides gracilis* and *Bacteroides ureolyticus* and reclassification of *B. gracilis* as *Campylobacter gracilis* comb. nov. Int J Syst Evol Microbiol 45:145–152. https://doi.org/10.1099/00207713-45-1-145.
- Stanley J, Burnens AP, Linton D, On SLW, Costas M, Owen RJ. 1992. *Campylobacter helveticus* sp. nov., a new thermophilic species from domestic animals: characterization, and cloning of a species-specific DNA probe. Microbiology 138:2293–2303. https://doi.org/10.1099/00221287 -138-11-2293.
- Van TTH, Elshagmani E, Gor MC, Scott PC, Moore RJ. 2016. Campylobacter hepaticus sp. nov., isolated from chickens with spotty liver disease. Int J Syst Evol Microbiol 66:4518–4524. https://doi.org/10.1099/ijsem.0.001383.
 Lawson AJ, On SL, Logan JM, Stanley J. 2001. Campylobacter hominis sp.

tinalis (new species) isolated from swine with lesions of proliferative iletis. Am J Vet Res 44:361–367.

nov., from the human gastrointestinal tract. Int J Syst Evol Microbiol

- On SLW, Block B, Holmes B, Hoste B, Vandamme P. 1995. Campylobacter hyointestinalis subsp. lawsonii subsp. nov., isolated from the porcine stomach, and an emended description of Campylobacter hyointestinalis. Int J Syst Evol Microbiol 45:767–774. https://doi.org/10.1099/00207713 -45-4-767.
- Gilbert MJ, Kik M, Miller WG, Duim B, Wagenaar JA. 2015. Campylobacter iguaniorum sp. nov., isolated from reptiles. Int J Syst Evol Microbiol 65:975–982. https://doi.org/10.1099/ijs.0.000048.
- Foster G, Holmes B, Steigerwalt AG, Lawson PA, Thorne P, Byrer DE, Ross HM, Xerry J, Thompson PM, Collins MD. 2004. *Campylobacter insulaenigrae* sp. nov., isolated from marine mammals. Int J Syst Evol Microbiol 54:2369–2373. https://doi.org/10.1099/ijs.0.63147-0.
- Steele TW, Owen RJ. 1988. Campylobacter jejuni subsp. doylei subsp. nov., a subspecies of nitrate-negative campylobacters isolated from human clinical specimens. Int J Syst Evol Microbiol 38:316–318. https:// doi.org/10.1099/00207713-38-3-316.
- Logan JM, Burnens A, Linton D, Lawson AJ, Stanley J. 2000. Campylobacter lanienae sp. nov., a new species isolated from workers in an abattoir. Int J Syst Evol Microbiol 50:865–872. https://doi.org/10.1099/ 00207713-50-2-865.
- Debruyne L, On SLW, De Brandt E, Vandamme P. 2009. Novel Campylobacter lari-like bacteria from humans and molluscs: description of Campylobacter peloridis sp. nov., Campylobacter lari subsp. concheus subsp. nov. and Campylobacter lari subsp. lari subsp. nov. Int J Syst Evol Microbiol 59:1126–1132. https://doi.org/10.1099/ijs.0.000851-0.
- Roop RM, Smibert RM, Johnson JL, Krieg NR. 1985. Campylobacter mucosalis (Lawson, Leaver, Pettigrew, and Rowland 1981) comb. nov.: emended description. Int J Syst Evol Microbiol 35:189–192. https://doi .org/10.1099/00207713-35-2-189.
- Cáceres A, Muñoz I, Iraola G, Díaz-Viraqué F, Collado L. 2017. Campylobacter ornithocola sp. nov., a novel member of the Campylobacter lari group isolated from wild bird faecal samples. Int J Syst Evol Microbiol 67:1643–1649. https://doi.org/10.1099/ijsem.0.001822.
- Gilbert MJ, Miller WG, St Leger J, Chapman MH, Timmerman AJ, Duim B, Foster G, Wagenaar JA. 2017. Campylobacter pinnipediorum sp. nov., isolated from pinnipeds, comprising Campylobacter pinnipediorum subsp. pinnipediorum subsp. nov. and Campylobacter pinnipediorum subsp. caledonicus subsp. nov. Int J Syst Evol Microbiol 67:1961–1968. https://doi.org/10.1099/ijsem.0.001894.
- Vandamme P, Falsen E, Rossau R, Hoste B, Segers P, Tytgat R, De Ley J. 1991. Revision of *Campylobacter*, *Helicobacter*, and *Wolinella* taxonomy: emendation of generic descriptions and proposal of *Arcobacter* gen. nov. Int J Syst Evol Microbiol 41:88–103. https://doi.org/10.1099/00207713-41 -1-88.
- Etoh Y, Dewhirst FE, Paster BJ, Yamamoto A, Goto N. 1993. Campylobacter showae sp. nov., isolated from the human oral cavity. Int J Syst Bacteriol 43:631–639. https://doi.org/10.1099/00207713-43-4-631.
- 33. On SLW, Atabay HI, Corry JEL, Harrington CS, Vandamme P. 1998. Emended description of *Campylobacter sputorum* and revision of its infrasubspecific (biovar) divisions, including *C. sputorum* biovar *paraureolyticus*, a urease-producing variant from cattle and humans. Int J Syst Evol Microbiol 48:195–206. https://doi.org/10.1099/00207713-48-1-195.
- Debruyne L, Broman T, Bergström S, Olsen B, On SLW, Vandamme P. 2010. Campylobacter subantarcticus sp. nov., isolated from birds in the sub-Antarctic region. Int J Syst Evol Microbiol 60:815–819. https://doi .org/10.1099/ijs.0.011056-0.
- Kaur T, Singh J, Huffman MA, Petrželková KJ, Taylor NS, Xu S, Dewhirst FE, Paster BJ, Debruyne L, Vandamme P, Fox JG. 2011. *Campylobacter troglodytis* sp. nov., isolated from feces of human-habituated wild chimpanzees (*Pan troglodytes schweinfurthii*) in Tanzania. Appl Environ Microbiol 77:2366–2373. https://doi.org/10.1128/AEM.01840-09.
- Sandstedt K, Ursing J. 1991. Description of *Campylobacter upsaliensis* sp. nov. previously known as the CNW group. Syst Appl Microbiol 14:39–45. https://doi.org/10.1016/S0723-2020(11)80359-0.
- Vandamme P, Debruyne L, De Brandt E, Falsen E. 2010. Reclassification of *Bacteroides ureolyticus* as *Campylobacter ureolyticus* comb. nov., and emended description of the genus *Campylobacter*. Int J Syst Evol Microbiol 60:2016–2022. https://doi.org/10.1099/ijs.0.017152-0.
- 38. Debruyne L, Broman T, Bergström S, Olsen B, On SLW, Vandamme P.

2010. Campylobacter volucris sp. nov., isolated from black-headed gulls (Larus ridibundus). Int J Syst Evol Microbiol 60:1870-1875. https://doi .org/10.1099/ijs.0.013748-0.

- 39. Zhang L, Budiman V, Day AS, Mitchell H, Lemberg DA, Riordan SM, Grimm M, Leach ST, Ismail Y. 2010. Isolation and detection of Campy lobacter concisus from saliva of healthy individuals and patients with inflammatory bowel disease. J Clin Microbiol 48:2965-2967. https://doi org/10.1128/JCM.02391-09.
- 40. Zhang L, Man SM, Day AS, Leach ST, Lemberg DA, Dutt S, Stormon M, Otley A, O'Loughlin EV, Magoffin A, Ng PHY, Mitchell H. 2009. Detection and isolation of Campylobacter species other than C. jejuni from children with Crohn's disease. J Clin Microbiol 47:453-455. https://doi.org/ 0.1128/JCM.01949-08
- 41. Mahendran V, Riordan SM, Grimm MC, Tran TAT, Major J, Kaakoush NO, Mitchell H, Zhang L. 2011. Prevalence of Campylobacter species in adult Crohn's disease and the preferential colonization sites of Campylobacter Species in the human intestine. PLoS One 6:e25417. https:// oi.org/10.1371/journal.pone.0025417.
- 42. Man SM, Zhang L, Day AS, Leach ST, Lemberg DA, Mitchell H. 2010. Campylobacter concisus and other Campylobacter species in children with newly diagnosed Crohn's disease. Inflamm Bowel Dis 16: 1008-1016. https://doi.org/10.1002/ibd.21157.
- 43. Matsheka MI, Elisha BG, Lastovica AL, On SL. 2002. Genetic heterogeneity of Campylobacter concisus determined by pulsed field gel electrophoresis-based macrorestriction profiling. FEMS Microbiol Lett 211:17-22. https://doi.org/10.1016/S0378-1097(02)00667-5.
- 44. Deshpande NP, Kaakoush NO, Mitchell H, Janitz K, Raftery MJ, Li SS, Wilkins MR. 2011. Sequencing and validation of the genome of a Campylobacter concisus reveals intra-species diversity. PLoS One 6:e22170. https://doi.org/10.1371/journal.pone.0022170.
- 45. Kaakoush NO, Deshpande NP, Wilkins MR, Raftery MJ, Janitz K, Mitchell H. 2011. Comparative analyses of Campylobacter concisus strains reveal the genome of the reference strain BAA-1457 is not representative of the species. Gut Pathog 3:15. https://doi.org/10.1186/1757-4749-3-15. 46. Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ. 2007. Type VI
- secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc Natl Acad Sci U S A 104:15508-15513. ttps://doi.org/10.1073/pnas.0706532104.
- 47. Chung HKL, Tay A, Octavia S, Chen J, Liu F, Ma R, Lan R, Riordan SM, Grimm MC, Zhang L. 2016. Genome analysis of Campylobacter concisus strains from patients with inflammatory bowel disease and gastroenteritis provides new insights into pathogenicity. Sci Rep 6:38442. https://doi.org/10 .1038/srep38442
- 48. Zhang L, Lee H, Grimm MC, Riordan SM, Day AS, Lemberg DA. 2014. Campylobacter concisus and inflammatory bowel disease. World J Gastroenterol 20:1259-1267. https://doi.org/10.3748/wjg.v20.i5.1259.
- 49. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315:1709-1712. https://doi.org/ 10.1126/science.1138140.
- 50. Parsonnet J, Friedman GD, Orentreich N, Vogelman H. 1997. Risk for gastric cancer in people with CagA positive or CagA negative Helicobacter pylori infection. Gut 40:297-301. https://doi.org/10.1136/gut.40 3.297.
- 51. Fries BC, Varshney AK. 2013. Bacterial toxins-staphylococcal enterotoxin B. Microbiol Spectr 1:AID-0002-2012. https://doi.org/10.1128/ nicrobiolspec.AID-0002-2012.
- 52. Liu F, Ma R, Tay CYA, Octavia S, Lan R, Chung HKL, Riordan SM, Grimm MC, Leong RW, Tanaka MM, Connor S, Zhang L. 2018. Genomic analysis of oral *Campylobacter concisus* strains identified a potential bacterial molecular marker associated with active Crohn's disease. Emerg Microbes Infect 7:64. https://doi.org/10.1038/s41426-018-0065-6
- 53. Kirk KF, Méric G, Nielsen HL, Pascoe B, Sheppard SK, Thorlacius-Ussing O, Nielsen H. 2018. Molecular epidemiology and comparative genomics of Campylobacter concisus strains from saliva, faeces and gut mucosal biopsies in inflammatory bowel disease. Sci Rep 8:1902. https://doi.org/ 10.1038/s41598-018-20135-4.
- 54. Wang Y, Liu F, Zhang X, Chung HKL, Riordan SM, Grimm MC, Zhang S, Ma R, Lee SA, Zhang L. 2017. Campylobacter concisus genomospecies 2 is better adapted to the human gastrointestinal tract as compared with Campylobacter concisus genomospecies 1. Front Physiol 8:543. https://doi.org/10.3389/fphys.2017.00543.
- 55. Ma R, Liu F, Yap SF, Lee H, Leong RW, Riordan S, Grimm M, Zhang L. 2018. The growth and protein expression of inflammatory bowel

October 2019 Volume 32 Issue 4 e00072-18

disease-associated Campylobacter concisus is affected by the derivatives of the food additive fumaric acid. Front Microbiol 9:896. https:// oi.org/10.3389/fmicb.2018.00896

- 56. Moolhuijzen PM, Lew-Tabor AE, Wlodek BM, Agüero 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. https:// doi.org/10.1186/1471-2180-9-86.
- 57. Iraola G, Hernández M, Calleros L, Paolicchi F, Silveyra S, Velilla A, Carretto L, Rodríguez E, Pérez 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. https://doi.org/10.4142/jvs.2012.13.4.371.
- 58. Bloois L, van der G, 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. https://doi.org/10.1371/journal.pone 0152832
- 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:e01246-13. https://doi.org/10.1128/genomeA.01246-13.
- 60. Iraola G, Pérez R, Naya H, Paolicchi F, Harris D, Lawley TD, Rego N, Hernández M, Calleros L, Carretto L, Velilla A, Morsella C, Méndez A, Gioffre A. 2013. Complete genome sequence of Campylobacter fetus subsp. venerealis biovar intermedius, isolated from the prepuce of a bull. Genome Announc 1:e00526-13. https://doi.org/10.1128/genomeA .00526-13.
- Bloois L, van der G, Miller WG, Yee E, Rijnsburger M, Wagenaar JA, Duim B. 2014. Inconsistency of phenotypic and genomic characteristics of Campylobacter fetus subspecies requires reevaluation of current diagnostics. J Clin Microbiol 52:4183-4188. https://doi.org/10.1128/JCM .01837-14.
- 62. Iraola G, Betancor L, Calleros L, Gadea P, Algorta G, Galeano S, Muxi P, Greif G. Pérez 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. https://doi.org/10.1007/s10096 -015-2393-y.
- 63. van der Graaf-van Bloois L, Duim B, Miller WG, Forbes KJ, Wagenaar JA, Zomer A. 2016. Whole genome sequence analysis indicates recent diversification of mammal-associated Campylobacter fetus and implicates a genetic factor associated with H₂S production. BMC Genomics 17:713. https://doi.org/10.1186/s12864-016-3058-7.. 64. Iraola G, Forster SC, Kumar N, Lehours P, Bekal S, García-Peña FJ,
- Paolicchi F, Morsella C, Hotzel H, Hsueh P-R, Vidal A, Lévesque S, Yamazaki W, Balzan C, Vargas A, Piccirillo A, Chaban B, Hill JE, Betancor L, Collado L, Truyers I, Midwinter AC, Dagi HT, Mégraud F, Calleros L, Pérez R, Naya H, Lawley TD. 2017. Distinct Campylobacter fetus lineages adapted as livestock pathogens and human pathobionts in the intestinal microbiota. Nat Commun 8:1367. https://doi.org/10.1038/s41467 -017-01449-9
- 65. 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. https://doi.org/10.1093/cid/
- 66. Marchand-Senécal X, Bekal S, Pilon PA, Sylvestre J-L, Gaudreau C. 2017. Campylobacter fetus cluster among men who have sex with men, Montreal, Quebec, Canada, 2014-2016. Clin Infect Dis 65:1751-1753. https://doi.org/10.1093/cid/cix610.
- 67. Dingle KE, Blaser MJ, Tu Z-C, Pruckler J, Fitzgerald C, van Bergen MAP, Lawson AJ, Owen RJ, Wagenaar JA. 2010. Genetic relationships among reptilian and mammalian Campylobacter fetus strains determined by multilocus sequence typing. J Clin Microbiol 48:977-980. https://doi .org/10.1128/JCM.01439-09.
- 68. Patrick ME, Gilbert MJ, Blaser MJ, Tauxe RV, Wagenaar JA, Fitzgerald C. 2013. Human infections with new subspecies of Campylobacter fetus. Emerg Infect Dis 19:1678-1680. https://doi.org/10.3201/eid1910.130883.
- 69. Gilbert MJ, Miller WG, Yee E, Blaser MJ, Wagenaar JA, Duim B, 2013. Complete genome sequence of Campylobacter fetus subsp. testudinum strain 03-427T. Genome Announc 1:e01002-13. https://doi.org/10 .1128/genomeA.01002-13. 70. Wang C-M, Wu Z-Y, Shia W-Y, Jhou Y-J, Tung K-C, Shyu C-L. 2015. Com-
- plete genome sequence of Campylobacter fetus subsp. testudinum strain

Pet-3, isolated from a lizard (Hydrosaurus pustulatus). Genome Announc 3:e01420-14. https://doi.org/10.1128/genomeA.01420-14.

- Gilbert MJ, Miller WG, Yee E, Zomer AL, van der Graaf-van Bloois L, Fitzgerald C, Forbes KJ, Méric 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 Biol Evol 8:2006–2019. https://doi.org/10.1093/gbe/evw146.
- Gilbert MJ, Duim B, van der Graaf-van Bloois L, Wagenaar JA, Zomer AL. 2018. Homologous recombination between genetically divergent *Campylobacter fetus* lineages supports host-associated speciation. Genome Biol Evol 10:716–722. https://doi.org/10.1093/gbe/evy048.
- Gregory M, Klein B, Sahin O, Girgis G. 2018. Isolation and characterization of *Campylobacter hepaticus* from layer chickens with spotty liver disease in the United States. Avian Dis 62:79–85. https://doi.org/10 .1637/11752-092017-Reg.1.
- 74. Petrovska L, Tang Y, Jansen van Rensburg MJ, Cawthraw S, Nunez J, Sheppard SK, Ellis RJ, Whatmore AM, Crawshaw TR, Irvine RM. 2017. Genome reduction for niche association in Campylobacter hepaticus, a cause of spotty liver disease in poultry. Front Cell Infect Microbiol 7:354. https://doi.org/10.3389/fcimb.2017.00354.
- Miller WG, Chapman MH, Yee E, On SLW, McNulty DK, Lastovica AJ, Carroll AM, McNamara EB, Duffy G, Mandrell RE. 2012. Multilocus sequence typing methods for the emerging Campylobacter species C. hyointestinalis, C. lanienae, C. sputorum, C. concisus, and C. curvus. Front Cell Infect Microbiol 2:45. https://doi.org/10.389/fcimb.2012.00045.
- Harrington CS, On S. 1999. Extensive 16S rRNA gene sequence diversity in *Campylobacter hyointestinalis* strains: taxonomic and applied implications. Int J Syst Evol Microbiol 49:1171–1175. https://doi.org/10.1099/ 00207713-49-3-1171.
- Miller WG, Yee E, Chapman MH. 2016. Complete genome sequences of *Campylobacter hyointestinalis* subsp. *hyointestinalis* strain LMG 9260 and *C. hyointestinalis* subsp. *lawsonii* strain LMG 15993. Genome Announc 4:e00665-16. https://doi.org/10.1128/genomeA.00665-16.
- Wilkinson DA, O'Donnell AJ, Akhter RN, Fayaz A, Mack HJ, Rogers LE, Biggs PJ, French NP, Midwinter AC. 2018. Updating the genomic taxonomy and epidemiology of Campylobacter hyointestinalis. Sci Rep 8:2393. https://doi.org/10.1038/s41598-018-20889-x.
- Kamei K, Hatanaka N, Asakura M, Somroop S, Samosornsuk W, Hinenoya A, Misawa N, Nakagawa S, Yamasaki S. 2015. Campylobacter hyointestinalis isolated from pigs produces multiple variants of biologically active cytolethal distending toxin. Infect Immun 83:4304–4313. https://doi.org/10.1128/IAI.00997-15.
- Samosornsuk W, Kamei K, Hatanaka N, Taguchi T, Asakura M, Somroop S, Sugimoto N, Chaicumpa W, Yamasaki S. 2015. A new variant of cytolethal distending toxin in a clinical isolate of *Campylobacter hyointestinalis*. J Med Microbiol 64:1124–1134. https://doi.org/10.1099/jmm .0.000145.
- Bian X, Huynh S, Chapman MH, Szymanski CM, Parker CT, Miller WG. 2018. Draft genome sequences of nine *Campylobacter hyointestinalis* subsp. *lawsonii* strains. Microbiol Resour Announc 7:e01016-18. https:// doi.org/10.1128/MRA.01016-18.
- Costa D, Lévesque S, Kumar N, Fresia P, Ferrés I, Lawley TD, Iraola G. 2019. Pangenome analysis reveals genetic isolation in Campylobacter hyointestinalis subspecies adapted to different mammalian hosts. bioRxiv https:// doi.org/10.1101/600403.
- Guévremont E, Normand V, Lamoureux L, Côté C. 2008. Genetic detection of *Campylobacter lanienae* in fecal matter and stored manure from swine and dairy cattle. Foodborne Pathog Dis 5:361–364. https://doi .org/10.1089/fpd.2007.0054.
- Turowski EE, Shen Z, Ducore RM, Parry NMA, Kirega A, Dewhirst FE, Fox JG. 2014. Isolation of a *Campylobacter lanienae*-like bacterium from laboratory chinchillas (*Chinchilla laniger*). Zoonoses Public Health 61: 571–580. https://doi.org/10.1111/zph.12107.
- Miller WG, Yee E, Lopes BS, Chapman MH, Huynh S, Bono JL, Parker CT, Strachan NJC, Forbes KJ. 2017. Comparative genomic analysis identifies a *Campylobacter* clade deficient in selenium metabolism. Genome Biol Evol 9:1843–1858. https://doi.org/10.1093/gbe/evx093.
- Chua K, Gürtler V, Montgomery J, Fraenkel M, Mayall BC, Grayson ML. 2007. Campylobacter insulaenigrae causing septicaemia and enteritis. J Med Microbiol 56:1565–1567. https://doi.org/10.1099/jmm.0.47366-0.
- Kweon OJ, Lim YK, Yoo B, Kim HR, Kim T-H, Lee M-K. 2015. First case report of *Campylobacter volucris* bacteremia in an immunocompromised patient. J Clin Microbiol 53:1976–1978. https://doi.org/10.1128/ JCM.00442-15.

Y. 2001. Campylobacter lari bacteremia. Clin Microbiol Infect 7:96–97. https://doi.org/10.1046/j.1469-0691.2001.00212.x.

 Werno AM, Klena JD, Shaw GM, Murdoch DR. 2002. Fatal case of Campylobacter lari prosthetic joint infection and bacteremia in an immunocompetent patient. J Clin Microbiol 40:1053–1055. https://doi .org/10.1128/jcm.40.3.1053-1055.2002.

88. Martinot M, Jaulhac B, Moog R, Martino SD, Kehrli P, Monteil H, Piemont

- Miller WG, Wang G, Binnewies TT, Parker CT. 2008. The complete genome sequence and analysis of the human pathogen *Campylobacter lari*. Foodborne Pathog Dis 5:371–386. https://doi.org/10.1089/fpd.2008.0101.
- Miller WG, Yee E, Chapman MH, Smith TPL, Bono JL, Huynh S, Parker CT, Vandamme P, Luong K, Korlach J. 2014. Comparative genomics of the *Campylobacter lari* group. Genome Biol Evol 6:3252–3266. https://doi .org/10.1093/gbe/evu249.
- Bullman S, Corcoran D, O'Leary J, Lucey B, Byrne D, Sleator RD. 2011. Campylobacter ureolyticus: an emerging gastrointestinal pathogen? FEMS Immunol Med Microbiol 61:228–230. https://doi.org/10.1111/j .1574-695X.2010.00760.x.
- Bullman S, Lucid A, Corcoran D, Sleator RD, Lucey B. 2013. Genomic investigation into strain heterogeneity and pathogenic potential of the emerging gastrointestinal pathogen *Campylobacter ureolyticus*. PLoS One 8:e71515. https://doi.org/10.1371/journal.pone.0071515.
- Lucid A, Bullman S, Koziel M, Corcoran GD, Cotter PD, Sleator RD, Lucey B. 2014. Draft genome sequence of Campylobacter ureolyticus strain CIT007, the first whole-genome sequence of a clinical isolate. Genome Announc 2:e00262-14. https://doi.org/10.1128/genomeA.00262-14.
- Miller WG, Yee E, On SLW, Andersen LP, Bono JL. 2015. Complete genome sequence of the *Campylobacter ureolyticus* clinical isolate RIGS 9880. Genome Announc 3:e01291-15. https://doi.org/10.1128/genomeA.01291-15.
- Bourke B, Chan VL, Sherman P. 1998. Campylobacter upsaliensis: waiting in the wings. Clin Microbiol Rev 11:440–449. https://doi.org/10.1128/ CMR.11.3.440.
- Hald B, Madsen M. 1997. Healthy puppies and kittens as carriers of Campylobacter spp., with special reference to Campylobacter upsaliensis. J Clin Microbiol 35:3351–3352.
- Damborg P, Guardabassi L, Pedersen K, Kokotovic B. 2008. Comparative analysis of human and canine *Campylobacter upsaliensis* isolates by amplified fragment length polymorphism. J Clin Microbiol 46:1504–1506. https://doi.org/10.1128/JCM.00079-08.
- Baig A, McNally A, Dunn S, Paszkiewicz KH, Corander J, Manning G. 2015. Genetic import and phenotype specific alleles associated with hyper-invasion in *Campylobacter jejuni*. BMC Genomics 16:852. https:// doi.org/10.1186/s12864-015-2087-y.
- Palyada K, Threadgill D, Stintzi A. 2004. Iron acquisition and regulation in *Campylobacter jejuni*. J Bacteriol 186:4714–4729. https://doi.org/10 .1128/JB.186.14.4714-4729.2004.
- 101. Sheppard SK, Didelot X, Meric G, Torralbo A, Jolley KA, Kelly DJ, Bentley SD, Maiden MCJ, Parkhill J, Falush D. 2013. Genome-wide association study identifies vitamin B₅ biosynthesis as a host specificity factor in *Campylobacter*. Proc Natl Acad Sci U S A 110:11923–11927. https://doi.org/10.1073/pnas.1305559110.
- 102. Colles FM, Dingle KE, Cody AJ, Maiden M. 2008. Comparison of *Campylobacter* populations in wild geese with those in starlings and freerange poultry on the same farm. Appl Environ Microbiol 74:3583–3590. https://doi.org/10.1128/AEM.02491-07.
- 103. Carvalho ACT, Ruiz-Palacios GM, Ramos-Cervantes P, Cervantes L-E, Jiang X, Pickering LK. 2001. Molecular characterization of invasive and noninvasive Campylobacter jejuni and Campylobacter coli isolates. J Clin Microbiol 39:1353–1359. https://doi.org/10.1128/JCM.39 .4.1353-1359.2001.
- Sheppard SK, McCarthy ND, Jolley KA, Maiden M. 2011. Introgression in the genus *Campylobacter*: generation and spread of mosaic alleles. Microbiology 157:1066–1074. https://doi.org/10.1099/mic.0.045153-0.
- Tenover FC, Williams S, Gordon KP, Nolan C, Plorde JJ. 1985. Survey of plasmids and resistance factors in *Campylobacter jejuni* and *Campylobacter coli*. Antimicrob Agents Chemother 27:37–41. https://doi.org/10 .1128/aac.27.1.37.
- Alfredson DA, Korolik V. 2007. Antibiotic resistance and resistance mechanisms in *Campylobacter jejuni* and *Campylobacter coli*. FEMS Microbiol Lett 277:123–132. https://doi.org/10.1111/j.1574-6968.2007 .00935.x.
- 107. Tang Y, Dai L, Sahin O, Wu Z, Liu M, Zhang Q. 2017. Emergence of a plasmid-borne multidrug resistance gene *cfr*(C) in foodborne pathogen

cmr.asm.org 21

October 2019 Volume 32 Issue 4 e00072-18

Campylobacter. J Antimicrob Chemother 72:1581–1588. https://doi.org/ 10.1093/jac/dkx023.

- Abril C, Brodard I, Perreten V. 2010. Two novel antibiotic resistance genes, tet(44) and ant(6)-lb, are located within a transferable pathogenicity island in Campylobacter fetus subsp. fetus. Antimicrob Agents Chemother 54:3052–3055. https://doi.org/10.1128/AAC.00304-10.
- Backert S, Meyer TF. 2006. Type IV secretion systems and their effectors in bacterial pathogenesis. Curr Opin Microbiol 9:207–217. https://doi .org/10.1016/j.mib.2006.02.008.
- Bacon DJ, Alm RA, Burr DH, Hu L, Kopecko DJ, Ewing CP, Trust TJ, Guerry P. 2000. Involvement of a plasmid in virulence of *Campylobacter jejuni* 81-176. Infect Immun 68:4384–4390. https://doi.org/10.1128/IAI .68.8.4384-4390.2000.
- Worby CA, Mattoo S, Kruger RP, Corbeil LB, Koller A, Mendez JC, Zekarias B, Lazar C, Dixon JE. 2009. The Fic domain: regulation of cell signaling by adenylylation. Mol Cell 34:93–103. https://doi.org/10.1016/ j.molcel.2009.03.008.
- 112. Blaser MJ, Perez-Perez GI, Kleanthous H, Cover TL, Peek RM, Chyou PH, Stemmermann GN, Nomura A. 1995. Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res 55:2111–2115.
- 113. van Doorn LJ, Figueiredo C, Rossau R, Jannes G, van Asbroek M, Sousa JC, Carneiro F, Quint WG. 1998. Typing of *Helicobacter pylori vacA* gene and detection of *cagA* gene by PCR and reverse hybridization. J Clin Microbiol 36:1271–1276.
- Toft C, Andersson SG. 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. Nat Rev Genet 11:465–475. https://doi .org/10.1038/nrq2798.
- McCutcheon JP, Moran NA. 2011. Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol 10:13–26. https://doi.org/10.1038/ nrmicro2670.
- 116. Iraola G, Pérez R, Naya H, Paolicchi F, Pastor E, Valenzuela S, Calleros L, Velilla A, Hernández M, Morsella C. 2014. Genomic evidence for the emergence and evolution of pathogenicity and niche preferences in the genus *Campylobacter*. Genome Biol Evol 6:2392–2405. https://doi .org/10.1093/gbe/evu195.
- Mikonranta L, Mappes J, Laakso J, Ketola T. 2015. Within-host evolution decreases virulence in an opportunistic bacterial pathogen. BMC Evol Biol 15:165. https://doi.org/10.1186/s12862-015-0447-5.
- Iraola G, Kumar N. 2018. Surveying what's flushed away. Nat Rev Microbiol 16:456. https://doi.org/10.1038/s41579-018-0047-7.
- 119. Fresia P, Antelo V, Salazar C, Giménez M, D'Alessandro B, Afshinnekoo E, Mason C, Gonnet GH, Iraola G. 2019. Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. Microbiome 7:35. https://doi.org/10.1186/s40168-019-0648-z.
- Garcia MM, Lior H, Stewart RB, Ruckerbauer GM, Trudel JR, Skljarevski A. 1985. Isolation, characterization, and serotyping of *Campylobacter jejuni* and *Campylobacter coli* from slaughter cattle. Appl Environ Microbiol 49:667–672.
- Müller W, Böhland C, Methner U. 2011. Detection and genotypic differentiation of *Campylobacter jejuni* and *Campylobacter coli* strains from laying hens by multiplex PCR and fla-typing. Res Vet Sci 91:e48–e52. https://doi.org/10.1016/j.rvsc.2011.01.028.
- 122. Chaban B, Ngeleka M, Hill JE. 2010. Detection and quantification of 14 *Campylobacter* species in pet dogs reveals an increase in species richness in feces of diarrheic animals. BMC Microbiol 10:73. https://doi.org/10.1186/ 1471-2180-10-73.
- 123. Nonga HE, Muhairwa AP. 2010. Prevalence and antibiotic susceptibility of thermophilic *Campylobacter* isolates from free range domestic duck (*Cairina moschata*) in Morogoro municipality, Tanzania. Trop Anim Health Prod 42:165–172. https://doi.org/10.1007/s11250-009-9401-0.
- Hutchinson DN, Bolton FJ, Jelley WCN, Mathews WG, Telford DR, Counter DE, Jessop EG, Horsley SD. 1985. *Campylobacter* enteritis associated with consumption of raw goat's milk. Lancet i:1037–1038.
 Sestak K, Merritt CK, Borda J, Saylor E, Schwamberger SR, Cogswell F,
- 125. Sestak K, Merritt CK, Borda J, Saylor E, Schwamberger SR, Cogswell F, Didier ES, Didier PJ, Plauche G, Bohm RP, Aye PP, Alexa P, Ward RL, Lackner AA. 2003. Infectious agent and immune response characteristics of chronic enterocolitis in captive rhesus macaques. Infect Immun 71:4079–4086. https://doi.org/10.1128/iai.71.7.4079-4086.2003.
- 126. Petersen RF, Harrington CS, Kortegaard HE, On SL. 2007. A PCR-DGGE method for detection and identification of *Campylobacter*, *Helicobacter*, *Arcobacter* and related Epsilobacteria and its application to saliva samples from humans and domestic pets. J Appl Microbiol 103:2601–2615. https://doi.org/10.1111/j.1365-2672.2007.03515.x.

October 2019 Volume 32 Issue 4 e00072-18

- tion of the flagellin gene from isolates of urease-positive thermophilic *Campylobacter*. Res Microbiol 155:185–191. https://doi.org/10.1016/j .resmic.2003.12.003.
 128. Açik MN, Çetinkaya B. 2006. Heterogeneity of *Campylobacter jejuni* and
- Campylobacter coll strains from healthy sheep. Vet Microbiol 115:370–375. https://doi.org/10.1016/j.vetmic.2006.02.014.
- Kaakoush NO, Castaño-Rodríguez N, Mitchell HM, Man SM. 2015. Global epidemiology of *Campylobacter* infection. Clin Microbiol Rev 28: 687–720. https://doi.org/10.1128/CMR.00006-15.
- Tanner ACR, Dzink JL, Ebersole JL, Socransky SS. 1987. Wolinella recta, Campylobacter concisus, Bacteroides gracilis, and Eikeneila corrodens from periodontal lesions. J Periodontal Res 22:327–330. https://doi.org/ 10.1111/j.1600-0765.1987.tb01593.x.
- Vandamme P, Falsen E, Pot B, Hoste B, Kersters K, Ley JD. 1989. Identification of EF group 22 campylobacters from gastroenteritis cases as *Campylobacter concisus*. J Clin Microbiol 27:1775–1781.
- Wetsch NM, Somani K, Tyrrell GJ, Gebhart C, Bailey RJ, Taylor DE. 2006. Campylobacter curvus-associated hepatic abscesses: a case report. J Clin Microbiol 44:1909–1911. https://doi.org/10.1128/JCM .44.5.1909-1911.2006.
- 133. Mannering SA, West DM, Fenwick SG, Marchant RM, Perkins NR, O'Connell K. 2004. Pulsed-field gel electrophoresis typing of *Campylo-bacter fetus* subsp. *fetus* isolated from sheep abortions in New Zealand. N Z Vet J 52:358–363. https://doi.org/10.1080/00480169.2004.36452.
- Holst E, Wathne B, Hovelius B, Mårdh P-A. 1987. Bacterial vaginosis: microbiological and clinical findings. Eur J Clin Microbiol 6:536–541. https://doi.org/10.1007/BF02014242.
 Inglis GD, Boras VF, Houde A. 2011. Enteric campylobacteria and RNA
- Inglis GD, Boras VF, Houde A. 2011. Enteric campylobacteria and RNA viruses associated with healthy and diarrheic humans in the Chinook Health Region of southwestern Alberta, Canada. J Clin Microbiol 49: 209–219. https://doi.org/10.1128/JCM.01220-10.
- 136. Gebhart CJ, Edmonds P, Ward GE, Kurtz HJ, Brenner DJ. 1985. "Campylobacter hyointestinalis" sp. nov.: a new species of Campylobacter found in the intestines of pigs and other animals. J Clin Microbiol 21:715–720.
- Oporto B, Hurtado A. 2011. Emerging thermotolerant *Campylobacter* species in healthy ruminants and swine. Foodborne Pathog Dis 8:807–813. https://doi.org/10.1089/fpd.2010.0803.
- Edmonds P, Patton CM, Griffin PM, Barrett TJ, Schmid GP, Baker CN, Lambert MA, Brenner DJ. 1987. *Campylobacter hyointestinalis* associated with human gastrointestinal disease in the United States. J Clin Microbiol 25:685–691.
- Gorkiewicz G, Feierl G, Zechner R, Zechner EL. 2002. Transmission of Campylobacter hyointestinalis from a pig to a human. J Clin Microbiol 40:2601–2605. https://doi.org/10.1128/jcm.40.7.2601-2605.2002.
- 140. Hänninen M-L, Sarelli L, Sukura A, On SLW, Harrington CS, Matero P, Hirvelä-Koski V. 2002. Campylobacter hyointestinalis subsp. hyointestinalis, a common Campylobacter species in reindeer. J Appl Microbiol 92:717–723. https://doi.org/10.1046/j.1365-2672.2002.01574.x.
- Miller WG, Yee E, Huynh Š, Chapman MH, Parker CT. 2016. Complete genome sequence of *Campylobacter iguaniorum* strain RM11343, isolated from an alpaca. Genome Announc 4:e00646-16. https://doi.org/ 10.1128/genomeA.00646-16.
- 142. González M, Villanueva MP, Debruyne L, Vandamme P, Fernández H. 2011. Campylobacter insulaenigrae: first isolation report from South American sea lion (Otaria flavescens, (Shaw, 1800). Braz J Microbiol 42:261–265. https://doi.org/10.1590/S1517-83822011000100033.
- 143. Stoddard RA, Miller WG, Foley JE, Lawrence J, Gulland FMD, Conrad PA, Byrne BA. 2007. *Campylobacter insulaenigrae* isolates from northern elephant seals (*Mirounga angustirostris*) in California. Appl Environ Microbiol 73:1729–1735. https://doi.org/10.1128/AEM.01816-06.
- 144. Steele TW, Lanser JA, Sangster N. 1985. Nitrate-negative *Campylobacter*like organisms. Lancet i:394.
- 145. Kasper G, Dickgiesser N. 1985. Isolation from gastric epithelium of Campylobacter-like bacteria that are distinct from "Campylobacter pyloridis." Lancet i:111–112.
- 146. Figura N, Guglielmetti P, Zanchi A, Partini N, Armellini D, Bayeli PF, Bugnoli M, Verdiani S. 1993. Two cases of *Campylobacter mucosalis* enteritis in children. J Clin Microbiol 31:727–728.
- 147. Rams TE, Feik D, Slots J. 1993. Campylobacter rectus in human periodontitis. Oral Microbiol Immunol 8:230–235. https://doi.org/10.1111/j .1399-302X.1993.tb00565.x.
- 148. Atabay HI, Corry J. 1999. Clonality of Campylobacter sputorum bv.

paraureolyticus determined by macrorestriction profiling and biotyping, and evidence for long-term persistent infection in cattle. Epidemiol Infect 122:175–182. https://doi.org/10.1017/S0950268898001824.

- 149. Jay-Russell MT, Bates A, Harden L, Miller WG, Mandrell RE. 2012. Isolation of *Campylobacter* from feral swine (*Sus scrofa*) on the ranch associated with the 2006 *Escherichia coli* 0157:H7 spinach outbreak investigation in California. Zoonoses Public Health 59:314–319. https:// doi.org/10.1111/j.1863-2378.2012.01465.x.
- 150. Platts-Mills JA, Liu J, Gratz J, Mduma E, Amour C, Swai N, Taniuchi M, Begum S, Yori PP, Tilley DH, Lee G, Shen Z, Whary MT, Fox JG, McGrath M, Kosek M, Haque R, Houpt ER. 2014. Detection of *Campylobacter* in stool and determination of significance by culture, enzyme immunoassay, and PCR in developing countries. J Clin Microbiol 52:1074–1080. https://doi.org/10.1128/JCM.02935-13.
- Burnens AP, Nicolet J. 1992. Detection of *Campylobacter upsaliensis* in diarrheic dogs and cats, using a selective medium with cefoperazone. Am J Vet Res 53:48–51.
- Fox JG, Maxwell KO, Taylor NS, Runsick CD, Edmonds P, Brenner DJ. 1989. "Campylobacter upsaliensis" isolated from cats as identified by DNA relatedness and biochemical features. J Clin Microbiol 27:2376–2378.
- 153. Koziel M, Lucey B, Bullman S, Corcoran GD, Sleator RD. 2012. Molecularbased detection of the gastrointestinal pathogen *Campylobacter ureolyticus* in unpasteurized milk samples from two cattle farms in Ireland. Gut Pathog 4:14. https://doi.org/10.1186/1757-4749-4-14.
- Hariharan H, Richardson G, Homey B, Heaney S, Bryenton J, Moore I. 1994. Isolation of *Bacteroides ureolyticus* from the equine endometrium. J Vet Diagn Invest 6:127–130. https://doi.org/10.1177/104063879400600130
- Diagn Invest 6:127–130. https://doi.org/10.1177/104063879400600130.
 155. Nielsen HL, Ejlertsen T, Engberg J, Nielsen H. 2013. High incidence of *Campylobacter concisus* in gastroenteritis in North Jutland, Denmark: a population-based study. Clin Microbiol Infect 19:445–450. https://doi.org/10.1111/j.1469-0691.2012.03852.x.
- 156. Vries JJC, Arents NLA, Manson WL. 2008. Campylobacter species isolated from extra-oro-intestinal abscesses: a report of four cases and literature review. Eur J Clin Microbiol Infect Dis 27:1119–1123. https:// doi.org/10.1007/s10096-008-0550-2.
- Kaakoush NO, Mitchell HM. 2012. Campylobacter concisus—a new player in intestinal disease. Front Cell Infect Microbiol 2:4. https://doi .org/10.3389/fcimb.2012.00004.
- Macfarlane S, Furrie E, Macfarlane GT, Dillon JF. 2007. Microbial colonization of the upper gastrointestinal tract in patients with Barrett's esophagus. Clin Infect Dis 45:29–38. https://doi.org/10.1086/518578.
- Deshpande NP, Kaakoush NO, Wilkins MR, Mitchell HM. 2013. Comparative genomics of *Campylobacter concisus* isolates reveals genetic diversity and provides insights into disease association. BMC Genomics 14:585. https://doi.org/10.1186/1471-2164-14-585.
 Mendz GL, Petersen R, Quinlivan JA, Kaakoush NO. 2014. Potential
- Mendz GL, Petersen R, Quinlivan JA, Kaakoush NO. 2014. Potential involvement of *Campylobacter curvus* and *Haemophilus parainfluenzae* in preterm birth. Case Rep 2014:bcr2014205282. https://doi.org/10 .1136/bcr-2014-205282.
- Horio Y, Shiraishi Y, Watanabe N, Inoue S, Imanishi T, Asano K. 2017. Empyema associated with Campylobacter curvus infection. Respirol Case Rep 5:e00234. https://doi.org/10.1002/rcr2.234.
- 162. Han XY, Tarrand JJ, Rice DC. 2005. Oral Campylobacter species involved in extraoral abscess: a report of three cases. J Clin Microbiol 43: 2513–2515. https://doi.org/10.1128/JCM.43.5.2513-2515.2005.
- Blaser MJ, Newell DG, Thompson SA, Zechner EL. 2008. Pathogenesis of Campylobacter fetus, p 401–428. *In* Nachamkin I, Szymanski CM, Blaser MJ (ed), Campylobacter, 3rd ed. ASM Press, Washington, DC.
- 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–e230. https://doi.org/10.1111/j.1439-0531.2009.01546.x.
- 165. Shinha T. 2015. Fatal bacteremia caused by Campylobacter gracilis, United States. Emerg Infect Dis 21:1084–1085. https://doi.org/10.3201/ eid2106.142043.
- 166. Lee D, Goldstein EJ, Citron DM, Ross S. 1993. Empyema due to Bacteroides gracilis: case report and in vitro susceptibilities to eight antimicrobial agents. Clin Infect Dis off Publ Infect Dis Soc Am 16(Suppl 4):S263–S265. https://doi.org/10.1093/clinids/16.Supplement_4.S263.

- Johnson CC, Reinhardt JF, Edelstein MA, Mulligan ME, George WL, Finegold SM. 1985. *Bacteroides gracilis*, an important anaerobic bacterial pathogen. J Clin Microbiol 22:799–802.
- 168. Mukhopadhya I, Thomson JM, Hansen R, Berry SH, El-Omar EM, Hold GL. 2011. Detection of *Campylobacter concisus* and other *Campylobacter* species in colonic biopsies from adults with ulcerative colitis. PLoS One 6:e21490. https://doi.org/10.1371/journal.pone.0021490.
- Miller WG, Yee E. 2015. Complete genome sequence of Campylobacter gracilis ATCC 33236T. Genome Announc 3:e01087-15. https://doi.org/ 10.1128/genomeA.01087-15.
- Broczyk A, Thompson S, Smith D, Lior H. 1987. Water-borne outbreak of Campylobacter laridis-associated gastroenteritis. Lancet i:164–165. https://doi.org/10.1016/S0140-6736(87)92003-4.
- Morris CN, Scully B, Garvey GJ. 1998. Campylobacter lari associated with permanent pacemaker infection and bacteremia. Clin Infect Dis 27: 220–221. https://doi.org/10.1086/517683.
- 172. Lastovica AJ, Le Roux E. 2000. Efficient isolation of campylobacteria from stools. J Clin Microbiol 38:2798–2799.
- 173. Siqueira JJ, Rôças IN. 2003. Campylobacter gracilis and Campylobacter rectus in primary endodontic infections. Int Endod J 36:174–180. https:// doi.org/10.1046/j.1365-2591.2003.00636.x.
- 174. Castillo DM, Sánchez-Beltrán MC, Castellanos JE, Sanz I, Mayorga-Fayad I, Sanz M, Lafaurie GI. 2011. Detection of specific periodontal microorganisms from bacteraemia samples after periodontal therapy using molecular-based diagnostics. J Clin Periodontol 38:418–427. https:// doi.org/10.1111/j.1600-051X.2011.01717.x.
- Mahlen SD, Clarridge JE. 2009. Oral abscess caused by *Campylobacter* rectus: case report and literature review. J Clin Microbiol 47:848–851. https://doi.org/10.1128/JCM.01590-08.
- 176. Lam JYW, Wu AKL, Ngai DC, Teng JLL, Wong ESY, Lau SKP, Lee RA, Woo P. 2011. Three cases of severe invasive infections caused by *Campylobacter rectus* and first report of fatal *C. rectus* infection. J Clin Microbiol 49:1687–1691. https://doi.org/10.1128/JCM.02487-10.
- Vandenberg O, Houf K, Douat N, Vlaes L, Retore P, Butzler JP, Dediste A. 2006. Antimicrobial susceptibility of clinical isolates of non-jejuni/coli campylobacters and arcobacters from Belgium. J Antimicrob Chemother 57:908–913. https://doi.org/10.1093/jac/dkl080.
- Lindblom G-B, Sjögren E, Hansson-Westerberg J, Kaijser B. 1995. Campylobacter upsaliensis, C. sputorum sputorum and C. concisus as common causes of diarrhoea in Swedish children. Scand J Infect Dis 27: 187–188. https://doi.org/10.3109/00365549509019006.
- On SLW, Ridgwell F, Cryan B, Azadian BS. 1992. Isolation of Campylobacter sputorum biovar sputorum from an axillary abscess. J Infect 24:175–178. https://doi.org/10.1016/0163-4453(92)92902-U.
- Tee W, Luppino M, Rambaldo S. 1998. Bacteremia due to Campylobacter sputorum biovar sputorum. Clin Infect Dis 27:1544–1545. https:// doi.org/10.1086/517748.
- Miller WG, Yee E, Chapman MH, Bono JL. 2017. Comparative genomics of all three *Campylobacter sputorum* biovars and a novel cattle-associated C. *sputorum* clade. Genom Biol Evol 9:1513–1518. https://doi.org/10.1093/ gbe/evx112.
- Jenkin GA, Tee W. 1998. Campylobacter upsaliensis-associated diarrhea in human immunodeficiency virus-infected patients. Clin Infect Dis 27:816–821. https://doi.org/10.1086/514957.
- Bullman S, O'Leary J, Corcoran D, Sleator RD, Lucey B. 2012. Molecularbased detection of non-culturable and emerging campylobacteria in patients presenting with gastroenteritis. Epidemiol Infect 140:684–688. https://doi.org/10.1017/S0950268811000859.
- 184. Goossens H, Vlaes L, De Boeck M, Levy J, De Mol P, Butzler J-P, Kersters K, Pot B, Vandamme P. 1990. Is "Campylobacter upsaliensis" an unrecognised cause of human diarrhoea? Lancet 335:584–586. https://doi.org/10.1016/0140-6736(90)90359-D.
- Gurgan T, Diker KS. 1994. Abortion associated with Campylobacter upsaliensis. J Clin Microbiol 32:3093–3094.
- Lastovica AJ, Roux EL, Penner JL. 1989. "Campylobacter upsaliensis" isolated from blood cultures of pediatric patients. J Clin Microbiol 27:657–659.
- Gaudreau C, Lamothe F. 1992. Campylobacter upsaliensis isolated from a breast abscess. J Clin Microbiol 30:1354–1356.

Continued next page

October 2019 Volume 32 Issue 4 e00072-18

3

Obtención de genomas de *Campylobacter fetus* aisladas en Uruguay

3.1. Introducción

... (A) knowledge of sequences could contribute much to our understanding of living matter.

Frederick Sanger, 1980

La palabra *genoma* se refiere a el conjunto completo de ADN de un organismo. En sus inicios, la secuenciación del ADN era aplicable sólo para genes o fragmentos de estos, debido a su altísimo costo y complejidad, por lo que la genómica tal como la conocemos hoy en día era inimaginable.

Al acceder a la información almacenada en el material genético de un organismo, abrimos la posibilidad de incontables aplicaciones. En el caso de los genomas bacterianos, la secuenciación ha permitido avances inconmensurables con respecto a la era pre-genómica, lo que ha significado un cambio de paradigma en la forma de investigar en microbiología. Tan solo 28 años atrás, en el año 1995, comenzaba la era genómica con la publicación del genoma de 1,8 millones de bases (Mb) de *Haemophilus influenzae* [107], hito al que le sucedieron las publicaciones de los genomas de las principales especies modelo [108].

3.1.1. De reads a genomas

La información almacenada en un genoma pasa por múltiples instancias de procesamiento y evaluación antes de plasmarse en una secuencia representativa. La primera etapa consiste en la lectura de las bases en una molécula de ADN, para generar un producto intangible como lo son las lecturas o *reads*. Estos serán las piezas de un rompecabezas armado a ciegas, sin más información de la contenida en las piezas y sus bordes, en un proceso llamado ensamblado *de novo*, que se describe más adelante (Ver Subsubsección 3.1.1).

En los últimos 40 años se han diseñado varias estrategias de secuenciación, transformándola en un recurso cada vez más accesible, gracias a los avances que permitieron la generación de metodologías cada vez más baratas y eficientes. La historia de estas tecnologías puede contarse en tres capítulos correspondientes a las tres generaciones de tecnologías de secuenciación, las cuales se describen a continuación. En la Tabla 3.1 se comparan algunas de las plataformas de secuenciación más utilizadas.

Primera generación: Podemos sentar las bases de la genómica en la década de 1970, con el desarrollo de una técnica para secuenciar el ADN. Esta técnica, que conocemos hasta ahora como el método de Sanger en honor a su creador, Frederick Sanger, fue la tecnología dominante en la secuenciación de genomas hasta entrada la década de los 2000, constituyendo lo que conocemos como la primera generación de tecnologías de secuenciación.

Método de Sanger: Este método se basa en la síntesis de fragmentos de ADN de diferentes longitudes, generados *in vitro* a partir de una cadena

molde, mediante la adición de nucleótidos (dNTPs) convencionales y didesoxinucleótidos (ddNTPs) marcados. Este método también es conocido como de "terminación de cadena", ya que la síntesis prosigue hasta la adición de un ddNTP, a partir del cual no puede continuar la polimerización. Los fragmentos de diferentes longitudes generados en la fase de síntesis son luego separados mediante electroforesis, y la secuencia del ADN molde se infiere a partir del patrón de bandas resultante.

En la década de 1980 se generó un plataforma automática de secuenciación basada en este método, marcando el inicio de la primera generación de plataformas de secuenciación que fueron lanzadas al mercado.

Esta tecnología sigue siendo relevante para algunas aplicaciones, ya que genera secuencias de aproximadamente 1000 bases con una calidad muy alta, requiriendo un mínimo procesamiento computacional, características que la convierten en una opción atractiva en el ámbito del diagnóstico clínico [109].

Segunda generación: La secuenciación comenzó a popularizarse, pero seguía siendo muy costosa y laboriosa. Es entonces que comienza una carrera impulsada por la necesidad de disminuir los costos y aumentar la eficiencia de la secuenciación. El lanzamiento al mercado de la plataforma desarrollada por 454 Life Technologies, basada en la pirosececuenciación, dio inicio a la segunda generación de tecnologías de secuenciación, conocida también como tecnologías de secuenciación de próxima generación, o NGS por sus siglas en ingles. Las características de esta segunda generación de tecnologías de secuenciación de tecnologías de secuenci

 No se requiere un paso previo de clonación, eliminando los sesgos inherentes a este paso y aumentando su aplicabilidad.

- La secuenciación se produce en forma masiva y paralela, ya que se secuencian al mismo tiempo millones de fragmentos de ADN.
- No se requiere del procesamiento posterior para inferir las secuencias del ADN, ya que estas se infieren a partir de información recopilada durante la secuenciación.

Illumina La plataforma de segunda generación más utilizada y que predomina en el mercado es la desarrollada por Solexa, compañía que luego adquirió Illumina, nombre con el que se asocia esta tecnología. Pueden distinguirse cuatro etapas en el proceso de secuenciación con Illumina:

- La primera etapa consiste en generar las librerías de secuenciación, que consisten en fragmentos de ADN con longitudes que generalmente se encuentran en el rango de entre 300 y 600 bases, y secuencias adaptadoras unidas a ambos extremos. Luego, se transfiere la biblioteca a una celda de flujo o *flowcell*, que consiste en una superficie que contiene oligonucleótidos anclados a su superficie, a los que se unirán los fragmentos mediante complementariedad de bases con una región del adaptador.
- Se realiza la amplificación de cada fragmento sobre la superficie de la celda, sobre la cual también permanecen unidos los productos de amplificación, generando *clústers* compuestos por fragmentos de ADN idénticos con sus secuencias adaptadoras.
- Un *read* se genera a partir de cada *clúster* mediante secuenciación por síntesis, en la que cada base es incorporada a las cadenas en crecimiento en forma sincronizada. Ocurren típicamente 150 ciclos que consisten en:

- a) incorporación de un nucleótido conjugado a un terminador reversible y un fluoróforo
- b) lavado de la celda para eliminar el excedente de nucleótidos no unidos
- c) excitación del fluoróforo y captura de imagen
- *d*) clivado del terminador, para permitir la incorporación de la siguiente base y el inicio de un nuevo ciclo.
- 4. Al finalizar los ciclos de secuenciación, el equipo cuenta con las imágenes capturadas, y genera a partir de éstas las secuencias correspondientes a cada *clúster*, en un proceso denominado llamado de bases o *base calling*.

La secuenciación produce *reads* o lecturas cortas a partir de uno o ambos extremos de los fragmentos, según lo cual decimos que los *reads* son simples (SE o *single end*) o pareados (PE o *pair end*). Estos últimos son más convenientes ya que no solo brindan información acerca de la secuencia en los extremos de los fragmentos, sino que se conoce la distancia que separa al par de *reads* entre si, lo cual facilita el procesamiento *downstream*.

La aplicación de esta tecnología genera millones de *reads* cortos pero de muy buena calidad. Sin embargo no está exenta de errores o sesgos, incorporados en cada etapa del proceso. Una de las fuentes de error más frecuentes que se produce durante la secuenciación, es la generación de señales ambiguas debido al desfase de la síntesis en los *clústers*, que dificultan el *base calling* y se hace más notable a medida que avanza la secuenciación, razón por la cual es frecuente observar una disminución en la calidad de la secuencias hacia el final de los *reads*. **Tercera generación:** La tercera generación de tecnologías de secuenciación es dominada por dos plataformas: Nanopore y PacBio. Estas tecnologías abren una tercera generación de secuenciación debido a dos características esenciales: la secuencia se obtiene en tiempo real a partir de una única molécula. Esto implica que no es necesario un paso de amplificación previo a la secuenciación, ya que cada fragmento es secuenciado de forma individual, eliminando sesgos asociados a este procedimiento.

Nanopore: Esta tecnología, desarrollada por Oxford Nanopore Technologies se basa en la lectura de potenciales eléctricos generados durante el paso de una molécula por un poro que está embebido en una membrana. Puede utilizarse tanto para ADN como para ARN, y tiene la ventaja de que, al no ser necesaria la amplificación, las moléculas son "leídas" en su forma nativa. El *base calling* se realiza de forma simultánea a la lectura de las variaciones de potencial eléctrico a través de la membrana durante el paso de la molécula. Esta tecnología tiene la ventaja de que generar lecturas muy largas, de entre 5 y 30 kilobases (Kb), con la posibilidad de llegar hasta 4 Mb. Si bien la calidad de la secuenciación es limitada, la incorporación de mejoras en la detección de señales y en el *base calling* ha permitido mejorar significativamente la calidad de las lecturas[110]. Se ha reportado que la tecnología Nanopore es poco eficiente en la secuenciación de regiones de baja complejidad y en la secuenciación de genomas o regiones de alto contenido de GC [111].

PacBio: Esta tecnología utiliza la secuenciación por síntesis como estrategia, en la que se utilizan nucleótidos marcados con fluoróforos y se detecta la incorporación de cada nucleótido detectando la fluorescencia emitida. PacBio produce *reads* largos, y no requiere la amplificación de las señales de fluorescencia, debido a que cada fragmento de ADN es secuenciado individualmente. Esto es posible ya que la ADN polimerasa se encuentra unida al fondo de pocillos de escala nanométrica llamadas *zero-mode waveguides* o ZMW, en los que ocurre la detección de forma individual. Se obtiene un perfil temporal de fluorescencia en los cuatro canales, que corresponden a los cuatro bases del ADN. Si bien esta tecnología se asociaba con grandes tasas de error, no se han encontrado errores sistemáticos, con lo que al aumentar la profundidad de la secuenciación su efecto disminuye. Una mejora que aumentó considerablemente la calidad de la secuenciación es la circularización de los fragmentos a secuenciar, que permite una lectura continua con lo que cada base es leída muchas veces.

Cada lectura continua que se genera por el paso de la molécula por la ADN polimerasa se conoce como lecturas de la polimerasa, o *polymerase reads*. Estas lecturas son procesadas mediante un programa que elimina las secuencias adaptadoras y extrae cada secuencia individual, correspondiente al fragmento original. Estas lecturas individuales o *subreads* provenientes del mismo *polymerase read* son comparados y se obtiene una secuencia consenso de alta calidad [112].

Calidad: El proceso de secuenciación no es perfecto, ya que cada una de sus etapas es propensa a incorporar sesgos y errores que repercuten negativamente en el producto de secuenciación y en sus aplicaciones posteriores. Independientemente de la tecnología utilizada es necesario asegurar la calidad de los datos de secuenciación antes de continuar el procesamiento hacia la obtención de los genomas. Como producto de la secuenciación se obtiene no solo información sobre la secuencia de los *reads*, sino también la calidad asociada a cada base, que refleja la confianza de que la base llamada sea la correcta. La evaluación de la calidad de las lecturas nos permite asegurarnos de que se obtengan secuencias de calidad, que puedan ser utilizadas con confianza en procesos posteriores, como el mapeo o el ensamblado de los

Plataforma	Longitud máxima	Output (bases)	Output (reads)	Ventajas	Desventajas
Sanger ABi 3500 S	>800 bp	-	Hasta 384	- Alta calidad (Error: 0,01%)	- Alto costo
					- Sesgos en la clonación
					- Bajo rendimiento
					- Errores en homopolímeros
Illumina MiSeq	$2 \times 300 \text{ bp}$	0,3-15 Gb	$1-25 \times 10^{6}$	- Bajo costo	- Reads cortos
				- Reads de alta calidad (Error: 0,613%)	- Sesgos en la amplificación
					- Errores acumulados sobre el final de los reads
					- No permite resolver regiones repetidas
					- Errores en secuencias con alto GC
Illumina HiSeq	$2\times 150 \ bp$	650-750 Gb	2,5 (SE), 5 (PE) $\times 10^7$	- Reads de alta calidad (Error:0,112%)	
PacBio RSII	>1 Kb	1 Gb	1-10 ×10 ⁶	- Reads largos	- Alta tasa de error (10-15%)
				- No requiere amplificación	
				- Costo medio	
ONT MinION	900 kb	48 Gb	Hasta 1 $\times 10^{6}$	- Bajo costo	- Alta tasa de error (0,1-1,7%) [1]
				- Reads ultra largos	- Errores en secuencias de baja complejidad

Tabla 3.1: Comparación de las principales plataformas de secuenciación utilizadas

[1] Valor depende de la tecnología: 1,7% (R9.4.1), 0,1-1% (R10.4.1)

genomas. Herramientas informáticas como por ejemplo FastQC [113] nos permiten visualizar la calidad de un experimento de secuenciación masiva e identificar errores y/o sesgos. Por ejemplo, es frecuente encontrar que la calidad de las bases hacia el final de los *reads* producidos por Illumina disminuye de forma sostenida, producto del desfasaje de la secuenciación. Al visualizar esta información es posible diseñar estrategias para remediar este y otros problemas de forma consciente o automática, en un proceso conocido como *trimming*, teniendo en cuenta que la elección del método elegido no resulte en una pérdida significativa de información.

Ensamblado de novo

Los *reads* no suelen tener información sobre a qué región del genoma pertenecen, por lo que determinar su orden en el material genético original supone un gran problema bioinformático. Todos los abordajes de ensamblado de *reads* parten del mismo concepto: las secuencias similares probablemente provengan de la misma región del genoma [114]. El ensamblado *de novo* es el procesamiento mediante el cual se busca reproducir la secuencia genómica original a partir de los datos de secuenciación, sin utilizar un genoma de referencia. En este proceso los *reads* son ordenados en base a superposiciones en sus secuencias para dar cadenas ininterrumpidas más largas denominadas *contigs*, que a su vez pueden asociarse para generar *scaffolds*. Podemos clasificar los métodos de ensamblaje en tres categorías según el largo de los *reads* [115].

Sólo reads cortos: Esta estrategia es la más utilizada, ya que parte únicamente de *reads* cortos, por ejemplo, los generados con Illumina. Primero, los *reads* son divididos en subsecuencias de k bases o k-meros y se construyen grafos, generalmente basados en los grafos de de Bruijn [115], en los que se representan los k-meros y sus uniones si comparten secuencias solapantes. Para obtener el ensamblado se reconstruye el camino que atraviesa todos los nodos (o conectores) del grafo una sola vez. Uno de los programas más populares que utiliza esta estrategia es SPAdes [116].

Sólo reads largos: Al utilizarse una tecnología de tercera generación los problemas relacionados con la longitud de los *reads* y los sesgos de amplificación se resuelven, pero, debido a la alta tasa de error asociada a estas tecnologías, se realiza un paso de corrección de errores. Luego, el ensamblado puede realizarse mediante superposición directa de *reads*, con o sin la compilación de k-meros. Luego del ensamblado se realiza una etapa de "pulido" de los genomas o *polishing*, con el propósito de corregir los errores que pudieran quedado sin resolver en etapas anteriores [117].

Ensamblado híbrido: las estrategias antes mencionadas pueden ser insuficientes para generar un genoma de buena calidad, ya que los *reads* cortos son insuficientes para resolver regiones de baja complejidad y por la baja calidad de los *reads* largos. Existen dos estrategias para realizar el ensamblado híbrido, según el tipo de *reads* que se ensamble primero: largos + cortos: Primero se realiza el ensamblado utilizando exclusivamente los *reads* largos, para luego mejorar la calidad utilizando la información brindada por los *reads* cortos. Este caso es similar a la estrategia que sólo utiliza *reads* largos, ya que los *reads* cortos son incorporados a la ecuación durante la fase de *polishing*.

cortos + largos: En este caso se realiza el ensamblado con *reads* cortos, y el grafo resultante es mejorado con la información almacenada en los *reads* largos, lo que aumenta la resolución de las regiones de baja complejidad en el genoma.

Por ejemplo, Unicycler, un programa especializado en el ensamblado híbrido, toma los *contigs* producidos a partir del ensamblado de los *reads* cortos y los combina con los *reads* largos para procesarlos de una manera similar a la utilizada por la estrategia que sólo emplea *reads* largos. Las regiones del grafo correspondientes a secuencias presentes en más de una copia por ejemplo, se observan como nodos ramificadores, que vuelven a converger a un mismo nodo. Para resolver esta estructura, Unicycler utiliza una estrategia de construcción de *puentes* mediante alineamiento global de los *contigs* involucrados y los *reads* largos, simplificando así el grafo [118].

En busca del ensamblado perfecto El ensamblado, es un proceso complejo, computacionalmente costoso y sensible a los errores y sesgos en el proceso de adquisición de datos. Un error que incide en el ensamblado es el producido por una cobertura de *reads* no uniforme, producto de sesgos en la generación de la librería de secuenciación o durante la amplificación. Si la cobertura no es uniforme es probable una mayor inserción de *gaps* en el ensamblado. Quizás la limitación más relevante en los ensamblados proviene de la presencia de regiones repetidas o de baja complejidad, cuya longitud excede la longitud de los *reads*, por lo que no pueden ser resueltas y son responsables en gran medida de la alta fragmentación de los genomas. [115].

El ensamblado ideal de un genoma bacteriano es aquel que está completo, ya que contiene tantos *contigs* circularizados como replicones, y está libre de errores, *gaps* o bases indeterminadas [119]. Existen varias herramientas para evaluar la calidad de los genomas ensamblados. En general todas tienen en cuenta unos pocos parámetros, tales como la continuidad de la secuencia, el número de *contigs* y el N50, y otros que dependen del taxo, como los que reflejan la completitud y contaminación de un ensamblado:

- Número de *contigs*: El número de *contigs* en un genoma debe ser bajo y corresponderse con el número de replicones si el genoma obtenido se encuentra completo. Sin embargo, esto es muy difícil de lograr, sobre todo si sólo se parte de *reads* cortos. Esta dificultad se refleja en la bajísima proporción de genomas completos encontrados en la base de datos *Datasets* del NCBI, ya que sólo el 2,44 % de los genomas bacterianos depositados en esta base de datos se encuentran completos¹.
- N50: Esta métrica indica que el 50 % de un genoma dado está contenido en *contigs* que son iguales o mayores que este valor. Es deseable que este número sea igual o cercano a la longitud total del genoma, en los genomas compuestos por un único cromosoma.
- **Completitud y contaminación:** Estos parámetros se obtienen evaluando la presencia de un conjunto de genes marcadores ubicuos y de copia única para un determinado nivel taxonómico. En un genoma completo y sin contaminaciones deben encontrarse todos los genes marcadores en una única copia. Si, por ejemplo, todos los genes marcadores están dos veces en el genoma su nivel de contaminación será del 100 %.

¹Valor correspondiente al porcentaje de genomas que figuran como "completos" sobre el total de los genomas bacterianos depositados en esta base de datos.
Mejoramiento de genomas fragmentados Si bien es más frecuente obtener genomas fragmentados, en especial si se utilizan sólo *reads* cortos para el ensamblado, se pueden tomar algunas acciones para mejorar su calidad, en ausencia de datos de secuenciación adicionales. Existen para este fin muchas estrategias, generalmente como parte de flujos de trabajo o *pipelines in house*. Una estrategia eficiente y fácil de implementar, desarrollada por el Instituto Sanger llamada Improve Assembly [120] consiste en aplicar primero la herramienta SSPACE [121] de forma iterativa, que extiende los *contigs* utilizando los *reads* no pareados (*singletons*), para luego mapear los *reads* pareados a los extremos de los *contigs* y así generar uniones entre ellos insertando *gaps*. En una segunda etapa del *pipeline*, se ejecuta GapFiller [122] de forma iterativa para rellenar estos *gaps* utilizando información contenida en los *reads*.

Al implementar una estrategia de mejoramiento del ensamblado se pueden obtener *contigs* más extensos, e incluso disminuir el número de *contigs*, al formar uniones entre ellos. Esto puede marcar la diferencia, sobre todo en estudios basados en genómica comparativa y en análisis funcionales de los genomas, al mejorar considerablemente la calidad de sus anotaciones [123].

3.1.2. De genomas a genes

La anotación es el proceso en el que se localizan secuencias características o *features*, incluidos marcos abiertos de lectura (ORFs por sus siglas en inglés), ARNs ribosomales o de transferencia, entre otros, en un genoma seguido de la asignación funcional de los *features* encontrados. Un marco abierto de lectura, o simplemente ORF puede definirse como la secuencia de ADN que

se encuentra circunscrita entre un codón de inicio y uno de finalización o *stop*, cuya longitud es múltiple de tres [124].

Existen varias herramientas para realizar este proceso [125, 126, 127], que constituyen *pipelines* en los que se ejecutan de forma secuencial y automática varios programas. Por ejemplo, PROKKA [125], predice la posición de los ORFs utilizando el programa Prodigal [128], para luego identificar sus productos mediante homología de secuencias de forma jerárquica, utilizando primero bases de datos más pequeñas y específicas y por último las más grandes y generales. En el caso de no encontrar *hits* en ninguna de las bases de datos utilizadas, el producto se designa como "proteína hipotética". Otros tipos de *features* son predichos con otras herramientas: RNAmmer [129], Aragorn [130] e Infernal [131] para anotar ARNs ribosomales, de transferencia y no codificantes, respectivamente.

Cabe resaltar que la precisión de la anotación depende de la calidad de los genomas ensamblados que forman parte de las bases de datos de referencia y de los que se están anotando. La calidad de los genomas presentes en las bases de datos de referencia constituyen un gran problema, ya que sus errores se propagan rápidamente. En condiciones ideales, las anotaciones de los genomas deben ser curadas manualmente, pero el volumen de datos de secuenciación masiva que se generan y depositan en las bases de datos públicas diariamente impide este tipo de acciones [132, 133, 123].

3.2. Hipótesis y objetivos

3.2.1. Hipótesis

Se ha reportado la presencia de *C. fetus* en Uruguay desde hace al menos 5 décadas, generando grandes pérdidas económicas en el sector productivo, pero también en el ámbito clínico, donde se ha aislado principalmente en individuos inmunocomprometidos.

Para evaluar el impacto de este y otros patógenos resulta imperativo avanzar más allá de los métodos de diagnóstico convencionales. Un enfoque genómico es útil, no solo para identificar la presencia de patógenos en la población, sino también para comprender su virulencia, establecer relaciones entre casos y determinar fuentes comunes.

La secuenciación genómica emerge como una herramienta esencial en este contexto, especialmente con la accesibilidad actual a estas tecnologías. Sin embargo, para garantizar la calidad de los datos, es esencial contar con información genómica local y de alta calidad.

Cada tecnología de secuenciación tiene sus ventajas y limitaciones. La combinación estratégica de tecnologías de segunda y tercera generación superará las limitaciones inherentes a cada tecnología, y permitirá obtener genomas completos y de alta calidad.

3.2.2. Objetivos

Objetivo general

Ensamblar y analizar secuencias genómicas de *C. fetus* de alta calidad a partir de aislamientos uruguayos, procedentes de muestras de campo o biobancos.

Objetivos específicos

- 1. Ensamblar los genomas de *C. fetus* secuenciados mediante tecnología Illumina.
- Completar genomas, mediante la utilización de una estrategia de ensamblado híbrido, combinando datos de secuenciación obtenidos con las tecnologías Illumina y PacBio RSII.
- 3. Anotar los genomas obtenidos.

3.3. Metodología

3.3.1. Antecedentes

Las nuevas secuencias genómicas fueron obtenidas a partir de aislamientos cedidos por colaboradores en el INIA, la Facultad de Ciencias (Universidad de la República), quienes participaron del proyecto de la ANII FS-SA_X_2014_1_105252. También fueron secuenciados aislamientos de origen clínico provistos por colaboradores en distintos hospitales de Montevideo, y el Instituto de Higiene de la Facultad de Medicina (Universidad de la República) y otros. El listado de las cepas, su procedencia y otros datos se muestra en la Tabla 3.2.

Tabla 3.2: Información asociada a los genomas de C. fetus obtenidos en este trabajo

Genoma	Cepa	Año	País	Hospedero	Muestra	Subespecie	Procedencia
15031_\$95	15031		Uruguay				INIA
15671_S81	15671	2015	Uruguay	Bovino	Aborto	Cfv	INIA
17133_S82	17133	2017	Uruguay	Bovino	Aborto	Cfv	INIA
17144_S83	17144	2017	Uruguay	Ovino	Aborto	Cff	INIA
2740_S96	2740		Uruguay				DILAVE
95258_S94	95258	1995	Argentina				INTA
ATCC_S93	ATCC-19438	1962	Reino Unido	Bovino	Vagina	Cfv	ATCC [134]
Felv5	H1-EC	2016	Ecuador	Humano	Sangre	Cff	Chávez et al. [135]
FETOS	RN	2017	Argentina	Humano	Sangre	Cff	Novogen
H1_S84	H1	2016	Uruguay	Bovino	Prepucio	Cff	FCIEN/IH
HB74	H7-UY	2015	Uruguay	Humano	Sangre	Cff	Clínica
HC	H2-UY	2014	Uruguay	Humano	Sangre	Cff	Clínica
HC71	H5-UY	2015	Uruguay	Humano	Fluido de ascitos	Cff	Clínica
HC73	H6-UY	2015	Uruguay	Humano	Sangre	Cff	Clínica
KCH09208	H1-JP	2015	Japón	Humano	Sangre		Kitamura et al. [136]
KCH09951	H2-JP	2015	Japón	Humano	Sangre		Kitamura et al. [136]
M8_S86	M8	2017	Uruguay	Bovino	Vagina	Cfv	FCIEN
HM1	H8-UY	2018	Uruguay	Humano	Sangre	Cff	Clínica
R18_S87	R18	2017	Uruguay	Bovino	Prepucio	Cfv	FCIEN
U10_S88	U10	2017	Uruguay	Bovino	Prepucio	Cfv	FCIEN
V28_S89	V28	2017	Uruguay	Bovino	Prepucio	Cfv	FCIEN
W11_S91	W11	2017	Uruguay	Bovino	Prepucio	Cfv	FCIEN
W8_S90	W8	2017	Uruguay	Bovino	Prepucio	Cfv	FCIEN
X29_S92	X29	2017	Uruguay	Bovino	Prepucio	Cfv	FCIEN

Obtención del ADN genómico y librerías de secuenciación²

Los aislamientos fueron obtenidos a partir de diferentes fuentes (detalladas en la Tabla 3.2), cultivando las muestras en medio Skirrow sólido en condiciones de microaerobiosis por 48 horas a 37°C. Su identificación fue realizada mediante métodos bioquímicos, según se detalla en [137]. El ADN genómico utilizado para la secuenciación se obtuvo a partir de cultivos puros, utilizando el kit QIAamp DNA minikit (Qiagen, Inc., Valencia, CA). Las librerías para secuenciación Illumina se prepararon utilizando el kit "Nextera XT library prep kit", siguiendo las instrucciones del fabricante. La secuenciación *paried end* (PE) se realizó utilizando la plataforma Illumina MiSeq en el Institut Pasteur Montevideo, o bien, utilizando la plataforma HiSeq en el Instituto Sanger de Inglaterra. Paralelamente, el ADN extraído de las cepas 17144 y U10 fueron enviados a Macrogen Inc. (Seoul, Corea del Sur) para su secuenciación con tecnología PacBio RSII.

El trabajo plasmado en esta tesis parte de los datos crudos de secuenciación masiva obtenidos mediante tecnología Illumina o PacBio, cedidos por diferentes colaboradores. Su procesamiento se detalla a continuación.

3.3.2. Procesamiento de datos crudos de secuenciación masiva

Los *reads*, productos de secuenciación masiva obtenidos mediante la tecnología Illumina o PacBio, se procesaron según se detalla en la Figura 3.1 y en las secciones correspondientes a cada paso.

²La obtención de los aislamientos y su secuenciación no forma parte de esta Tesis. Sólo se describen a modo informativo.



Fig. 3.1: Procesamiento de datos de secuenciación masiva, obtenidos mediante tecnología Illumina y PacBio. En el bloque de la izquierda (A) se muestra el procesamiento de los *reads* cortos, desde la obtención de los *reads* hasta el ensamblaje, mientras que en el bloque de la izquierda (B) se muestra la obtención del ensamblaje mediante la metodología híbrida, utilizada en los casos en los que se dispusiera de ambos tipos de datos de secuenciación. Los últimos pasos de este flujo de trabajo, mostrados en el bloque C, aplican a ambas estrategias de ensamblaje y consisten en el análisis de calidad y la anotación de los genomas ensamblados. Ver detalles en el texto.

Tecnología Illumina

Los *reads* cortos, productos de la secuenciación masiva con tecnología Illumina fueron sometidos a un pre-procesamiento que incluyó el *pipeline* mostrado en la Figura 3.2 y en el bloque A de la Figura 3.1. Brevemente, los pasos seguidos fueron los siguientes:

- Identificación de *reads* contaminantes con Centrifuge v1.0.3-beta
 [138]: Centrifuge es ejecutado sobre cada par de archivos de secuenciación, para identificar, mediante un algoritmo que utiliza *kmeros*, el taxón al que se asemeja cada *read*. Los *reads* identificados como taxa diferentes a *C. fetus* fueron descartados.
- Análisis de calidad de *reads* con FastQC v0.11.9 [113]. Los resultados de este análisis son considerados para elegir la estrategia de filtrado de *reads* o posiciones de baja calidad que se ejecuta a continuación.
- 3. Filtrado de *reads* cortos y bases terminales de baja calidad, utilizando Trimmomatic v.0.36 [139]. Se ejecuta Trimmomatic para cada par de *reads*, utilizando los parámetros TRAILING=5 y MINLEN=50, que eliminan las bases de los extremos 3' de los *reads* con calidad menor a 5, y los *reads* de menos de 50 bases de largo, respectivamente.
- 4. FastQC es ejecutado sobre los *reads* filtrados, para comprobar la eficacia de la estrategia de filtrado utilizada.

Tecnología PacBio

Los datos de secuenciación masiva provistos por la empresa Macrogen Inc., correspondientes a los aislamientos U10 y 17144 fueron pre-procesados utilizando el protocolo "RS Subreads" del portal de análisis SMRT, con el software smrtanalysis_2.3.0.140936.p5.167094.



3.3.3. Ensamblaje de los genomas

Se utilizaron dos estrategias, según la disponibilidad de *reads* largos obtenidos mediante secuenciación PacBio:

Ensamblaje de reads cortos: Se utilizó SPAdes v3.5.0 [116] con parámetros por defecto y la opción careful, que permite minimizar el número de *mismatches* e *indels* cortos, y es recomendado para el ensamblaje de genomas pequeños como los bacterianos. Posteriormente se aplicó el *pipeline* Assembly Improvement v1.160490 [120], que utiliza los *reads* y los *contigs* para producir *scaffolds* y rellenar *gaps*, utilizando para esto los programas SSPACE [121] y GapFiller [122]. En este paso también se removieron los *contigs* menores a 1000 bases.

Ensamblaje híbrido: Los *reads* largos y cortos, previamente filtrados, fueron la base del ensamblaje híbrido, que se efectuó utilizando Unicycler v.0.4.7 [118], con el modo bold, que maximiza la unión de los *contigs*.

3.3.4. Post-procesamiento de genomas ensamblados

Los pasos que comprenden el bloque C en la Figura 3.1 son comunes para ambas estrategias de secuenciación, y se describen a continuación:

Control de calidad de los genomas ensamblados Se evaluó el desempeño de las estrategias empleadas utilizando el programa Quast v.4.3 [140], revisando las métricas relacionadas al número y largo de los *contigs*, número de bases indeterminadas y el N50. También se utilizó Checkm v.1.1.3 [141] para establecer el grado de completitud y contaminación de las secuencias.

Anotación de los genomas La anotación de los genomas se realizó con el programa PROKKA v. 1.14.6 [125], utilizando opciones por defecto.

Determinación de los sitios de inicio y final de replicación en secuencias completas

La función oriloc [142], perteneciente al paquete de R Seqinr [143] se empleó para determinar el origen y el término de replicación en un genoma completo, utilizando como *input* el genoma anotado.

Orden y orientación de contigs

Para determinar el orden y las orientaciones más probables de cada *contig* en un genoma poco fragmentado se ejecutó Mauve Contig Mover (MCM) [144], utilizando como genoma de referencia a la secuencia cfv84112 (GCA_000967135.1), asumiendo ausencia de reordenamientos.

Comparación de secuencias con plásmidos conocidos

Para caracterizar *contigs* plasmídicos en el ensamblado híbrido, se realizó una búsqueda utilizando BLASTN [145, 146], y una base de datos se secuencias nucleotídicas en la que se incluyeron secuencias pertenecientes al género *Campylobacter* (TaxID:194). Para visualizar el mejor *hit* obtenido se utilizó Artemis v18.1.0 y ACT [147, 148].

3.4. Resultados

3.4.1. Origen de los aislamientos

Se obtuvieron datos crudos de secuenciación masiva de 24 aislamientos de *C. fetus*, procedentes de diferentes orígenes y fuentes, entre los años 1962 y 2018, de los cuales 18 son de origen nacional y 6 de origen extranjero (Argentina, Ecuador, Japón y Reino Unido). En la Tabla 3.2 se detallan los datos correspondientes a cada aislamiento.

Los aislamientos nacionales no clínicos fueron obtenidos en el marco del proyecto financiado por la ANII FSSA X 2014 1 105252, "Desarrollo y validación de metodologías para el diagnóstico y control de la campylobacteriosis genital bovina", o brindados por diferentes instituciones: Instituto de Investigación Agropecuaria (INIA, cuatro aislamientos), Instituto Nacional de Tecnología Agropecuaria (INTA, Argentina, dos aislamientos), División de Laboratorios Veterinarios (DILAVE, un aislamiento) y de la Colección Americana de Cultivos Tipo (cepa ACTC-19438, cedida por la empresa Zurgen). Estos aislamientos fueron obtenidos a partir de muestras de fetos abortados, placenta, raspaje prepucial o mucus vaginal de bovinos u ovinos. Los 5 aislamientos nacionales de origen humano fueron obtenidos a partir de pacientes infectados ingresados en diferentes hospitales locales. En el Capítulo 6 se profundiza sobre el origen y análisis genómico de estos aislamientos. Los 4 aislamientos extranjeros de origen humano fueron brindados por diferentes colaboradores de Ecuador y Japón, respectivamente [135, 136], y por Novogen.

3.4.2. Métricas de los datos de secuenciación masiva

Los experimentos de secuenciación masiva con tecnología Illumina dieron como resultado *reads* pareados de entre 30 y 151 bases, con contenidos GC de entre 32 y 35 %. Los números mínimos y máximos de *reads* obtenidos en cada experimento fueron de 1,57 y 11,90 millones de bases respectivamente, con una media de 5,06 millones de bases. Las métricas de los *reads* pre y post filtrado se detallan en la Tabla 3.3.

	S	in filtrar		Filtrados			
Genoma de ºCº. fetusº	# reads	Rango de longitudes (pb)	%GC	# reads	Rango de longitudes (pb)	%GC	% reads retenidos
15031_S95	2.405.372	30-151	35	2.347.892	50-151	35	97,61
15671_S81	3.499.250	30-151	35	3.452.622	50-151	35	98,67
17133_S82	3.531.056	30-151	34	3.453.244	50-151	34	97,80
17144_S83	2.886.128	30-151	35	2.842.970	50-151	35	98,50
2740_S96	2.167.548	30-151	35	2.119.896	50-151	35	97,80
95258_S94	1.979.508	30-151	35	1.947.770	50-151	35	98,40
ATCC_S93	3.230.964	30-151	35	3.163.510	50-151	35	97,92
Felv5	8.298.984	150	33	8.297.412	150	33	99,98
FETOS	10.769.016	150	33	10.768038	150	33	99,99
H1_S84	3.356.416	30-151	34	3.265.054	50-151	34	97,28
HB74	11.906.316	150	33	11.905.724	150	33	100
HC	9.774.386	150	33	9.773.896	150	33	99,99
HC71	9.847.238	150	33	9.844.374	150	33	99,97
HC73	9.524578	150	33	9.524.254	150	33	100
KCH09208	8.788.712	150	32	8.780.254	150	32	99,9
KCH09951	9.809.712	150	32	9.803.232	150	32	99,94
M8_S86	9.809.430	30-151	34	2.309.626	50-151	34	98,45
HM1	2.644.366	30-151	35	2.580.048	50-151	35	97,57
R18_S87	2.511.128	30-151	34	2.471.762	50-151	34	98,43
U10_S88	2.444.254	30-151	35	2.396.242	50-151	35	98,04
V28_S89	1.570.500	30-151	35	1.544.088	50-151	35	98,32
W11_S91	2.529.778	30-151	35	2.562.424	50-151	35	97,34
W8_S90	2.937.712	30-151	34	2.854.116	50-151	35	97,15
X29_S92	2.708.414	30-151	34	2.641.712	50-151	34	97,54

Tabla 3.3: Análisis de calidad de reads cortos pre y post-filtrado

3.4.3. Preprocesamiento de reads cortos

Identificación de reads contaminantes

La primera evaluación de los *reads* consistió en identificar contaminación, es decir, buscar *reads* que tienen baja probabilidad de pertenecer al organismo secuenciado. Por ejemplo, pueden identificarse *reads* de origen humano, producto de contaminación durante la manipulación de la muestra. Este análisis, efectuado con Centrifuge, no devolvió cantidades significativas de *reads* contaminantes, por lo cual se prosiguió con el siguiente paso del pipeline mostrado en la Figura 3.2.

Análisis de calidad de los reads

El análisis de calidad de los *reads* se realizó ejecutando FastQC sobre cada archivo *fastq*. Este programa devuelve un reporte para cada archivo, donde se exponen diferentes métricas cuya evaluación permite definir una estrategia de filtrado. En los reportes obtenidos se observó una buena calidad en general.

Filtrado de reads por calidad

A pesar de la buena calidad en general obtenida en los reportes producidos por FastQC, se realizó un *trimming* en los extremos 3' de los *reads*, que se caracterizan por presentar una disminución en la calidad, además de eliminar los *reads* de longitud menor a 50 bases. Los *reads* resultantes tuvieron una longitud de entre 50 y 151 bases, con una remoción de entre 0 y 2,85 % de los *reads* para cada experimento de secuenciación (ver Tabla 3.3).

Pre-procesamiento de reads largos

Dos aislamientos, uno obtenido a partir de muestras extraídas de la placenta de un aborto ovino y el otro a partir del raspaje prepucial de un toro, fueron seleccionados para resecuenciar utilizando la tecnología de secuenciación PacBio SMRT (Single Molecule Real Time), que fue realizada por la empresa Macrogen Inc. Para la cepa 17144, proveniente de placenta de un aborto ovino, se obtuvieron 143.668 *reads*, con una longitud media de 10.704 bases. La secuenciación de la cepa bovina U10 generó 150.292 *reads* de 7.141 bases en promedio. La estrategia de filtrado, llevada a cabo en el portal de análisis SMART, retuvo más de 90% de los *reads* iniciales. El largo de los *reads* luego del pre-procesamiento aumentó 44% para 17144 y 100% para U10, con aumentos también significativos en la calidad de las secuencias. Las métricas de los *reads* largos antes y después del filtrado se detallan en la Tabla 3.4.

3.4.4. Ensamblado y control de calidad

Los ensamblados de los genomas fueron realizados utilizando SPAdes. Los genomas obtenidos únicamente a partir de *reads* cortos se procesaron adicionalmente con el programa Improve_assembly, con el fin de minimizar la

	17144			U10		
	Pre-	Post-	Diferencia	Pre-	Post-	Diferencia
	filtrado	filtrado	(%)	filtrado	filtrado	(%)
# bases (×10 ⁶)	1608,86	1563,02	-2,85	1073,25	982,96	-8,41
# reads	143.668	101.447	-29,39	150.292	68.739	-54,26
N50 (bases)	21.336	21.464	0,60	19.522	20.102	2,97
Longitud media de <i>reads</i> (bases)	10.704	15.407	43,94	7.141	14.299	100,24
Calidad de reads	0,58	0,844	45,50	0,43	0,86	100,00

Tabla 3.4: Análisis de calidad de reads largos pre y post filtrado

fragmentación de los genomas y eliminar *contigs* menores a 1000 bases de longitud.

Según los reportes de calidad de los ensamblados producidos por Quast se obtuvieron genomas de entre 26 y 80 *contigs*, con largos totales de entre 1,74 y 1,95 Mb y contenidos GC de entre 32,94 y 34,11 %. En cuanto a los parámetros de calidad determinados con Checkm, vemos rangos de entre

Genoma	# Contigs	Long. máx. por <i>contig</i> (pb)	Long. total (pb)	% GC	N50	Ns/100kb	Comp. (%)	Cont. (%)
15031_895	43	319.143	1.885.228	32,98	245.747	16	98,9	2,03
15671_S81	39	318.012	1.872.384	33,03	245.736	11	98,75	1,19
17133_S82	45	280.218	1.877.419	33,02	190.791	16	98,86	1,22
17144_S83	27	594.528	1.776.353	33,2	236.884	11	95,77	0,36
2740_S96	44	280.040	1.877.491	33,04	190.791	17	98,86	1,22
95258_S94	57	320.997	1.917.338	33,11	100.937	11	98,9	1,19
ATCC_S93	35	258.900	1.868.561	33,2	136.975	6	96,36	0,61
Felv5	34	601.262	1.745.241	33,27	225.696	0	97,61	0,71
FETOS	80	685.110	1.832.423	34,11	238.163	0	95,39	4,6
H1_S84	44	280.216	1.879.095	33,04	190.792	16	98,86	1,22
HB74	30	594.537	1.750.479	33,2	236.882	0	94,57	1,89
HC	29	601.275	1.754.052	33,08	238.709	0	96,36	0,46
HC71	26	343.398	1.741.168	33,25	232.617	0	96,28	0,36
HC73	31	601.275	1.751.651	33,1	238.295	0	96,36	0,62
HM1	29	650.726	1.808.299	33,23	261.814	22	97,61	0,41
KCHO9208	56	600.120	1.795.576	33,56	225.711	0	96,36	2,57
KCHO9951	51	322.444	1.760.408	33,45	205.469	0	96,36	1,53
M8_S86	39	318.013	1.872.462	33,03	245.736	6	98,75	1,19
R18_S87	48	318.011	1.884.047	33,05	175.925	6	98,75	1,19
U10_S88	40	281.622	1.881.976	33,02	190.790	6	98,86	1,22
V28_S89	49	281.057	1.875.107	33,03	190.791	11	98,83	1,22
W11_S91	46	281.055	1.872.393	33,04	190.791	16	98,86	1,3
W8_S90	41	323.079	1.884.872	33,01	227.430	6	98,86	1,19
X29_S92	46	280.040	1.903.166	32,94	190.791	6	98,86	1,22
Min	1	258.900	1.741.168	32,9	100.937	0	94,57	0,4
Median	40,5	321.720	1.872.388,5	33,1	226.570,5	6	98,8	1,2
Mean	39	480.829	1.841.229,3	33,2	285.076,9	7	97,6	1,3
Max	80	1.849.237	1.955.536	34,1	1.849.237	22	98,9	4,6
U10_S88 (H)	5	917.244	1.955.536	33,06	508.499	0	98,86	2,47
17144_S83 (H)	1	1.849.237	1.849.237	33,28	1.849.237	0	95,7	1,61

Tabla 3.5: Análisis de calidad de los genomas ensamblados.

(H): ensamblado híbrido. Long.: Longitud. Comp.: Completitud. Cont.: Contaminación

	Mín	Q1	Mediana	Media	Q3	Máximo
CDSs	1724	1823	1926	1902	1972	2059
ARNr	2	2	2	2,95	4	6
ARNt	36	40	42	41,69	43	49

Tabla 3.6: Medidas resumen de las métricas de anotación de los genomas

Q1 y Q3 se refieren al primer y tercer cuartil, respectivamente.

94,57 y 98,90 % de completitud y entre 0,36 y 4,6 % de contaminación. Los detalles para cada genoma ensamblado se encuentran en la Tabla 3.5.

Métricas de Anotación

Se extrajo la información referente al número de CDSs y ARNs ribosomales y de transferencia a partir del resultado de la anotación de los genomas. Las medidas resumen se muestran en la Tabla 3.6.

3.4.5. Procesamiento de los ensamblados híbridos

Los aislamientos U10 y el 17144 fueron seleccionados para complementar su secuenciación con datos obtenidos con PacBio, con el fin de lograr genomas cerrados. A continuación se describen las diferencias en los productos obtenidos mediante ambas estrategias de ensamblaje.

Cepa C. fetus subesp. fetus 17144

La cepa de *C. fetus subesp. fetus* 17144 fue responsable de al menos un aborto ovino en el departamento de Colonia. El estudio de este caso fue publicado por Dorsch et al. en el 2022 [149].

Origen de replicación: Este genoma pudo completarse en el ensamblaje híbrido, pero no fue posible determinar de forma automática el sitio de origen de replicación. Se utilizó para este fin la función oriloc, perteneciente al paquete de R Seqinr. Brevemente, oriloc considera la densidad de CDSs y los desvíos en los contenidos de GC para determinar el origen y el final de replicación de una secuencia de ADN. Para esto se asume que el cromosoma al que se aplica tiene un único origen a partir del cual la replicación ocurre en forma bidireccional, y que existe asimetría en los patrones de sustitución entre las hebras de ADN [142]. Según los resultados obtenidos, que se muestran en la Figura 3.3, los sitios de inicio y final de la replicación predichos se encuentran alrededor de las posiciones 650.000 y 1.550.000 respectivamente. La aplicación de BLASTn contra la base de datos de nucleótidos no redundantes de las regiones circundantes al origen de replicación predicho permitió encontrar los genes que codifican para las proteínas DnaA, DnaN y GyrB, cuya presencia se asocia a las regiones de inicio de la replicación [150].

El ensamblado híbrido para esta cepa resultó entonces en un único *contig* compuesto por 1.849.237 pb, cuyo mapa circular es mostrado en la Figura 3.4.

Calidad: El análisis de calidad realizado con checkM determinó que este genoma presenta 1,61 % de contaminación y 95,7 % de completitud.

Anotación: La anotación del genoma completo de la cepa 17144 identificó las tres copias del operón ribosomal, además de 44 tRNAs (aumento de 9 tRNAs respecto al ensamblado simple) y 1851 CDSs, lo que significó un aumento neto de 63 CDSs.



Cepa C. fetus subesp. venerealis U10

Esta cepa de Cfv de origen bovino, provino de un establecimiento que registraba bajo rendimiento reproductivo [2].



Ensamblado: El ensamblado híbrido dio como resultado 5 *contigs* y un N50 superior a las 500.000 pb, lo que supone aproximadamente el 28 % del largo medio de un genoma de *C. fetus* y un aumento neto de 73.560 pb con respecto al ensamblado con *reads* cortos.

Orden y orientación de los contigs: Los tres contigs más largos se clasificaron como incompletos, por lo que fueron procesados con Mauve Contig Mover, con el fin de ordenar y reorientarlos, utilizando el genoma completo **GCA_000967135.1** como referencia. En la Figura 3.5 se muestra el alineamiento de los contigs del genoma de la cepa U10 a diferentes regiones del genoma utilizado como referencia, información que se utiliza para reordenar y reorientar los contigs. Los contigs en el genoma de la cepa U10 fueron procesados en base a estos resultados. En la Figura 3.6 se muestra el mapa genómico resultante.

Contigs plasmídicos: Los dos *contigs* restantes de menor longitud fueron cerrados automáticamente por Unicycler, aunque no se encontraron en ellos los sitios de inicio y final de la replicación. Estos elementos circulares, de 59,666 y 4,436 pb respectivamente, fueron sometidos a una búsqueda mediante BLASTN contra una base de datos de nucleótidos que incluye sólo secuencias plasmídicas pertenecientes al género *Campylobacter* (taxid:194), con el fin de conocer si estas secuencias corresponden a plásmidos ya descritos en otras cepas. Según los resultados obtenidos, el *contig* mayor se alinea de forma discontinua en un 65% de su extensión y con una identidad cercana al 95% con el plásmido pCfviMP1 (CP007000.1), de 91,3 kb de longitud, perteneciente a la cepa *C. fetus subsp. venerealis* cfvi03/293 (CP007000.1) (Ver Figura 3.6 y Figura 3.7). La búsqueda mediante BLASTN del *contig* menor tuvo como mejor *hit* al plásmido pCFViADRI1362_P4 (CP059436.1) de la cepa *C. fetus* CFViADRI1362, con una cobertura de 76% y una identidad superior al 99,9%.

Anotación: Se encontraron diferencias en la anotación de los genomas de la cepa U10: el ensamblado híbrido representó una ganancia neta de



60 genes, 3 tRNA y 2 operones ribosomales con respecto al ensamblado simple.

Calidad: Los índices de completitud determinados con checkM fueron muy similares, pero en el ensamblado híbrido se encontró el doble de contaminación que en el simple.



Fig. 3.6: Mapa del genoma de la cepa bovina U10. Se muestran a la izquierda los tres *contigs* obtenidos, cada uno con los CDSs en la hebra *forward* (azul oscuro) y en la *reverse* (celeste) y el GC-*skew*, con valores por encima y por debajo del promedio del *contig*. En el recuadro de la derecha se muestran los plásmidos cerrados obtenidos (no a escala).



3.5. Discusión y conclusiones

Se obtuvieron los genomas de 24 cepas de *C. fetus*

En este capítulo se describe la obtención de 24 secuencias genómicas de *C. fetus* tomando como punto de partida los datos de secuenciación masiva mediante tecnología Illumina de ADN genómico de aislamientos de *C. fetus* de origen clínico y veterinario. A continuación se discuten los resultados obtenidos según su origen.

Genomas de aislamientos clínicos:

Los aislamientos clínicos provienen de individuos inmunocomprometidos que desarrollaron una infección sistémica. El estudio de los casos desarrollados en Montevideo se describe detalladamente en el Capítulo 6. Tres de los casos producidos fuera del país fueron documentados en trabajos publicados por Kitamura et al. y Chávez et al. respectivamente [136, 135].

Se obtuvieron en total 9 genomas de *C. fetus* de origen clínico, mayoritariamente Cff según los metadatos disponibles, a partir de datos de secuenciación masiva Illumina. Todos resultaron en ensamblados con buenos índices de completitud y contaminación, con la excepción del genoma de la cepa de origen argentino, para el que se obtuvieron 80 *contigs*, el máximo en este trabajo, e índices de contaminación y completitud de 95,39 y 4,6 % respectivamente. En este genoma además se encontró el mayor contenido de GC, 0,9 % superior a la media de los demás genomas aquí obtenidos, por lo que no puede descartarse la presencia de contaminación con material genético de especies de mayor contenido GC.

Genomas bovinos

Se ensamblaron en total 10 genomas bovinos, todos de origen uruguayo y tipificados como Cfv. Se secuenciaron dos cepas aisladas de abortos bovinos y dos cepas del tracto vaginal de vacas. Las 7 cepas restantes provienen de muestras de raspaje prepucial de toros, obtenidas en el marco del proyecto "Desarrollo y validación de metodologías para el diagnóstico y control de la campylobacteriosis genital bovina", el cual se describe en la Capítulo 1.

La combinación de tecnologías de secuenciación genera genomas de mejor calidad

Según lo descrito anteriormente, la calidad de los ensamblados híbridos fue significativamente mayor, viéndose reflejada en el número de *contigs*, la longitud de los genomas y sus anotaciones.

Número de contigs: Los genomas fragmentados se componen de un número de *contigs* mayor al número de replicones en la célula. Hay varias causas que impiden la obtención de la secuencia de un replicón completo. Por ejemplo, una librería se secuenciación no uniforme dará lugar a coberturas desiguales, generando regiones de muy baja o nula cobertura que impiden el ensamblado. Otro factor, quizás el más influyente, es la presencia de repetidos o regiones de baja complejidad en los genomas:

Presencia de repetidos y *reads* cortos: Los repetidos producen estructuras similares a burbujas en los grafos que impiden la resolución de un camino único en el ensamblado. Este es un problema especialmente importante

en ensamblados generados a partir de tecnologías de secuenciación que producen *reads* cortos, debido a que el largo de las regiones repetidas es generalmente mayor a la longitud de los *reads*. La secuenciación PE es útil para minimizar el problema sin recurrir al uso de una plataforma de secuenciación de tercera generación, siempre que el largo del inserto lo permita, pero generalmente no es suficiente. La aplicación de *pipelines* de mejoramiento de los ensamblados también son útiles para mejorar su calidad, ya que permiten aumentar la longitud de los *contigs* e incluso disminuir su número, pero en algunos casos puede ser necesario el curado manual de los ensamblados [123].

Repetidos y *reads* largos: Al utilizar tecnologías que generan *reads* largos, el problema de la fragmentación es eliminado o al menos minimizado. El principio es simple: un *read* largo puede abarcar con facilidad una región repetida o de baja complejidad y sus alrededores, que permiten el solapamiento con otros *reads*, y por lo tanto aumentar la probabilidad de encontrar un camino único en los grafos de ensamblado.

Largo de los *contigs*: En ambos ensamblados híbridos se obtuvo una ganancia neta de alrededor de 73.000 pb y más de 60 CDSs. Los genomas fragmentados generalmente contienen genes truncados en los extremos de los *contigs*, que no logran recuperarse durante la anotación. Al aumentar el largo de los *contigs* y por ende disminuir su número, son menos los genes afectados por la discontinuidad de la secuencia, razón por la cual se encontró un número significativo de CDSs adicionales en los ensamblados híbridos.

Completitud de los genomas: Este parámetro se estima en base a la presencia de un *set* de genes precomputados para determinado taxón, que

tienen las características de ser ubicuos y de copia única. Es una práctica común estimar la completitud de los genomas utilizando checkM [141]. Para el caso de *C. fetus* este programa utiliza en total 1.143 genes marcadores. Sin embargo ninguno de los genomas utilizados en esta tesis resultó en un 100% de completitud, ya que para todos ellos faltaron al menos 12 marcadores. Incluso el genoma de referencia GCF_011600945.2 presenta una completitud³ de 98.56%. Debido a que un genoma con más de 90% de completitud ya se considera de buena calidad, no se consideran significativos los valores reportados por CheckM [141].

Operones ribosomales: A diferencia de los ensamblados simples, en los ensamblados híbridos fue posible recuperar las tres copias del operón ribosomal, característicos de la especie *C. fetus* [152], parámetro que refleja la mayor completitud en estos genomas.

Desventajas del ensamblado híbrido: Si bien esta estrategia es fácil de implementar en cuanto a su procesamiento computacional, supone un costo económico alto, ya que cada muestra es secuenciada en dos plataformas de secuenciación diferentes, requiriendo entonces una mayor cantidad de la muestra a secuenciar, más insumos y recursos humanos capacitados para ambas tecnologías.

Afortunadamente la calidad de los *reads* producidos por las plataformas de Nanopore y PacBio ha aumentado significativamente en los últimos años, al incorporarse mejoras tanto en la química de secuenciación como en el *base calling*. Nanopore anuncia una tasa de error menor a 1 % al utilizase la química R10.4.1, mientras que en Illumina se estima una tasa de error menor al 0,1 % en todas sus plataformas actuales. Las mejoras implementadas en

³Según reporte del NCBI realizado con checkM v1.2.2

las tecnologías de secuenciación de tercera generación en conjunto con una mayor profundidad no tardarán en ser suficientes para obtener genomas bacterianos completos sin implementar abordajes híbridos.

Aporte de los genomas generados

C. fetus en las bases de datos genómicos: situación actual

Una búsqueda utilizando el término "Campylobacter" en la base de datos de Genomas del NCBI dio como resultado la presencia de 110.975 genomas, mientras que la búsqueda combinada de *C. coli* y *C. jejuni* generó 107.519 registros, lo que representa el 96,89 % del total de los genomas del género. De los 3.456 genomas restantes, 478 pertenecen a *C. fetus* (460 genomas si se excluyen los genomas atípicos), y sólo 31 de ellos están completos.

Uruguay cuenta actualmente con 4 genomas de *C. fetus* depositados en la base de datos del NCBI, incluyendo el de la cepa ovina descrito en este trabajo (Ver Tabla 3.7).

Сера	Accession	Hospedero	Referencia
H1UY	GCF_001399955.1	Humano	[153]
HC1	GCF_003426005.1	Humano	[154]
HC2	GCF_003426015.1	Humano	[154]
17144	GCF_007723545.1	Ovino	Este trabajo, [137]

Tabla 3.7: Genomas de C. fetus uruguayos depositados en el NCBI

Aporte de los genomas de aislamientos uruguayos

Los genomas obtenidos en este trabajo representan un aumento del 500 % en la disponibilidad de genomas de aislamientos de *C. fetus* de origen nacional, sumando 21 genomas en total.

El aumento de secuencias genómicas de *C. fetus* tiene el potencial para contribuir al establecimiento de un sistema de vigilancia genómica nacional, importante para describir la situación epidemiológica del país en relación a *C. fetus*. El conocimiento generado tiene potencial para descubrir posibles fuentes y rutas de transmisión de la enfermedad, permitiendo a su vez construir estrategias de prevención y contención efectivas y discretas.

El análisis de los genomas permitirá conocer, por ejemplo la presencia y distribución de factores de virulencia y genes de resistencia a antibióticos. Estas y otras características nos permitirán visibilizar a *C. fetus* en el ámbito clínico, donde su frecuencia es subestimada debido a las características del proceso de diagnóstico.

Aporte de la secuencia completa de la cepa 17144

El genoma completo de la cepa 17144 de origen ovino fue depositado con el nombre INIA/17144 y el *accession* GCF_007723545.1 y se encuentra disponible desde julio de 2019.

Este genoma constituyó un gran aporte a la base de datos, ya que es uno de los 31 genomas completos de *C. fetus*. Al filtrar por hospedero, vemos que este es uno de los 7 genomas de origen ovino, de los cuales sólo 3 (incluído 17144) están completos. Además se publicó el anuncio de este genoma [137], que se muestra a continuación.



GENOME SEQUENCES



Complete Genome Sequence of *Campylobacter fetus* Isolated from a Sheep

Daniela Costa,^{a,b} Virginia Aráoz,^c Maila Barcellos,^a Rubén Darío Caffarena,^c Martín Fraga,^c Federico Giannitti,^c Cecilia Monesiglio,^c Ruben Pérez,^a Caroline da Silva Silveira,^c ^(D) Lucía Calleros^a

^aSección Genética Evolutiva, Instituto de Biología, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay
 ^bLaboratorio de Genómica Microbiana, Institut Pasteur de Montevideo, Montevideo, Uruguay
 ^cPlataforma de Investigación en Salud Animal, Instituto Nacional de Investigación Agropecuaria Estación Experimental La Estanzuela, Colonia, Uruguay

ABSTRACT *Campylobacter fetus* is an important reproductive pathogen of ruminants that occasionally infects humans. Here, we describe the complete circularized genome of a strain of *Campylobacter fetus* subsp. *fetus* isolated from a sheep. The final assembly consisted of a unique contig with a length of 1,849,237 bp.

C ampylobacter species are *Epsilonproteobacteria* adapted to vertebrate hosts. Many of them cause disease in a wide range of livestock species and have extensive reservoirs in wildlife. *Campylobacter fetus* is an important animal pathogen and an opportunistic human pathogen. It produces considerable economic losses as a major reproductive pathogen of cattle and sheep. This species is currently divided into *C. fetus* subsp. *fetus*, *C. fetus* subsp. *venerealis* (1), and *C. fetus* subsp. *testudinum* (2). *C. fetus* subsp. *fetus* causes abortion mainly in sheep (3) and, to a lesser extent, cattle.

Here, we describe the closed whole genome of a strain of *C. fetus* subsp. *fetus* isolated from a sheep, representing an important resource for evolutionary and epidemiological studies of the species.

Fresh placenta was received at the Instituto Nacional de Investigación Agropecuaria (INIA) Animal Health Platform (Uruguay) for bacteriological analyses. The sample was spiked into Skirrow agar and incubated in a microaerobic atmosphere for 48 h at 37°C using CampyGen (Oxoid) (4). Biochemical testing was performed, including catalase, oxidase, and hydrogen sulfide production (triple-sugar iron medium), growth in the presence of 3.5% sodium chloride or 1% glycine, and growth at 25 and 42°C in Skirrow agar (4). Results were consistent with those of *Campylobacter fetus* subsp. *fetus*.

Bacterial colonies from five petri dishes of pure culture were suspended in 500 μ l of phosphate-buffered saline solution (pH 7.4), and DNA was extracted using the QIAamp DNA minikit (Qiagen, Inc., Valencia, CA). Whole-genome sequencing (WGS) was performed with Illumina and Pacific Biosciences (PacBio) sequencing technologies. An Illumina library was prepared with the Nextera XT library prep kit without modifications. Illumina reads were obtained using a MiSeq platform at Institut Pasteur de Montevideo, obtaining 1,443,064 2 \times 150-bp paired-end (PE) reads. All bioinformatic analysis software was set to default values. Read quality was assessed using FastQC v0.11.7 (5). Bases at the ends of the reads with a quality score lower than 20 and reads shorter than 50 bases were trimmed using Trimmomatic v0.39 (TRAILING:20 MINLEN:50) (6). A total of 1,421,381 PE reads remained after filtration (mean length, 145 bases) (Table 1).

Long reads were obtained using PacBio RS II single-molecule real-time (SMRT) technology with P6-C4 chemistry (Pacific Biosciences) at Macrogen, Inc. (Seoul, South Korea). A library was prepared with the 20-kb SMRTbell template library kit without modifications. DNA was sheared, and 20-kb fragments were selected. Preprocessing

November 2020 Volume 9 Issue 45 e01008-20

Citation Costa D, Aráoz V, Barcellos M, Caffarena RD, Fraga M, Giannitti F, Monesiglio C, Pérez R, da Silva Silveira C, Calleros L. 2020. Complete genome sequence of *Campylobacter fetus* isolated from a sheep. Microbiol Resour Announc 9:e01008-20. https://doi.org/10.1128/ MRA.01008-20.

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2020 Costa et al. This is an openaccess article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Lucía Calleros, calleros@fcien.edu.uy. Received 28 August 2020 Accented 20 October 2020

Published 5 November 2020

Microbiology mra.asm.org 1

Costa et al.

Sequencing platform	Metric	Prefilter	Postfilter
Illumina MiSeq	No. of reads	1,443,064	1,421,381
	Read length range (bp)	30-151	50-151
	Read Q-score range	21–36	22–36
PacBio RS II	No. of polymerase read bases	1,608,855,393	1,563,018,757
	No. of polymerase reads	150,292	101,447
	Polymerase read N ₅₀ (bp)	21,336	21,464
	Polymerase read length (bp)	10,704	15,407
	Polymerase read quality	0.583	0.844

TABLE 1 Short- and long-read preprocessing summary

was done using the SMRT Portal (smrtanalysis_2.3.0.140936.p5.167094) RS Subreads 1 protocol. It obtained 1,608,855,393 bases from 150,292 reads, with an average read length of 10,704 bp and an N_{so} value of 16,155 bp. After filtering, 97% of these data remained (Table 1).

A hybrid assembly approach was performed with Unicycler v0.4.7 assembler (7), using *bold* mode. The quality of the assembly was addressed using Quast v4.3 (8).

Assembly resulted in a unique 1,849,237-bp contig with an average G+C content of 33.3%, as expected for *C. fetus* subsp. *fetus* genomes.

The sequences were annotated with NCBI Prokaryotic Genome Annotation Pipeline (PGAP v4.13) (9). The genome contains 1,884 coding sequences: 1,811 hypothetical proteins, 44 tRNA genes, 3 rRNA operons as in all other *Campylobacter* species, and 2 CRISPR arrays.

Data availability. These whole-genome sequencing, assembly, and raw data have been deposited in DDBJ/ENA/GenBank under the BioProject number PRJNA554155.

ACKNOWLEDGMENTS

This work was supported by the Uruguayan Agencia Nacional de Innovación e Investigación (grant ANII FSSA-105252 and Ph.D. fellowship for D.C.) and the Instituto Nacional de Investigación Agropecuaria (grants N-15156_PL_15_0_00 and N-23398).

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. We declare no conflict of interest.

L.C., F.G., M.F., and R.P. conceived the study. V.A. and R.D.C. collected the sample. M.B., C.M., C.D.S.S., and L.C. performed the experiments. D.C. assembled the genome. L.C. and D.C. wrote the manuscript. All the authors revised the manuscript.

REFERENCES

- Veron M, Chatelain R. 1973. Taxonomic study of the genus *Campylobacter* Sebald and Veron and designation of the neotype strain for the type species, *Campylobacter fetus* (Smith and Taylor) Sebald and Veron. Int J Syst Bacteriol 23:122–134. https://doi.org/10.1099/00207713-23-2 -122.
- 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. https://doi.org/10.1099/ijs.0.057778-0.
- Sahin O, Yaeger M, Wu Z, Zhang Q. 2017. Campylobacter-associated diseases in animals. Annu Rev Anim Biosci 5:21–42. https://doi.org/10.1146/annurev -animal-022516-022826.
- OIE. 2008. Bovine genital campylobacteriosis, p 661–670. In Manual of diagnostic tests and vaccines for terrestrial animals (mammals, birds and bees), 6th ed, vol 2. OIE, Paris, France.
- Andrews S, Lindenbaum P, Howard B, Ewels P. 2010. FastQC high throughput sequence version 0.11.7. http://www.bioinformatics.babraham.ac.uk/ projects/fastqc/.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10 .1093/bioinformatics/btu170.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genomics 3: e000132. https://doi.org/10.1099/mgen.0.000132.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/ 10.1093/bioinformatics/btt086.
- Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi .org/10.1093/nar/gkw569.

November 2020 Volume 9 Issue 45 e01008-20

95

A Microbiology

4

Estudio comparativo de genomas de *Campylobacter fetus* subsp. fetus y *Campylobacter fetus* subsp. venerealis

4.1. Introducción

4.1.1. Impacto de la disponibilidad de secuencias genómicas en la microbiología

El inicio de la era genómica marca un cambio de paradigma en el estudio de la microbiología. La profundidad del cambio generado se refleja en el número de genomas que se secuencian año a año y se depositan en las bases de datos públicas. Quizás uno de los cambios más profundos se dio en el área de la taxonomía y sistemática.

De un gen a un genoma

Históricamente la identificación de una bacteria se basó (y lo sigue haciendo) en base de una batería de pruebas fenotípicas sobre los cultivos puros. Un método que ha sido el *golden standard* en la identificación es el de hibridación ADN-ADN, una técnica molecular muy laboriosa y poco reproducible en la que se compara una cepa de referencia contra una cepa problema. Al comenzar a ser accesible la secuenciación, se optó por la secuenciación del gen que codifica la subunidad 16S del ARNr, un marcador de presencia universal en *Bacteria* y *Archaea*. Un 98 % de identidad de secuencia de este gen se correlaciona con una hibridación de ADN-ADN de 70 % [108]. Sin embargo, esta metodología no discrimina entre subespecies, debido a la poca variabilidad en la región secuenciada. La baja resolución de esta técnica se debe a que sólo cubre una región del genoma, razón por la cual se diseñaron métodos que tienen en cuenta la secuencia de fragmentos de más genes, como el MLST, descrito anteriormente (ver Párrafo 2.3.2).

La disponibilidad de recursos genómicos permitió establecer con mayor resolución los límites de especie, y hoy es posible la aplicación del índice de identidad nucleotídica promedio (ANI), al demostrarse que un 95 % de ANI se corresponde con el 98 % de identidad de secuencia del gen 16S [155].

De un genoma a genómica poblacional

A medida que se secuenciaban más genomas, se evidenciaba la creciente diversidad presente, incluso dentro de una misma especie. La disponibilidad de un único genoma para un taxón determinado era insuficiente para una descripción completa [108]. Por ende, es imperativo reconsiderar la sensibilidad y especificidad de los métodos de diagnóstico que se han diseñado a partir de un número limitado de genomas.

La diversidad genética dentro de un taxón o población se caracteriza mediante el concepto de pangenoma, introducido por Tettelin et al. en 2005. Este término describe el conjunto de genes que se pueden encontrar en una población o taxón determinado. Se definen tres categorías de genes o grupos de genes ortólogos en un pangenoma: los genes *core*, que abarcan aquellos presentes en todos o la mayoría de los individuos de la población; los genes accesorios, identificados en algunos genomas pero no en todos; y los genes individuales o *singletons*, específicos de un genoma en particular. Generalmente los genes que forman parte del genoma *core* codifican funciones básicas de las células, mientras que el genoma accesorio tendría una función en la adaptación de linajes a nuevos ambientes o condiciones particulares. Los *singletons* por otro lado, probablemente correspondan a material genético recientemente incorporado, que aún no se ha fijado en la población.

Además de la identificación de especies y subespecies, el análisis de todo el genoma permite extraer y comparar genes o grupos de genes asociados, por ejemplo, al potencial patogénico, mediante la identificación de factores de virulencia o resistencia a antimicrobianos. La aplicación de la genómica comparativa permite además establecer vínculos con diferentes cepas al reconstruir la estructura poblacional, que, en conjunto con la filogenómica, nos permite trazar su historia evolutiva.

4.1.2. *C. fetus*: del diagnóstico fenotípico a la genómica comparativa

El primer paso antes tomar medidas ante un proceso infectivo en curso consiste en la detección e identificación del agente patogénico, para lo cual es necesario contar con una metodología de diagnóstico confiable. Un diagnóstico rápido permite la adopción de medidas de control tempranas, que impiden el avance de la enfermedad en el individuo afectado, así como la propagación de la enfermedad en la población, para lo cual los métodos de diagnóstico microbiológico clásicos, basados en el aislamiento y cultivo del patógeno, pueden resultar inadecuados.

C. fetus es un claro ejemplo de esto, ya que la contaminación presente en las muestras a partir de las cuales se aísla el patógeno, así como sus necesidades nutricionales, dificultan la obtención de los cultivos puros necesarios para la tipificación, proceso que consiste en la realización de pruebas fenotípicas descritas en 1959 [61] que son recomendadas hasta el día de hoy [62].

Las técnicas de diagnóstico molecular basadas en PCR a tiempo final y en tiempo real constituyen alternativas muy atractivas en este escenario, ya que permitirían un diagnóstico rápido y objetivo, imprescindible para la vigilancia epidemiológica. Se han desarrollado varias herramientas diagnósticas basadas en PCR con el fin de distinguir entre las subespecies de *C. fetus*, las cuales fueron abordadas en la Subsección 2.3.2 (ver pág. 23).

Abordajes poblacionales

La efectividad de la mayoría de los métodos de diagnóstico desarrollados ha sido cuestionada, principalmente debido a que su diseño se basó en un número limitado de muestras o que estas provinieron de orígenes poco variados. Por ejemplo, Abril et al. no incluye ningún aislamiento Cfvi [99]; Iraola et al. y Moolhuijzen et al. diseñan la metodología en base a una cepa Cfv y una Cff [101, 100]; Van Bergen et al. se basan en el estudio de 65 cepas, de las cuales 12 fueron identificadas como Cfv y las 53 restantes como Cff, según el protocolo de AFLP de Wagenaar et al. [157], sin ningún representante sudamericano [90]. Por consiguiente, se requieren enfoques poblacionales que incluyan muestras de diferentes orígenes geográficos y diferentes hospederos para conseguir metodologías robustas que puedan ser implementadas a nivel global.

En 2005 se llevó a cabo uno de los primeros estudios poblacionales por Van Bergen et al., en el que se analizó la tipificación de *C. fetus* mediante MLST en relación al hospedero, origen geográfico, subespecie y tipo *sap*, sobre una colección de 140 aislamientos [90]. Los autores describen a *C. fetus* como una especie poco diversa, ya que la magnitud de la diversidad de STs encontrada en la especie es comparable a la de un único complejo clonal de *C. jejuni*. También describen asociaciones entre el ST y subespecie y hospedero, observando que los aislamientos pertenecientes a la subespecie Cfv se corresponden en su mayoría con los STs 4, 7 y 12, mientras que en Cff se encuentran más STs representados. En cuanto a la relación con el hospedero, los autores reportan que los STs 1 y 4 se asocian mayoritariamente a aislamientos de origen bovino. Estos resultados apoyan la hipótesis de Veron y Chatelain, que describe a Cfv como un clon derivado de Cff adaptado a bovinos [90].
4.2. Hipótesis y objetivos

4.2.1. Hipótesis

Los estudios poblacionales de las subespecies Cff y Cfv de *C. fetus* no han sido abordados con un número de genomas lo suficientemente alto para describir la mayor parte de la diversidad genética existente, y han resultado en metodologías moleculares incapaces de discriminar de forma robusta los genotipos asociados a su potencial patogénico.

Este trabajo surge de la hipótesis de que es posible derivar, a partir del análisis comparativo de un gran número de genomas de aislamientos de *C*. *fetus*, una estructura poblacional en la que se distinguen linajes con diferente asociación a hospedero.

La distinción entre los diferentes linajes permitirá el diseño de metodologías de diagnóstico molecular para predecir el potencial patogénico de nuevos aislamientos de *C. fetus*, que resulten más robustos y específicos que la distinción entre subespecies que se realiza rutinariamente, y es recomendada por la OMSA [62].

4.2.2. Objetivos

Objetivo general

Identificar características genómicas que permitan diferenciar los aislamientos de *C. fetus* origen bovino de los de origen humano y que puedan utilizarse como posibles blancos moleculares de diagnóstico de la CGB.

Objetivos específicos

- 1. Caracterizar la estructura poblacional de *C. fetus* a partir de una colección de genomas compuesta por datos públicos y derivados de nuestro trabajo, pertenecientes a aislamientos de diferentes orígenes geográficos y hospederos.
- Diseñar métodos moleculares que permitan genotipar aislamientos de *C. fetus*, localizándolo en la estructura poblacional detectada.
- Caracterizar los linajes, describiendo su diversidad genética, sus perfiles de recombinación, su pangenoma y determinar patrones de asociación a hospedero.

4.3. Metodología

4.3.1. Resumen de la estrategia utilizada

A continuación se resume la estrategia utilizada, la cual se esquematiza en la Figura 4.1. Los detalles de cada etapa se encuentran en las secciones correspondientes.

Se partió de un set de datos compuesto por secuencias genómicas de C. fetus y sus metadatos asociados, obtenidos de bases de datos de acceso público. En una primera instancia, se evaluó la calidad de los genomas, y se descartaron los que no superaron niveles aceptables de completitud y contaminación. Se caracterizaron los genomas con el fin de evaluar la diversidad de los datos recolectados en asociación con los metadatos disponibles. Se obtuvo el genoma core mediante el alineamiento de los genomas, y a partir de este se identificaron regiones con evidencia de recombinación, analizando su posición y extensión. Para obtener el marco clonal se filtraron del alineamiento los bloques recombinantes. Se infirió la estructura poblacional utilizando un método de clusterización jerárquica y mediante la construcción de la filogenia, y se caracterizaron los linajes obtenidos. Paralelamente, se anotaron los genomas y a partir estos se construyó el pangenoma, el cual fue caracterizado, evaluando entre otras cosas el número y la función de los genes accesorios en el marco de la estructura poblacional obtenida. Por último, se identificaron blancos moleculares capaces de diferenciar los linajes. Se evaluó in silico la especificidad y sensibilidad de los cebadores obtenidos, y se comenzó su evaluación in vitro.



4.3.2. Obtención del Dataset

El *dataset* utilizado en este trabajo contiene secuencias genómicas de *C. fetus*, públicas y disponibles en las bases de datos NCBI [158] y PATRIC [159, 160], así como también secuencias brindadas por colaboradores y otras obtenidas en este trabajo (Ver Capítulo 3). Los metadatos asociados a estos genomas se encuentran disponibles de forma interactiva en https://microreact.org/project/1azjvnxljc3xuhaghssek2.

Análisis de calidad

Antes de realizar los análisis poblacionales se comprobó la calidad de los genomas del *dataset*, según se describió en la página 74 (Control de calidad de los genomas ensamblados).

4.3.3. Caracterización de los genomas

ANI

Se obtuvieron los índices de identidad nucleotídica promedio (ANI), utilizando el programa FastANI v.1.33 [155], con parámetros por defecto, utilizando como referencia y como *query* todos los genomas del *set* de datos, de manera de obtener los valores de ANI para cada par de genomas. Los resultados obtenidos se procesaron y visualizaron utilizando los paquetes tidyverse y ggplot en R, y se relacionaron con los datos disponibles de hospedero y subespecie. Para determinar el *sequence type* (ST) [161] al que pertenece cada aislamiento se utilizó la herramienta MLSTar [162], y como referencia las secuencias almacenadas en la base de datos PubMLST [92].

4.3.4. Obtención del genoma core

Las secuencias genómicas se utilizaron para construir alineamientos de genoma entero (*Whole Genome Alignment-WGA*), con parsnp v.1.2 [163], utilizando el genoma GCA_000967135.1, correspondiente a la cepa de *C. fetus subsp. venerealis* 84/112 como referencia, agregando genomas completos de *C. fetus testudinum* como grupos externos. El alineamiento obtenido con parsnp se convirtió al formato fasta con harvestools, y se utilizó Trimal v.1.2 para eliminar de forma automatizada regiones mal alineadas.

4.3.5. Marco clonal, recombinaciones y filogenias

Las regiones con evidencia de recombinación fueron identificadas utilizando Gubbins v.2.4.1 [164], y se eliminaron de los alineamientos, para obtener el marco clonal.

A partir del marco clonal se construyó un árbol filogenético utilizando el método de máxima verosimilitud (ML) y la aproximación GTR-CAT en RaxML v8.2.11 [165]. Se determinaron las distancias en SNPs entre pares de secuencias de los alineamientos clonales utilizando snpdist v.0.7.0 [166]. La anotación y visualización de los árboles filogenéticos se realizaron en R, utilizando el paquete ggtree v.1.14.6 [167, 168]. Los *outputs* de Gubbins y RaxML también se visualizaron en Phandango [169]. Se utilizó la implementación en R de hierBAPS [170] RhierBAPS [171] con el fin de identificar agrupamientos de genomas, y la estructura poblacional así obtenida se comparó con la filogenia inferida a partir del marco clonal.

4.3.6. Anotación de los genomas

La anotación se realizó según se detalla anteriormente (Ver página 74).

4.3.7. Pangenomas

A partir de los genomas anotados se reconstruyeron los pangenomas utilizando Pewit v.1.1.0 (https://github.com/iferres/pewit). Se utilizó Pagoo v.0.3.13 [172] para analizar y visualizar los pangenomas. La anotación funcional de los grupos de ortólogos se realizó utilizando eggNOG-Mapper v.2 [173], sobre la base de datos eggNOG v.5.0 [174], y/o KEGG mapper v.4.3 [175], que utiliza datos provenientes de KEGG [176, 177]. Los análisis de correspondencias múltiples y de componentes principales se realizaron utilizando los paquetes de R FactoMineR v.2.4 [178] y stats v.4.2.0. Los análisis discriminantes de componentes principales (DAPC) se realizaron con el paquete de R adegenet v.2.1.5 [179]. El número de ejes retenidos fue determinado utilizando el método de validación cruzada implementado en la función xvalDapc del mismo paquete.

4.3.8. Diseño de cebadores específicos de linaje

A partir de los genomas agrupados en los linajes identificados se ejecutó el programa RUCS v.1.0.3 [180], con el fin de identificar posibles blancos moleculares y diseñar cebadores para PCR con alta especificidad y sensibilidad, utilizando para esto Primer3 [181]. La especificidad de los cebadores se comprobó *in silico* utilizando BLASTn [146], FastPCR [182] y Unipro UGENE v.43.0 [183] sobre genomas completos de diferentes linajes.

Optimización y validación de los cebadores diseñados

Para obtener la temperatura de annealing óptima se realizó una PCR en gradiente de temperaturas, con rangos entre 52 y 53°C para L1 y de entre 54 y 55°C para los cebadores específicos para L2. Con el fin de validar los cebadores diseñados, se realizaron amplificaciones mediante PCR, utilizando como molde muestras de ADN de 26 aislamientos de C. fetus secuenciados y 27 no secuenciados (Ver Tabla 4.4), pertenecientes al banco de muestras del Laboratorio de Genética Evolutiva de la Facultad de Ciencias, Universidad de la República. Se utilizaron las cepas Cfv06/195 y Cff70L como controles positivos para L1 y L2 respectivamente, y la cepa 17144 como control negativo en ambos casos. Las mezclas de reacción se prepararon según las especificaciones del fabricante de la Taq polimerasa utilizada, incorporando 0,05 nmoles de cada cebador y 0,5 μ l de ADN molde por reacción. El programa utilizado consiste en un ciclo incial de desnaturalización de 3 minutos a 95°C, seguido por 30 ciclos de desnaturalización (95°C 30 segundos), annealing (temperatura óptima según resultados de PCRs en gradiente, 35 segundos) y extensión (72°C, 1,5 minutos), finalizando con un ciclo de extensión de 5 minutos a 72°C. Los productos de PCR se separaron mediante electroforesis en gel de agarosa 1 % teñido con bromuro de etidio, y se visualizaron utilizando transiluminadores de luz UV.

4.4. Resultados

4.4.1. *Dataset*: distribución geográfica de los aislamientos, fuentes, hospederos

El *dataset* completo comprende genomas de aislamientos de *C. fetus* de origen clínico o animal disponibles en bases de datos públicas, o secuenciados en el marco del proyecto FSSA_X_2014_1_105252, descrito en la Capítulo 1. Los metadatos asociados a cada uno de los genomas se encuentran disponibles en forma interactiva en https://microreact.org/project/1azjvnxljc3xuhaghssek2. En la Figura 4.2 se muestra la distribución geográfica de los aislamientos, donde el diámetro de cada burbuja representa el número de genomas secuenciados por país. Se observa allí, que si bien los cinco continentes están representados, la distribución es desigual, con la mayoría de los aislamientos provenientes de Europa, y en menor medida, de las Américas. Estos datos se representan también en la Figura 4.3 (B), donde se muestra el número de aislamientos por país, y puede desprenderse que el 90 % de estos provienen de sólo ocho países.

Por otra parte, la Figura 4.3 (A) muestra la distribución en cuanto al hospedero y la fuente del aislamiento, observándose que los hospederos predominantes son bovinos, con 156 aislamientos, y humanos, con 81 aislamientos, representado 53 % y 37 % del total, respectivamente. La principal fuente de aislamientos en bovinos fue el prepucio, mientras que en humanos la muestra predominante fue sangre.



Fig. 4.2: Origen geográfico de los 219 genomas que componen el *set* de datos, donde el diámetro de cada burbuja es proporcional al número de genomas por país.

Calidad de los genomas

La mayoría de los genomas de este conjunto de datos están ensamblados a nivel de *contigs*, con la excepción de 6 genomas cerrados, entre los que se encuentra la cepa ovina 17144, cuya obtención se describe en el Capítulo 3 y en el artículo "Complete Genome Sequence of *Campylobacter fetus* Isolated from a Sheep" [137] (Ver pág. 94).

El rango de *contigs* por genoma es de 5 hasta 759, con valores de N50 de entre 5000 y 500.000 bases para los genomas ensamblados a nivel de *draft*. Los resultados del análisis de calidad realizado con Checkm se muestran en la Tabla 4.1. Según estos, los genomas 17059_2.35 y 18048_2.27 tienen baja calidad de ensamblado, por lo cual fueron descartados en los análisis posteriores.



	Mín.	Q1	Mediana	Media	Q3	Máx.	Outliers
# Contigs	1	33	41	58,40	56	759	759[1] 722[2]
Completitud (%)	65,41	99,69	99,69	99,33	99,69	99,88	65,41[1] 76,41[2]
Contaminación (%)	0	0	0,18	0,14	0,18	2,69	-
Ns/genoma (%)	0	0,002	0,008	0,128	0,0033	11,00	11,00[1] 8,15[2]

 Tabla 4.1: Medidas resumen de las métricas de calidad de los genomas ensamblados, obtenidas con CheckM.

Los genomas con valores extremos referidos en la columna *Outliers* son [1] 18048_2.27 y [2] 17059_2.35. Mín=mínimo; Máx=máximo; Q1=primer cuartil, Q3=tercer cuartil.

4.4.2. Caracterización de los genomas

ANI

Se utilizó el índice ANI (identidad nucleotídica promedio) como medida de similaridad entre genomas. El valor medio general de ANI encontrado con FastANI fue de 99,64 % (Rango intercuartil (IQR):99,76-99,52 %). Se evaluaron también los índices de ANI de los genomas, agrupados por subespecie y hospedero. Las distribuciones de ANI obtenidas se muestran en la Figura 4.4 y se describen a continuación:

- **Subespecies:** Se obtuvieron los índices de ANI de pares de genomas pertenecientes a la misma subespecie y entre subespecies, encontrándose diferentes distribuciones, como se observa en la Figura 4.4a. La Tabla 4.2 muestra un resumen de los resultados.
 - Cfv tiene una distribución que se aproxima a la de una campana, con valores de ANI de entre 99,16 y 100 %, y una mediana de 99,77.
 - Cff presenta una distribución que permite inferir una mayor variabilidad que la encontrada en Cfv, con un ANI mínimo de 98,5 % y un

Grupo ¹	Mín.	Media	Mediana	Max.	IQR	
Intra Cff	98,5	99,6	99,6	100	0,248	
Intra Cfv	99,2	99,7	99,7	100	0,169	
Cff-Cfv	98,8	99,6	99,6	100	0,229	

 Tabla 4.2: Medidas resumen de valores de ANI por subespecie.

Mín=mínimo; Máx=máximo; IQR=rango intercuartil. ¹ Cfv incluye también al biovar intermedio Cfvi.

rango intercuartil más amplio. Se identifica un grupo de genomas con muy altos índices de identidad.

- Cfv-Cff: Al comparar de forma pareada genomas de Cff y Cfv encontramos el rango intercuartil y la media similar a Cff. La distribución en ANI indica que son pocas las Cff muy similares a Cfv.
- Hospedero: Al comparar los índices de ANI entre genomas de aislamientos provenientes del mismo hospedero, se observan distribuciones similares para los hospederos ovino y humano, que consisten en la presencia de dos picos separados, ubicados en valores de ANI de 99,5 y 100% respectivamente. Esta distribución es compatible con la presencia de aislamientos de estos hospederos en diferentes grupos o linajes. Al evaluar los índices de ANI entre genomas pertenecientes a bovinos, se observa una distribución en forma de campana, con una mediana de 99,7%, y un rango intercuartil más estrecho, que denota una menor variabilidad. A su vez, la ausencia de picos suplementarios indica que los aislamientos bovinos son más homogéneos desde el punto de vista genómico, según el índice evaluado.



Fig. 4.4: Curvas de distribución de ANI, agrupados por subespecie (Figura 4.4a) y por hospedero (Figura 4.4b). Los valores mostrados en cada distribución corresponde a las medianas.

MLST

Se determinaron los ST de 205 de 217 genomas con MLStar: Se detectaron 11 STs que forman parte de la base de datos de referencia, siendo ST4 el mayoritario. 12 de los genomas no pudieron ser tipificados, ya que no se identificaron secuencias correspondientes a los alelos en los genes *aspA* (11 genomas) o *glyA* (1 genoma), con umbrales de identidad y/o cobertura iguales o superiores al 90 % con respecto a las secuencias de los alelos de referencia. Además de estos, en dos genomas no fue posible asignar un ST debido a que los alelos identificados (*gltA* en un caso y *uncA* en otro) difieren de los depositadas en la base de datos, por lo que constituirían una novedad.

MLST y hospederos Los dos ST más frecuentes fueron ST4 en bovinos (87 genomas), y ST6 en humanos (41 genomas). Además, se encontraron otros 6 STs en bovinos y 9 en humanos, con lo que la diversidad encontrada en humanos es mayor a la de bovinos (índice de Shannon: 0,67 en bovinos frente a 1,61 en humanos). Además, en humanos se incluye un ST no reportado previamente. En la Figura 4.5 se muestra un gráfico que representa, para cada ST identificado, el número de genomas de aislamientos bovinos y humanos. Allí puede apreciarse la diversidad de STs representados en aislamientos humanos, en contraposición con los STs encontrados en aislamientos bovinos, en los que predomina el ST4. En ovinos en cambio, el ST2 fue más frecuente, seguido por ST3, y, al igual que en humanos, se identificó un genoma con ST desconocido. En la Tabla 4.3 se muestra un resumen de los STs encontrados.

	ST1	ST2	ST3	ST4	ST5	ST6	ST11	ST70	ST20	ST35	Nuevo
Bovino	0	1	3	86	3	4	2	2	0	0	0
Humano	1	4	7	5	2	41	6	0	13	1	1
Mono	0	1	0	0	0	0	0	0	0	0	0
Ovino	1	9	4	0	0	1	0	0	0	0	1

 Tabla 4.3: Número de genomas por ST por hospedero



Métricas de Anotación

Al anotar los genomas se identificaron las secuencias codificantes (CDSs), ARNs ribosomales y de transferencia. Los genomas contienen en promedio 1.868 CDSs (SD=104,33), con una densidad media de 1.015 CDSs por millón de bases. El número de CDSs anotados como proteínas hipotéticas fue de 657 (SD=85), lo que representa el 35 % de los CDSs identificados. Se hallaron entre 2 y 6 genes ribosomales por genoma, siendo 2,2 el valor promedio. El número máximo de copias del operón ribosomal se encontró en los genomas menos fragmentados, compuestos por entre 1 y 5 *contigs*.

4.4.3. Estructura poblacional

Obtención del Genoma core y el marco clonal

El genoma core se construyó a partir de 216 genomas de *C. fetus* ¹. Además, se incluyó un genoma de la subespecie *testudinum* como grupo externo. Este proceso resultó en un alineamiento de 1.339.999 posiciones, lo que representa un 61 % del genoma utilizado como referencia. Asimismo, se obtuvo el genoma *core* del mismo conjunto de genomas excluyendo el grupo externo, lo cual dio como resultado un alineamiento de 1.528.587 posiciones, es decir, el 75,63 % del genoma de referencia.

El marco clonal se compone de las posiciones del alineamiento del genoma *core*, donde se filtraron las posiciones con evidencia de recombinación. Se filtraron 33.500 posiciones con evidencia de recombinación, que representan en promedio el 2,5 % de las posiciones en el genoma *core*, formado por 1.339.999 posiciones.

Filogenia

El marco clonal, compuesto por 97.713 posiciones polimórficas, se utilizó para construir una filogenia utilizando el método de máxima verosimilitud (ML). Con el propósito de mejorar la definición de los grupos internos, se eliminó la rama correspondiente al genoma utilizado como grupo externo. Las

¹Se excluyeron 3 genomas del conjunto de datos original (18048_2.27, 17059_2.35 y FETOS) debido a la presencia de indicadores que señalaban baja calidad.



figuras 4.6 y 4.7 muestran representaciones del árbol filogenético obtenido, donde se revela una estructura poblacional compuesta por 8 grupos de genomas, con un rango de entre 1 y 105 integrantes cada uno. A excepción de los grupos L1 y L2, todos los linajes están compuestos por genomas pertenecientes a un único ST. En el caso de L1 prevalece el ST4, pero se incluyen además genomas no tipificados o pertenecientes al ST70. Por otro lado, el grupo L2 está mayoritariamente compuesto por genomas ST6, con una minoría de genomas ST4. Ambos grupos se encuentran resaltados en el recuadro de la Figura 4.6.

Clusterización jerárquica y caracterización de los linajes

La estructura poblacional derivada de la filogenia fue comparada con la obtenida mediante la ejecución de rhierBAPS sobre el marco clonal. Esta clasificación resultó también en la formación de 8 *clusters* (Ver Figura 4.6),

de entre 3 y 105 genomas, los cuales pasaremos a nombrar como linajes L1 a L8, descritos a continuación.



Fig. 4.7: Árbol filogenético construido a partir del alineamiento del marco clonal de 216 genomas de aislamientos de *C. fetus* de diferentes orígenes y sus metadatos asociados.

- L1 es el linaje mayoritario (105 aislamientos), y está formado principalmente por aislamientos de origen bovino, pertenecientes al ST4. Además se nota en este linaje un claro enriquecimiento en aislamientos de la subespecie Cfv y el biovar intermedio, no observado en los demás linajes.
- L2 está conformado por 53 aislamientos, principalmente de origen clínico tipificados como Cff ST6. En la filogenia se distingue un pequeño grupo conformado por 6 aislamientos ST4, de los cuales 5 son de origen clínico y uno es bovino.
- L3 y L5 son también linajes predominantemente de origen clínico, caracterizados como Cff ST11 en el caso de L3 y ST20 en L5. Los aislamientos de L5 provienen mayoritariamente de Taiwan, con la excepción de la cepa HC71 de origen nacional.
- L4 merece una mención especial, ya que está enriquecido en aislamientos de origen ovino, identificados como Cff, con excepción de un aislamiento bovino, tipificado como Cfv. En este linaje todos los aislamientos son ST2.
- L6 es un linaje de aislamientos humanos y bovinos pertenecientes a ST5 e identificados como Cff. La particularidad de este grupo es que comparte un ancestro común con un grupo de aislamientos de Cff ST1 uruguayos clasificados por hierBAPS como L7.
- L7 es el linaje que representa menos individuos y constituye la única incongruencia encontrada con respecto a la topología del árbol filogenético. Se trata de un grupo polifilético, ya que dos de los genomas ST1 son más cercanos al linaje L8, mientras que el tercer genoma es ST35 y es más cercano al linaje L6. La formación de este linaje se debe a un fenómeno conocido como atracción de ramas largas, que tiende a agrupar secuencias muy divergentes al resto de la población [170], y sería por tanto un artefacto metodológico. Para ser congruente con la filogenia

obtenida, se puede dividir este linaje para formar el linaje 7a, con los genomas ST1 y el linaje 7b con el genoma ST35.

L8 está conformado por aislamientos Cff ST5 de origen bovino y humano. Al igual que L6, este linaje comparte un origen común con un aislamiento del linaje L7b.

Asociaciones derivadas de los linajes.

Existe una asociación aparente entre algunos metadatos y los linajes determinados a partir de la filogenia, que puede observarse en la Figura 4.7, y se resumen a continuación:

- **Asociación a hospedero:** Resulta evidente la asociación del linaje L1 a bovinos, ya que no sólo es el hospedero predominante en este linaje, sino también porque son pocos los genomas bovinos en otros linajes y su distribución no parece estar asociada a la filogenia. Así es que el 87% de los genomas bovinos se encuentran en L1 y el 99% de los genomas de L1 son bovinos. Otro claro caso de asociación linaje-hospedero lo constituye el linaje L4, enriquecido en genomas de origen ovino.
- **Asociación a ST:** En la mayoría de los linajes se observó una correspondencia directa entre linaje y ST. Si bien L1 y L2 presentan genotipos compuestos, ambos presentan un genotipo predominante que permite trazar asociaciones robustas entre linaje y ST, con L1 compuesto por 84 de 114 genomas ST4, y L2, con 47 de 53 genomas ST6.
- Asociación a subespecie: Si bien cada subespecie está representada a lo largo de todo el árbol filogenético, se observa que más del 60 % de los aislamientos del linaje L1 pertenecen a la subespecie Cfv y al biovar intermedio Cfvi, mientras que Cff es predominante en los demás linajes,

con más de 93 % de todos los aislamientos no L1 correspondientes a la subespecie Cff.

4.4.4. Distancias pareadas

Las distancias pareadas de los genomas en relación a la subespecie, el hospedero y el linaje pueden visualizarse de forma general en el *heatmap* de la Figura 4.8 y en la Figura 4.9.





Distribución de distancias intra y entre subespecies y en relación al hospedero

A partir del análisis de distancias de SNPs pareadas se observa que la subespecie Cff presenta el mayor rango, llegando a 2.013 SNPs en promedio (IQR=2.296), seguida por Cfv, con 443 SNPs de separación media (IQR=115). Por el contrario, la diversidad intra Cfvi es significativamente menor, con 115 SNPs de separación media (IQR=154). Las comparaciones inter subespecie muestran que el biovar intermedio se aproxima más a Cfv que a Cff (298 (IQR=104) vs. 1.474 (IQR=2605) SNPs en promedio respectivamente, significativo con prueba de Wilcoxon, p=2,2 × 10⁻¹⁶).

En cuanto a la relación distancia-hospedero, se pueden observar menores distancias entre aislamientos bovinos, con una media de 641 SNPs (IQR=131). Las distancias intra ovinos e intra humanos fueron considerablemente más altas (1.921 y 3.813 SNPs) y más dispersas (IQR: 3.111 y 3.770, respectivamente).

Distribución de distancias y estructura poblacional

En general, la distancia intra-linaje alcanza los 500 SNPs como máximo, con la excepción de L7, donde el genoma 16244_6.2, correspondiente a L7b, está separado de los genomas de L7a por más de 3.000 SNPs. El linaje L4 es el de menor distancia intralinaje, con una media de 280 SNPs.

4.4.5. Recombinación

La identificación de las regiones o bloques recombinantes se basa en las distribuciones de SNPs anómalas, con respecto al resto del genoma. De esta

manera se identificaron 35 bloques recombinantes, representando en promedio el 2,27% de las bases en el genoma *core*. Se estudió la distribución de las regiones con evidencia de recombinación en relación al linaje, encontrándose diferencias significativas en el número de bases recombinantes entre linajes, con el linaje L1 concentrando el mayor número de bases en regiones recombinantes (3,22%, IQR=1,31), seguido de L2 (1,76%, IQR=0) y L8 (1,81%, IQR=0). En la Figura 4.10 se muestra el *boxplot* donde se representa el porcentaje de posiciones recombinantes por genoma para cada linaje. Se encontró además que la mayoría de los bloques recombinantes son exclusivos de linaje, con los linajes L1 y L2 con un mayor número de bloques recombinantes exclusivos, algunos de ellos presentes en todos los genomas de cada linaje. También se encontraron bloques recombinantes compartidos por más de un linaje, los que se muestran como barras negras en la Figura 4.11.





4.4.6. Búsqueda de blancos moleculares para tipificación

Análisis de componentes principales discriminatorios

Un análisis discriminatorio de componentes principales (DAPC) permite conocer en qué medida las variables o combinaciones de variables contribuyen a la diferenciación de grupos preestablecidos, como pueden ser la subespecie o el hospedero del aislamiento, de forma de maximizar la separación entre grupos, al mismo tiempo que se maximiza la semejanza intra-grupo. Los DAPC se hicieron a partir del alineamiento del marco clonal y se utilizaron como variables de agrupamiento linaje, subespecie y hospedero.

- **Agrupamiento por linaje:** Al utilizar el linaje como variable de agrupamiento la separación es evidente (Figura 4.12-A), con grupos definidos tanto a nivel del primer como del segundo componente principal. Los linajes identificados en la filogenia y la clusterización jerárquica obtenida con hierBAPS se distinguen sin solapamiento. En este caso la asignación de los genomas a los linajes correspondientes fue correcta en todos los casos.
- **Agrupamiento por subespecie:** En la Figura 4.12-B se muestra el gráfico obtenido a partir de resultados de DAPC sobre las posiciones polimóficas del marco clonal, utilizando como variable de agrupamiento el resultado de la tipificación bioquímica de cada aislamiento. Los aislamientos con subespecie no determinada fueron excluidos del análisis. Se observa que, si bien existe una separación entre los grupos, hay individuos de diferentes subespecies que se solapan en el gráfico. Según el análisis realizado, sólo el 58 % de los genomas tipificados como Cfvi

se asignan correctamente. Para Cfv y Cff el porcentaje de asignación exitosa es mayor, con 78 y 89% de genomas asignados correctamente.

Agrupamiento por hospedero: En el gráfico resultante del DAPC obtenido para la variable de agrupamiento hospedero, mostrado en la Figura 4.12-C, se observa, al igual que en el caso anterior, que la separación es sólo parcial. Existen puntos solapantes entre categorías de hospederos, con excepción del genoma aislado a partir de un aislamiento de mono, que se agrupó con los aislamientos humanos, pero no se solapó con ningún otro genoma. El grupo hospedero reflejó de forma más acertada las diferencias entre los genomas en el alineamiento del marco clonal, con porcentajes de asignación correcta en el 81 % de los genomas ovinos, y mayores al 90 % en los genomas bovinos y humanos (91,23 % para bovinos y 91,03 % para humanos). La asignación del único genoma de mono se realizó correctamente, sin embargo, dos genomas de aislamientos humanos fueron asignados como pertenecientes a mono.

Pangenoma de la especie

Se partió de un grupo de 210 genomas seleccionados al azar a partir del conjunto de datos original, excluyéndose entonces los genomas 16244_6.17, 17059_2.25, 2740_S96, 32019.46, cffH1UY, y HB74. El pangenoma obtenido está compuesto por 4.901 grupos de genes ortólogos, de los cuales 1.531 (31%), por estar presentes en al menos el 95% de los genomas, pertenecen al grupo de genes *core*. La estimación de los parámetros del modelo de la ley de Heaps demostró que el pangenoma de esta especie está abierto, con $\alpha = 0,814$ [184], estimándose que el tamaño total del pangenoma sería de



6.580 grupos de ortólogos. Estos resultados pueden apreciarse en las curvas de rarefacción de los genes *core* y accesorios de la Figura 4.13(**A**).

Caracterización del pangenoma: Se evaluó el número de genes accesorios por genoma para cada uno de los linajes (Figura 4.13(**B**)). L1 contiene un número de genes accesorios significativamente mayor a los demás linajes (p ajustado entre 0,46 para L1-L7 y 7,5 × 10^{-21} para L1-L2) según resultados del test de Wilcoxon con corrección de Bonferroni. Resulta llamativo el punto



extremo presente en el linaje L6, que corresponde al genoma 32019.24, con un número de genes accesorios mayor a la media de L1.

A partir del pangenoma se construyó la *panmatriz*, que consiste en una matriz donde se representan los grupos de genes ortólogos presentes en cada genoma. A partir de esta se construyó el PCA mostrado en la Figura 4.14(**A**).



parte **B**.

Vemos allí como se separa el linaje L1 del resto de los linajes en el componente principal 1, con lo que se puede inferir que, además de diferir en el número de genes accesorios, el perfil de genes accesorios de este linaje es muy diferente al de los demás. Al eliminar los genomas pertenecientes al linaje L1, se puede visualizar el PCA de los demás linajes con mayor resolución, como se observa en la Figura 4.14(**B**). Vemos allí cómo los linajes se separan, pero la distancia del genoma 32010.24 al resto de los genomas impide visualizarlos con mayor resolución. Al eliminar este genoma del análisis vemos las separaciones entre los grupos con mayor claridad, indicativo de perfiles de genes accesorios diferenciales entre linajes, con el linaje L5 notablemente separado del resto, y los linajes L2 y L3 muy cercanos entre sí, indicando una mayor similaridad entre ellos.

Funciones de los genes accesorios: Se extrajeron y tradujeron secuencias representativas de cada grupo de ortólogos del pangenoma, con el fin de anotarlos con eggNOG-mapper v.2 [173], tomando como referencia la base de datos eggNOG v.5.0. Se obtuvo la anotación de 3.980 de un total de 5.261 grupos de genes ortólogos en el pangenoma, 2.132 de ellos pertenecientes al *subset* de genes accesorios.

A partir de la *panmatriz* y las categorías COG derivadas de la anotación de los grupos de ortólogos, se realizó un análisis de correspondencias múltiples (MCA), obteniéndose los gráficos mostrados en la Figura 4.15. En el gráfico de la parte **A** encontramos una clara separación entre los genomas del linaje L1 del resto, siguiendo el comportamiento observado tanto a nivel del número de genes accesorios como del perfil de presencia/ausencia de genes. Al excluir los genomas del linaje L1 y repetir el MCA se observan dos *clusters*. Por un lado, con valores negativos en el eje x, los linajes L2, L3, L4 y L5, y por el



otro L6, L7 y L8, con valores en x mayores a 0. El linaje L4 se aleja de los demás en el eje y, lo que denotaría una distinción en su perfil funcional.

Grupos de ortólogos linaje específicos: Partiendo de la panmatriz, se seleccionaron los grupos de ortólogos presentes o ausentes en por lo menos

el 90% de los genomas de cada linaje, para luego aplicar otro filtro que permitió seleccionar aquellos grupos cuya presencia o ausencia es exclusiva de linaje. Este procedimiento permitió identificar 194 grupos de ortólogos con perfiles de presencia/ausencia diferenciales. Estos se observan en el *heatmap* de la Figura 4.16.

Para conocer más sobre la función de estos grupos de ortólogos, se revisó su anotación funcional. Así encontramos que la mayoría de los grupos de ortólogos exclusivos de linaje pertenecen a tres categorías funcionales:

- E: Transporte y metabolismo de aminoácidos
- P: Transporte y metabolismo de iones inorgánicos
- U: Tráfico intracelular, secreción y transporte vesicular.

La mayoría de estos grupos exclusivos son, probablemente, genes ganados, ya que su presencia es exclusiva para un linaje.

La Figura 4.17 muestra para cada categoría, el número de grupos de ortólogos cuya presencia o ausencia es exclusiva para un determinado linaje. Como se observa allí, L4 es el linaje que presenta mayor diversidad de categorías funcionales en grupos de ortólogos exclusivos, tanto en cuanto a presencia como ausencia de genes con respecto a los demás linajes, mientras que L2 sólo presenta grupos exclusivos pertenecientes a la categoría COG E, de transporte y metabolismo de aminoácidos. En la Tabla Tabla 8.1 se muestran las definiciones de todas las categorías COG.




Diseño de primers específicos para linajes

Se han desarrollado diferentes metodologías de diagnóstico con el fin de diferenciar a Cff de Cfv y Cfvi, sin embargo, ninguna ha demostrado especificidad y sensibilidad suficiente. Dada la clara estructura poblacional encontrada, nos abocamos hacia la identificación de marcadores moleculares que permitan identificar mediante PCR el linaje de un aislamiento. Para esto, se diseñaron cebadores específicos para los linajes 1 y 2 utilizando RUCS [180], una herramienta que permite extraer kmeros presentes en un grupo de genomas "positivos" y ausentes en un grupo de genomas "negativos", a partir de los cuales se realiza un ensamblaje que resulta en regiones *core*, exclusivas del grupo de "positivos". A partir de estas regiones, se diseñan cebadores utilizando Primer3 [181]. Los cebadores obtenidos fueron validados mediante PCR *in silico* utilizando FastPCR [182], resultando en 100 % de especificidad y de sensibilidad. Se mandaron a sintetizar tres pares de cebadores para cada linaje a identificar, los cuales se detallan en la Tabla 4.4.

Los cebadores diseñados para identificar L1 con longitudes de entre 26 y 29 bases y temperaturas de *melting* (Tm) entre 53,7 y 54,7 °C, permiten amplificar una región de 277 a 286 bases en el gen que codifica una resolvasa, según el análisis BLASTN realizado a partir de los amplicones predichos. Una búsqueda BLASTP sobre la base de datos de proteínas no redundantes del NCBI da como mejor *hit* la secuencia identificada como WP_002850926, una proteína de la familia de trasposasas tipo IS607, presente en el elemento de inserción ISCfe1. Al combinar el resultado de la búsqueda mediante BLASTN del amplicón predicho con un BLASTN para identificar las secuencias de inserción ISCfe1 encontramos que en todos los casos, el amplicón está localizado entre las posiciones 40 y 325 del primer CDS del elemento de inserción ISCfe1. La Figura 4.18 (A) muestra los sitios donde se localizan las secuencias de inserción ISCfe1

Par	Secuencia	Long. (pb)	Tm (°C)	GC (%)	Long. amp. (pb)	Hits Swissprot	
1	F: ACTATTCAAACACTTAGAAACTGGGATAA	29 26	54,4 53.7	31.0 32.2	277	Resolvasa putativa R80;	
L1_2	F: TTACTATTCAAACACTTAGAAACTGGGAT	29 26	54,4	31,0	279	Resolvasa putativa R771; Óxido nitroso reductasa;	
L1_3	F: CITIGGGGTTACCIACCIAAGICI IGI	20 27 26	53,7 54,7	34,0 37,0	286	Sistema de eflujo regulado por glutatión	
1	F: GCTTCCACCACTAAAAGAGCAA	20	55,2	45,5	415		
	R: GCTGGACTGCAAATTCCATCA	21	55,4	47,6	417		
L2_2	F: GCTTCCACCACTAAAAGAGCAAG R: GCTGGACTGCAAATTCCATCA	23 21	56,3 55,4	47,8 47,6	417	Enzima de restricción de tipo 1 MjaXIP	
L2_3	F: GCTTCCACCACTAAAAGAGCAA R: GCTGGACTGCAAATTCCATCAT	22 22	55,2 55,6	45,5 45,5	417		

Tabla 4.4: Cebadores diseñados a partir de los resultados del análisis con RUCS, para la identificación de los linajes L1 y L2. Long.: longitud del cebador; Long. Amp.: longitud del amplicón esperado

y los amplicones predichos, tomando como referencia el cromosoma del genoma cfvi03293 (CP0069999.1).

Los cebadores diseñados para la identificación de L2 consisten en secuencias de 21 a 23 bases, con Tm estimado de entre 55,2 y 56,3 °C. Los tres pares diseñados permiten amplificar una región de 417 bases de un CDS. Según los resultados de BLASTP a partir de su secuencia aminoacídica contra la base de datos de proteínas no redundantes del NCBI, el amplicón se encuentra en un gen que codifica la subunidad S de una enzima de restricción del tipo 1. En la Figura 4.18 (**B**) se muestra el amplicón predicho y su entorno génico.

Validación de los cebadores diseñados: Con el fin de validar los cebadores diseñados, se realizaron amplificaciones mediante PCR, utilizando los pares compuestos por los cebadores L1_3F L1_1R y L2_1R L2_2F, que generan productos de amplificación específicos de 286 pb y 417 pb en aislamientos de los linajes L1 y L2 respectivamente. En primer lugar se realizó una PCR en gradiente de temperaturas, con el fin de seleccionar las mejores temperaturas de *annealing*, utilizando como molde el ADN genómico de la cepa cfv06/195 como control positivo del par L1 y la cepa Cff70L como control positivo para L2. Como control negativo de ambos cebadores, se utilizó la cepa 17144, perteneciente al linaje L7. Para L1 se probaron temperaturas de *annealing* de 52 y 53°C, mientras que para L2 las temperaturas ensayadas fueron de 54 y 55°C. Los productos de amplificación fueron separados mediante electroforesis en un gel de agarosa 1 % en buffer TAE, obteniéndose, en ambos casos, bandas correspondientes a productos de amplificación del tamaño esperado. Se seleccionaron las temperaturas de *annealing* de 52°C para L1 y 54°C para L2, en base a que las bandas correspondientes a los productos de amplificación a esas temperaturas fueron más definidas.

Se utilizaron 26 muestras de ADN genómico de aislamientos secuenciados (Ver Tabla 4.5a) y otras 27 de aislamientos no secuenciados (Tabla 4.5b) como molde para reacciones de PCR con cebadores específicos de L1, obteniéndose producto de amplificación específico en las 23 muestras de ADN de aislamientos secuenciados pertenecientes al linaje L1, y en 15 de los 27 aislamientos no secuenciados.

En general, se comprobó la especificidad de la amplificación del linaje L1, con excepción de la cepa 04/554, perteneciente al linaje L8 según la estructura poblacional descrita en este trabajo, de la cual se obtuvo el producto de PCR específico de L1, y sería el único falso positivo obtenido. La cepa 17144 fue la única secuenciada no perteneciente a L1 que se testeó con los cebadores L1, y resultó en ausencia de producto de amplificación. Los cebadores diseñados para la amplificación de L2 fueron testeados únicamente en las cepas 06/195 (L1), 70L (L2) y 17144 (L7), dando como resultado la producción de un producto de amplificación específico únicamente en la cepa 70L.



Fig. 4.18: (A) amplicones predichos a partir del par de cebadores L1_1, contra el cromosoma del genoma cfvi003293 (CP006999.1). En esta secuencia se esperan la amplificación de cuatro regiones del cromosoma, que se corresponden con el primer CDS de la secuencia de inserción ISCfe1 (marrón). (B) Para el genoma cff8240 se espera, a partir del par de cebadores L2_1, la amplificación de un único fragmento que se localiza en un CDS que codifica la subunidad S de una endonucleasa de restricción. En ambas figuras los amplicones predichos se muestran en naranja y los ORFs cercanos se muestran en azul. En celeste se muestran los ORFs interrumpidos por la secuencia de inserción ISCfe1.

(a) Aislamientos secuenciados	
-------------------------------	--

(b) No secuenciados

Сера	Genoma	Linaje	PCR L1	Сера	PCR L1
00/398	17059_2.17	1	+	73790	-
00/564	17059_2.18	1	+	A28	-
03/596	17059_2.22	1	+	71009	v
04/875	17059_2.23	1	+	11677	+
05/434	17059_2.25	1	+	Cff063	+
06/195	17059_2.27	1	+	MCR03	+
06/340	17059_2.26	1	+	06/341	+
07/379	17059_2.28	1	+	MCR01	-
10/247	17059_2.31	1	+	36	-
10/445	17059_2.32	1	+	650 H-	+
11/360	17059_2.33	1	+	11408	+
11/427	17059_2.34	1	+	cfv 1198	+
2740	2740_S96	1	+	38877	-
90/189	cff90189	1	+	F106	-
90/264	cfv90264	1	+	2432	+
95258	95258_S94	1	+	2106	+
97/608	cfvG	1	+	11356 e1	+
99/541	cfvi99541	1	+	2991	+
99/801	17059_2.16	1	+	71098	-
H1	H1_S84	1	+	4601	-
M8	M8_S86	1	+	CFV 3598	-
R18	R18_S87	1	+	305	+
U10	U10_S88	1	+	3596	+
BT_10/98	1273268.3	4	-	2370P	-
17144	17144_S83	7	-	90264	+
04/554	cff04554	8	+	CIN3	+
				70L	-

Tabla 4.5: Resultados de validación de los cebadores para la detección mediantePCR de aislamientos pertenecientes al linaje L1 a partir de ADN genómico de aislamientos de *C. fetus* secuenciados y no secuenciados.

4.5. Discusión

La OMSA recomienda la realización de pruebas bioquímicas para el diagnóstico de la CGB, sin embargo, no existen evidencias a nivel genómico que avalen la asociación entre los fenotipos diferenciales y las características ecológicas y la patogenia atribuida a las subespecies de *C. fetus*. Varios estudios han investigado la estructura poblacional de *C. fetus*, impulsados por la necesidad de desarrollar métodos de diagnóstico molecular sensibles y específicos para la CGB, basados en la identificación de *C. fetus subsp. venerealis* y del biovar intermedio. Este trabajo se suma a esta línea de investigación, concebido como un estudio exhaustivo a gran escala, que permita capturar la mayor diversidad posible en aislamientos de *C. fetus* de mamíferos, información imprescindible para el diseño de metodologías moleculares para el diagnóstico de la CGB.

Más de 200 genomas de buena calidad componen un conjunto de datos de alcance global, pero con desbalances

Partimos de un conjunto de secuencias genómicas obtenidas en distintos países en cinco continentes. Sin embargo, la globalidad de este conjunto de datos se ve comprometida debido al desbalance en el origen de los genomas, donde países europeos junto con Argentina y Uruguay aportan gran parte de los mismos. Estos países, junto a Brasil, se encuentran entre los mayores productores y proveedores de ganado bovino y sus productos derivados. Se estima que Brasil es un país con alta prevalencia de CGB, atribuyendo a esta enfermedad grandes pérdidas económicas que han impactado significativamente en el sector ganadero [185]. Sin embargo, se pudo acceder a sólo 4 secuencias genómicas de *C. fetus* de origen brasileño.

En este conjunto de datos también se encuentra un desbalance en cuanto al tipo de hospedero, y la subespecie, con un mayor número de genomas de aislamientos bovinos, que a su vez han sido en su mayoría tipificados como Cfv o Cfvi. Esto es esperable, dada la importancia de la CGB en bovinos, atribuida a esta subespecie. En humanos y ovinos se encuentra predominantemente Cff, y persiste el subdiagnóstico.

Un paso previo a la realización de los análisis de los genomas de *C. fetus* que forman parte de este estudio fue comprobar su calidad. Para esto se utilizaron como parámetros de calidad el número de *contigs*, la completitud, basada en la presencia de un *set* de genes *housekeeping*, y la contaminación, basada en la presencia de múltiples copias de genes descritos como de copia única (Detalles en el Párrafo 3.5). Si bien la mayor parte de los genomas estudiados se encuentra fragmentado, sólo en dos genomas se detectaron niveles bajos de completitud, razón por la cual fueron excluidos de los análisis posteriores.

Los estudios poblacionales revelan una estructura definida y robusta, confirmada por diferentes metodologías

La estructura poblacional *de C. fetus* ha sido estudiada previamente. En la revisión "Pathogenomics of Emerging *Campylobacter* Species" [58] se citan los principales trabajos en los que se analiza la estructura poblacional de *C. fetus*.

En este trabajo se obtiene la estructura poblacional a partir de 216 genomas de *C. fetus* de diferentes orígenes, mediante un abordaje múltiple, de forma similar a los trabajos de Van et al. [26], Iraola et al. [154] y Abdel-Glil et al. [186], obteniéndose resultados comparables. Esta se contrastó con resultados

de tipificación mediante MLST y en relación a la subespecies y el origen de los aislamientos.

La estructura poblacional inferida consiste en 8 linajes, y fue obtenida mediante el análisis de clusterización jerárquica y métodos multidimensionales basados tanto en el alineamiento del marco clonal, como en los perfiles de presencia/ausencia y funcionales del pangenoma. El linaje L7 fue el único que presentó inconsistencias con la filogenia, probablemente producidas por el fenómeno de acercamiento de ramas largas como se mencionó anteriormente (ver Subsubsección 4.4.3), razón por la cual se propone su división en los linajes L7a y L7b.

La tipificación mediante MLST demostró que cada linaje está asociado a un ST, con excepción de los linajes L1 y L2, en los que se encontraron varios ST, pero uno es predominante. El genotipo ST4 predomina en L1, pero ha sido encontrado también en aislamientos clínicos pertenecientes al linaje L2. En L2, en cambio, predomina el ST6, y existe un subgrupo de genomas ST4. Un caso de homoplasia de este tipo fue reportado anteriormente [153] y evidencia que el MLST por sí mismo es inadecuado para el genotipado de *C*. *fetus*, ya que no refleja las diferencias encontradas a nivel genómico.

Existe asociación entre linajes y hospederos, donde L1 se asocia a bovinos. Esta asociación es evidente al observar la gran proporción de bovinos en este linaje, y la baja ocurrencia de bovinos en otros linajes. El análisis de las distancias pareadas intra bovinos demuestra en general una muy baja variabilidad en este hospedero. Otra asociación menos evidente puede trazarse entre el linaje L4 y ovinos, ya que la mayoría de los genomas de ese grupo provienen de este hospedero. Sin embargo, al analizar la distribución

de las distancias pareadas, se observa un rango de valores muy grande, sugiriendo la distribución de genomas ovinos en diferentes linajes.

La asociación de la estructura poblacional y las subespecies no parece tan clara, sin embargo se encontró baja variabilidad genómica en Cfv y Cfvi, lo que sugiere que estos genomas no se distribuyen entre los linajes, como sí lo hace Cff. Sin embargo, el análisis de componentes principales discriminatorios no separa completamente a las subespecies, debido a la presencia de perfiles genómicos solapantes.

Los diferentes linajes tienen perfiles de genes accesorios y funciones diferenciales

Otra arista explorada en este trabajo fue el pangenoma, que se encuentra abierto, y está compuesto por 4.900 grupos de genes ortólogos, con un 31 % de ellos presentes en al menos el 95 % de los individuos, formando parte del conjunto de genes *core*. La apertura del pangenoma indica que la diversidad de esta especie aún no ha sido completamente explorada, por lo cual quedarían aún genes o grupos de ortólogos por descubrir. Las poblaciones con pangenomas abiertos también suelen coincidir con una mayor flexibilidad metabólica y la ocupación de múltiples nichos [187], sin embargo este probablemente no es el caso, debido al bajo número de genomas secuenciados disponibles y a la baja diversidad reportada para esta especie.

Se observaron diferencias significativas en el número de genes accesorios entre linajes, observándose una explosión de genes en el linaje L1. Esto se suma al aumento en la proporción de bases en regiones recombinantes encontrado en este linaje, denotando una mayor plasticidad en estos genomas. El tamaño aumentado del genoma accesorio del linaje bovino con respecto a los linajes no bovinos es discutido por Iraola et al. [154], quienes encontraron entre estos a genes pertenecientes a sistemas de secreción de tipo 4 (T4SS) y elementos movilizables, que jugarían un rol muy importante en el proceso de adaptación a hospedero. Graaf-Van Bloois et al. reportan en genomas de este linaje, correspondiente al clado 5 en su análisis, una mayor proporción de genes bajo selección diversificadora y la presencia de 24 SNPs específicos en el gen *cjeI*, perteneciente a un sistema de restricción/modificación de tipo II, cuya deleción se asocia al aumento en la capacidad de transformación de cepas de *C. jejuni*, por lo que los genomas de este linaje estarían más afectados por la transferencia horizontal de genes (THG)[189].

Tanto los perfiles de presencia/ausencia de grupos de ortólogos como sus anotaciones funcionales permitieron replicar los grupos predichos por hierBAPS en los análisis multidimensionales aplicados. En ambos, además, es notoria la separación del linaje L1, lo cual denota una gran divergencia con respecto a los demás linajes, tanto en cuanto al número y perfil de genes accesorios como en lo funcional. El linaje L4, en el que predominan genomas aislados de ovinos, tiene la mayor divergencia, después de L1, en los análisis multidimensionales, con perfiles funcionales y de presencia/ausencia de genes divergentes al resto de los linajes, como grupos de ortólogos exclusivos y funcionalmente más variables.

El análisis de los genes accesorios reveló grupos de ortólogos cuya presencia o ausencia es específica de linaje. Su anotación permitió reconocer que la mayoría se trata de genes relacionados al transporte de aminoácidos e iones inorgánicos y sistemas de secreción, categorías principalmente ligadas a la interacción con el hospedero.

Se diseñó una metodología para el diagnóstico de *C. fetus* L1 basada en PCR, con alta especificidad y sensibilidad *in silico* y resultados prometedores *in vitro*

Para cumplir el objetivo de obtener un método de diagnóstico para la CGB basado en la estructura poblacional obtenida, y asumiendo que L1 representa a la subespecie Cfv, se utilizó el programa RUCS, que identifica regiones presentes en un grupo de genomas y ausentes en otro, y diseña a partir de ellos cebadores capaces de amplificarlos de manera específica. Así se obtuvieron cebadores capaces de amplificar de manera específica y diferencial un fragmento en genomas de L1, que luego se confirmó que forma parte del primer CDS en el elemento de inserción ISCfe1. La validación de la técnica diseñada resultó en la identificación exitosa de genomas L1, con 100% de sensibilidad y 100% especificidad, al menos en los análisis computacionales realizados. Aunque la validación in vitro se realizó en forma parcial, los resultados obtenidos demostraron la simplicidad de la interpretación de los resultados, con amplificación de un único producto específico en muestras de ADN genómico de aislamientos L1, siendo estos muy prometedores para ser aplicados en el diagnóstico de rutina. Para que esto sea posible es necesario extender los ensayos de validación abarcando cepas más diversas y también comprobar su efectividad directamente a partir de colonias.

El elemento de inserción ISCfe1 fue descrito por primera vez por Abril et al. en 2007 [99], como un elemento exclusivo de Cfv. En trabajos posteriores, Abdel-Glil et al. [186], Graaf-Van Bloois et al. [85] y McGoldrick et al. [190] analizan la variabilidad genética del elemento y diseñan cebadores capaces de amplificar diferentes fragmentos del mismo. En la Figura 4.19 se muestran los sitios amplificados por los cebadores diseñados en este trabajo (ISCfe1_3F2R) y en otros. En un trabajo reciente de Abdel-Glil et al. [186] se evalúan *in silico* las diferentes métodologías basadas en PCR, utilizando para esto un *set* de 283 secuencias genómicas, donde demostraron que las metodologías diseñadas por McGoldrick et al. y Graaf-Van Bloois et al. tienen alta sensibilidad y especificidad, con lo que, al tener como blanco el mismo *locus*, podría extrapolarse a partir de éstos el potencial de los cebadores diseñados en este trabajo.

Si bien la metodología de diagnóstico propuesta se basa en cebadores ubicados dentro de un elemento de inserción, la detección de ISCfe1 en todos los genomas del linaje L1 indica que es estable. Sin embargo, la reactividad cruzada con la especie *C. hyointestinalis*, frecuentemente encontrada en bovinos, comprometen esta diana de diagnóstico, por lo que deberá formularse como una PCR multiplex, adicionando marcadores específicos de *C. fetus* [26].



Fig. 4.19: Mapa de una región del genoma de la cepa cfv84-112, correspondiente a una de las tres inserciones de ISCfe1, donde se muestran los fragmentos amplificados por cebadores diseñados por Abril et al. (Abril 2007), Graaf-Van Bloois et al. (VDG6 y VDG7), y McGoldrick et al. (McGoldrick), y los diseñados en estre tabajo (ISCfe1 3R2F)

4.6. Conclusiones

Los genomas de *C. fetus* asociados a mamíferos son altamente clonales, y las subespecies no se distinguen mediante técnicas como la hibridación ADN-ADN o secuenciación del gen ARNr 16S, por lo que la clasificación de Cff y Cfv como diferentes subespecies es cuestionable [191].

En este trabajo identificamos un linaje predominantemente bovino y enriquecido en cepas caracterizadas como Cfv del ST4. Este linaje se caracteriza además por un mayor número de genes accesorios, perfiles funcionales distintivos y una mayor proporción de sus genomas sujetas a recombinación. A pesar de esto, este linaje es considerablemente menos diverso que los demás, los que puede estar indicando un proceso de adaptación a hospedero bovino.

Es entonces que se sugiere el diseño de una metodología de diagnóstico basada, no en la identificación de subespecies, sino en función a la estructura poblacional descrita aquí y por otros autores [189, 154] para el diagnóstico de la CGB, basada en la identificación del linaje L1. En este sentido, los cebadores diseñados podrían ser una buena opción de diagnóstico basado en PCR, para lo cual se requiere una validación más exhaustiva.

5

Genómica poblacional aplicada al estudio de aislamientos de *C. fetus* bovinos

5.1. Introducción

La estructura poblacional encontrada en *C. fetus*, estudiada en el capítulo anterior, revela la presencia de 8 linajes. El linaje L1 presenta una fuerte asociación a bovinos y características genómicas que sugieren un proceso de adaptación que ha sido reportado anteriormente [188, 154]. Los hospederos representan combinaciones complejas de presiones selectivas sobre los genomas bacterianos, dadas por las respuestas inmunes, fluctuaciones en la concentración de metabolitos y nutrientes y la presencia de otros microorganismos [58].

Se han descrito diferentes factores involucrados en la adaptación de *C. fetus* al hospedero. A continuación se resumen algunos de ellos.

Factores de virulencia

Según la definición expuesta en la página web de la base de datos de factores de virulencia (VFDB por sus siglas en inglés), los factores de virulencia se entienden como:

Características, incluidos productos génicos, que le permiten a un microorganismo establecerse en o dentro de un hospedero particular y aumentar su capacidad de causar enfermedad. Entre estos se encuentran toxinas, proteínas y carbohidratos de superficie que protegen a la bacteria o median procesos que contribuyen a su patogenicidad.

Se han propuesto diferentes factores de virulencia en *Campylobacter*, entre los que se encuentran aquellos involucrados en la movilidad mediada por flagelos, la adherencia a la mucosa intestinal y la producción de toxinas [192]. Es frecuente encontrar en los genomas de *C. fetus* la mayoría de los factores de virulencia más frecuentes [193].

A continuación se detallan los factores de virulencia abordados en este capítulo.

CDT La toxina distensora citoletal, más conocida como CDT por sus siglas en inglés, comprende tres subunidades, A, B y C. Estos genes forman parte de un operón y forman una exotoxina que, al ingresar a una célula hospedera, produce daños en el ADN que comprometen el ciclo celular [194]. La subunidad B es la responsable de la actividad nucleasa, mientras que las subunidades A y C median la unión y la entrada a las células blanco. Se ha demostrado que son necesarias las tres subunidades para la actividad citotóxica [192]. CDT ha sido encontrada en diferentes especies bacterianas entéricas y no entéricas, incluyendo *Escherichia coli, Shigella spp., Haemophilus ducreyi* y *Actinobacillus actinomycetemcomitans* [192].

SAP Este término se refiere a proteínas de matriz de superficie, que conforman una estructura similar a una cápsula. Esta estructura, denominada

capa S, forma una cubierta proteica paracristalina en células de *C. fetus*, y es responsable de la evasión del sistema inmunológico del hospedero. Las proteínas de superficie están codificadas por el clústeres de genes *sapCDEF* y múltiples copias de *sapA* o *sapB* o recombinantes de estas variantes (*sapAB*) [195]. *C. fetus* contiene varias copias de *sapA* o *sapB*, cuya expresión varía en función de la posición relativa de una única secuencia promotora. En un momento dado una bacteria puede expresar un alelo de *sap* y, mediante inversiones en el *locus sap* que ocurren por recombinación dependiente o independiente de *recA*, puede modificar la posición relativa del promotor sap cambiando así el alelo expresado y por lo tanto la conformación de la capa S. Estos cambios le permiten a *C. fetus* evadir el sistema inmune y persistir en el hospedero, razón por la cuál es uno de los principales factores de virulencia.

Sistemas de secreción Los sistemas de secreción son maquinarias celulares que permiten la externalización de moléculas efectoras fuera del citoplasma o directamente hacia el citoplasma de otra célula. Estos se dividen en diferentes tipos, dependiendo del tipo de molécula efectora que secreten y la estructura del sistema. La Figura 5.1 muestra la estructura y composición de los sistemas de secreción bacterianos.

T1SS Los sistemas de secreción de tipo 1 transportan su sustrato a través de las membranas interna y externa en un único paso, y se asemejan a los sistemas de transporte de tipo *ATP binding cassette* o ABC [196]. En *C. fetus* un sistema de secreción de este tipo es el encargado de externalizar las proteínas que forman parte de la capa S [197].

T2SS Los sistemas de tipo 2 están conservados en la mayoría de las bacterias gram negativas, y son responsables del trasporte de proteínas plegadas desde el periplasma a través de la membrana externa. La salida hacia el periplasma de las proteínas plegadas es dependiente de un sistema de transporte del tipo Sec o Tat. Los T2SS son responsables de la secreción de diferentes sustratos, algunos de los cuales están involucrados en la patogénesis. Elementos de este tipo de sistemas son necesarios para la transformación natural en *C. jejuni* [198]. Estos sistemas forman parte de un grupo conocido como TFF (*type four filament*), que contiene además los sistemas T4P (*type four pili*), involucrados en la adhesión a células del hospedero, movilidad tipo *twiching* y en la formación de biofilms [199].

T3SS Estos sistemas de secreción forman una estructura compleja que se asemeja a una jeringa y una aguja, y tienen como función la inyección de diferentes efectores proteicos directamente desde el citoplasma de la bacteria hacia el interior de una célula eucariota del hospedero. Estos sistemas evolucionaron a partir de los sistemas responsables de la externalización de los flagelos bacterianos [199].

T4SS Los sistemas de secreción de tipo 4 son especialmente importantes, debido a su capacidad de externalizar tanto proteínas como ADN y su rol en la transferencia horizontal de genes (THG). Todos los T4SS están relacionados evolutivamente, y, por lo tanto, sus componentes y su funcionamiento son similares. El T4SS más estudiado es el de *Agrobacterium tumefaciens*, encargado de movilizar ADN oncogénico a células vegetales. Este sistema está compuesto por 12 proteínas VirB y VirD4. Algunas proteínas VirB forman parte del canal secretor, mientras que las proteínas VirB4, VirB11 y VirD4 son ATPasas que brindan la energía necesaria para el proceso de transporte [196]. Los sistemas de secreción de tipo 4 son importantes factores de virulencia en



muchas bacterias, debido a que posibilitan la adquisición de nuevos genes de virulencia y resistencia mediante THG, Estos sistemas permiten inyectar proteínas efectoras directamente sobre células del hospedero, para alterar su ciclo celular o comprometer sus estructuras [196]. T5SS Estos sistemas, a diferencia de los anteriores, que forman estructuras que permiten el transporte de moléculas efectoras a través de las membranas bacterianas, funcionan como autotransportadores de membrana externa. Las proteínas a exportar forman el sistema de secreción que permitirá su externalización, debido a que contienen una región que forma una estructura del tipo barril β que se inserta en la membrana externa formando un canal a través del cual pasa el resto de la proteína. Debido a que estas estructuras sólo se forman en la membrana externa, su transporte hacia el periplasma es dependiente de Sec [196]. En *C. ureolyticus* la presencia de sistemas de este tipo está relacionada con una mayor invasividad [200].

T6SS Los sistemas de tipo 6 están involucrados en el transporte de moléculas efectoras directamente a células del hospedero, e incluso a otras bacterias, por lo que se sugiere un rol en la comunicación y la interacción con el ambiente [196]. Se ha demostrado que proteínas dedicadas a la destrucción de otras células son transportadas por medio de estos sistemas. Esto se logra mediante la degradación de paredes celulares de bacterias o las membranas de hospederos eucariotas, nucleasas, o como parte de un sistema de defensa anti fagos en la que se induce la muerte de células infectadas [201]. En *Campylobacter* se ha demostrado la presencia de estos sistemas a cepas más invasivas y virulentas [202].

Fic Las proteínas Fic obtienen su nombre de "filamentación inducida por AMPc", debido al efecto que ejercen sobre las células al alterar la expresión de proteínas a nivel post-traduccional, mediante la adición de monofosfatos de nucleósidos. Estas proteínas actúan tanto en la propia célula bacteriana, por ejemplo frente a condiciones de *stress*, como en células del hospedero, u otras bacterias, al ser secretadas mediante sistemas de T3SS o T4SS [203]. *C. fetus*

codifica en su genoma varias proteínas Fic que actúan como sistemas toxinaantitoxina (T-AT) e interrumpen la maquinaria de traducción. Se distinguen diferentes clases de proteínas Fic, en función de la ubicación de una hélice inhibitoria que interfiere con la acción del dominio Fic: en las proteínas Fic de clase I esta hélice se ubica en una proteína accesoria, mientras que en las proteínas de clase II y III el dominio Fic y la hélice inhibitoria forman parte de la misma proteína, ubicándose hacia el extremo N o C terminal con respecto a Fic, respectivamente [204].

Moviloma

Los elementos genéticos móviles (EGM) presentes en un organismo son agrupados bajo el término moviloma. Se entiende por EGM a toda secuencia de ADN que puede movilizarse a través del cromosoma o bien trasladarse de un cromosoma a otro. Los EGM pueden transmitirse tanto de forma vertical, mediante replicación, o bien horizontal, a través de un grupo de procesos que se conocen en forma conjunta como transferencia horizontal de genes (THG). Entre los EGM se encuentran secuencias como plásmidos, elementos de inserción, islas génicas, trasposones y profagos. Otros actores importantes que actúan en la composición y el tamaño del moviloma son los sistemas de defensa bacterianos, incluyendo los sistemas CRISPR-Cas y los de restricciónmodificación (R-M), que limitan la incorporación de elementos foráneos a los genomas.

THG La THG constituye uno de los principales mecanismos de evolución, ya que a través de ésta, las bacterias pueden adquirir nuevas funciones que pueden ser útiles en un determinado ambiente, contribuyendo entonces a su *fitness*. La THG puede ocurrir mediante conjugación, a través de T4SS especializados, transfección, a través de fagos que interactúan con receptores

específicos en las superficies bacterianas inyectando al interior celular el material genético, o bien mediante transformación, mecanismo mediante el cual el material genético es internalizado directamente desde el medio extracelular, sin la intervención de terceros actores. Las especies termotolerantes de *Campylobacter*, como *C. jejuni*, *C. coli* y *C. lari* son naturalmente competentes, ya que son capaces de internalizar material genético directamente desde el medio extracelular a través de T2SS y T4P. Asimismo, se asume que la transfección es un proceso frecuente, debido a la presencia de fagos en diferentes ambientes [205].

Las secuencias de inserción (IS) son elementos genéticos móviles que pueden copiarse y movilizarse de una región del genoma a otra. Un ejemplo de este tipo de elementos en *C. fetus* es el elemento de inserción ISCfe1, presente en los genomas del linaje L1, según se discutió en el capítulo anterior, el cual se compone de dos ORFs que codifican trasposasas.

Los elementos de inserción que además pueden ser transferidos a otras células mediante conjugación son los llamados elementos integrativos conjugativos (ICE). Estos contienen todos los genes necesarios para su escisión e integración a un genoma, así como también de la maquinaria de conjugación, la cual es conformada por un T4SS [206].

CRISPR-Cas Para limitar la incorporación desmedida de elementos genéticos foráneos, las arqueas y algunos filos bacterianos han desarrollado un sistema inmune adaptativo, capaz de memorizar segmentos del material genético de virus y otros elementos móviles foráneos, con el fin de neutralizarlos en futuras intromisiones. Un sistema CRISPR (*Clustered regularly interspaced short palindromic repeats*) - Cas (CRISPR - *associated*), está formado por una sucesión de secuencias repetitivas intercaladas con secuencias provenientes de ADN foráneo (espaciadores) encontrados en la cercanía de



genes Cas [207]. La estructura básica de un *locus* CRISPR-Cas se muestra en la Figura 5.2. El número y composición del operón de genes *cas* y la arquitectura del *array* CRISPR son utilizados para la clasificación de estos sistemas, permitiendo su distinción en tipos y subtipos [208]. La identificación de sistemas CRISPR-Cas se basa en la identificación de los genes *cas* por homología de secuencia, y del *array* CRISPR, cuya detección requiere de diferentes estrategias. CRISPRCasFinder, por ejemplo, basa su estrategia en la identificación de los repetidos directos, que contienen secuencias muy similares entre sí. Así es que se identifican elementos repetitivos en los genomas que podrían constituir un CRISPR *array*, y se evalúan los candidatos según su longitud y similitud, y las distancias que los separan.

5.2. Hipótesis y Objetivos

5.2.1. Hipótesis

C. fetus presenta una estructura poblacional compleja. Su estudio permitirá reconocer sublinajes asociados a bovinos producidos por procesos evolutivos divergentes.

5.2.2. Objetivos

Objetivo general

Describir el linaje bovino y su diversidad, con énfasis en los factores involucrados en la asociación al hospedero y su distribución en el marco de su estructura poblacional.

Objetivos específicos

- 1. Obtener la estructura poblacional del linaje L1 bovino.
- Caracterizar los sublinajes obtenidos, describiendo las principales diferencias y similitudes, en cuanto a su repertorio de genes accesorio, elementos móviles, sus perfiles de recombinación y la presencia y distribución de factores de virulencia.

5.3. Metodología

5.3.1. Set de datos

Para la construcción del *dataset* a partir del cual se realizaron los análisis expuestos en este capítulo, se partió del *dataset* descripto en el Capítulo 4, el cual fue filtrado según los resultados obtenidos en la clusterización obtenida mediante hierBAPS y la tipificación mediante MLST, con el fin de retener los genomas pertenecientes a aislamientos bovinos ST4 del linaje L1.

5.3.2. Obtención del alineamiento del marco clonal y estructura poblacional

A partir de los genomas seleccionados según los criterios mencionados anteriormente, se realizó el alineamiento del genoma *core* tal como se describió en Subsección 4.3.5 para el *set* de datos completo. En este caso se utilizó como grupo externo el genoma GCA_000015085.1, perteneciente a la cepa cff82-40 del linaje L2.

Con el fin de describir la estructura poblacional intra-L1, se siguió la misma estrategia empleada anteriormente, en la que se utilizó el método bayesiano de clusterización jerárquica implementado en hierBAPS, además de realizarse PCA y DAPC, a partir de los sitios polimórficos derivados del alineamiento del marco clonal.

Se obtuvo el árbol filogenético del linaje L1 a partir del alineamiento de las posiciones polimórficas del marco clonal en RaxML utilizando el método de ML sobre el modelo GTR.

Filogeografía

A partir de la filogenia obtenida se realizó un análisis bayesiano de significancia de *tips* utilizando BaTS [209], utilizando como estados los orígenes geográficos de los aislamientos, con el fin de demostrar la correspondencia entre linajes y orígenes geográficos de los aislamientos.

5.3.3. Determinación de la antigüedad de los linajes

Diversidad nucleotídica

Se determinó a partir del alineamiento del marco clonal y utilizando la función pegas::nucdiv en R. Se determinó la significancia de los resultados utilizando un test estadístico t de Welch, utilizando la función stats::t. Se visualizaron los resultados utilizando ggplot2.

Datado de la filogenia

Para obtener una filogenia datada, se partió de un alineamiento de genes *core* neutrales concatenados, obtenidos a partir de un *set* de genomas extendido, según se describe a continuación:

Construcción del alineamiento

Set de datos Para determinar las edades de los nodos ancestrales de cada linaje, se partió de un *set* de datos extendido de 151 genomas en total, que contiene, además de los genomas de los aislamientos bovinos ST4, otros genomas representativos de diferentes hospederos y linajes, de los cuales se conoce la fecha de aislamiento y el origen geográfico. La lista de los 151 genomas utilizados se muestra en la Tabla 8.3.

Selección de genes A partir de los genomas anotados se reconstruyó el pangenoma utilizando Roary v.3.11.2 [210], con 95% de identidad. Se realizó el alineamiento de cada grupo de ortólogos utilizando la función DECIPHER::AlignTranslation() [211] en R. Se seleccionaron los grupos de genes *core*, presentes en todos los genomas, en los que no se encontraran evidencias de recombinación, según resultados obtenidos de la aplicación de PhiTest del paquete PhiPack v.1.0 [212]. En una segundo paso de filtrado, se eliminaron los genes en los que se encontrara evidencia de selección, al evaluar el estadístico *D* de Tajima, conservando los genes con $-0, 2 \le D \le 0, 2$. Finalmente, los alineamientos de los genes seleccionados se concatenaron para formar el alineamiento que será el *input* de pasos posteriores.

Análisis bayesianos Los análisis bayesianos realizados en BEAST v.2.6.3 [213] estimaron la topología, los parámetros del modelo de sustitución y las fechas de los eventos cladogénicos:

- Selección del modelo bayesiano Para obtener el mejor árbol se realizó la selección del modelo bayesiano mediante *nested sampling* (NS), utilizando el paquete NS v.1.1.0 [214], siguiendo el protocolo descrito en [215]. Se consideraron cuatro opciones de árboles *prior*: tamaño de población constante, crecimiento exponencial, *Bayesian Skyline* (BS) y BS extendido.
- Selección del modelo de sustitución nucleotídica El mejor modelo de sustitución nucleotídica se seleccionó según resultados de ModelFinder

[216] e IQ-TREE [217]. Se usaron reloj estricto y calibración del datado de *tips* para todas las corridas.

Procesamiento post-corrida Se analizaron los resultados de tres corridas independientes, con cadenas de MCMC de 400 millones de generaciones sampleadas cada 20.000 generaciones. Se chequeó la convergencia a valores estables utilizando Tracer v.1.7.1 [218], obteniendo un tamaño de muestra efectivo (ESS) mayor a 200 para todos los parámetros. Luego de aplicar un *burn-in* de 10%, se combinaron los resultados de las tres corridas utilizando LogCombiner [213]. Se compilaron los árboles para obtener el de mayor credibilidad de clados (maximum clade credibility, MCC) con TreeAnnotator v2.6.3 [219], manteniendo las medias de las edades y los intervalos de mayor densidad de posteriores (HPD por sus siglas en inglés) con 95 % de confianza para cada nodo. Se obtuvieron las distribuciones de posteriores para los nodos de interés utilizando la función beastio::getCladeHeight v.0.3.3 en R, y se visualizaron los resultados con ggplot2. Se compararon las medias de las HPD de diferentes nodos, y se comprobó la significancia de las diferencias mediante la función stats::t.test para aplicar un test t con la modificación de Welch.

5.3.4. Construcción y análisis del pangenoma

Se utilizaron pewit y pagoo, según se describe en Subsección 4.3.7.

Análisis funcional del genoma accesorio: Para identificar funciones enriquecidas en cada sublinaje, se realizó un análisis funcional del pangenoma. Para esto se extrajeron las secuencias aminoacídicas de cada grupo de genes ortólogos, los cuales fueron anotados utilizando eggNOG mapper [173],

utilizando la base de datos de referencia eggNOG v.5.0 [174]. Se realizó un análisis de correspondencias multivariado (MCA), a partir de la *panmatriz*, a la que se le sustituyó cada valor 1 por la categoría COG correspondiente a cada cluster de genes ortólogos, si se encontraba disponible. Paralelamente, se realizó un DAPC a partir de la suma de los genes de cada categoría.

5.3.5. Identificación de factores de virulencia y otros elementos de asociados a la patogenicidad

Factores de virulencia

Los factores de virulencia fueron detectados con abricate v.1.0.1 [220], utilizando como referencia la base de datos VFDB [221, 222], actualizada a diciembre de 2022. Alternativamente, se utilizaron los perfiles HMM específicos de cada factor de virulencia de interés, descargados desde Interpro para realizar la búsqueda utilizando HMMER v.3.3.2 [223], utilizando los *thresholds* especificados en los perfiles, siempre que estuvieran disponibles.

Sistemas de secreción

Para identificar los sistemas de secreción presentes en los genomas del *dataset*, se ejecutó el programa MacsyFinder v.2.0 [224] a partir de las secuencias aminoacídicas, utilizando las opciones dbtype gembase y replicon-topology linear para todos los genomas. Los reportes obtenidos se analizaron en R.

Proteínas Fic

Se utilizaron dos abordajes para la identificación de los genes que codifican las proteínas Fic en los genomas del *dataset*.

- BLASTp: Se realizó la búsqueda de proteínas Fic a partir de 18 secuencias aminoacídicas utilizadas en el artículo de Sprenger et al. [203] que se listan en la Tabla 5.1. Se utilizó BLASTp para la búsqueda, definiendo como mínimos una cobertura de 90 % y una identidad de 70 %.
- 2. Mapeo: Se mapearon los *reads* correspondientes a 79 experimentos de secuenciación masiva de genomas de *C. fetus* contra cada una de las secuencias nucleotídicas de las proteínas Fic listadas en la Tabla 5.1. Los alineamientos se hicieron con Bowtie2 y los resultados se procesaron con Samtools v.1.9 y su implementación en R Rsamtools v.2.2.1. Con el fin de estimar la frecuencia relativa de genes codificantes para proteínas Fic se calculó el RPKM (proporción de *reads* mapeados sobre totales (1 × 10⁶) cada 1000 bases).

5.3.6. Búsqueda de elementos móviles

Identificación de secuencias de origen plasmídico

Se emplearon dos estrategias para la identificación de secuencias plasmídicas en los genomas de *C. fetus* ST4.

PlaSquid: En un primer abordaje se utilizó PlaSquid [225], una herramienta desarrollada en nuestro laboratorio, que predice la procedencia (cromosomal o plasmídica) de los *contigs* de un genoma ensamblado en base a la identificación de secuencias típicamente plasmídicas, utilizando para

ID (NCBI)	Nombre	Alias	Longitud (pb)	Clase
AHE94549.1	Fic C. fetus subesp. venerealis cfvi03/293	Fic_01	933	Ι
AIR78561.1	Fic C. fetus subesp. venerealis 04/554	Fic_02	726	II
OCS28453.1	Fic C. fetus subesp. venerealis LMG6570	Fic_03	687	II
OCS25815.1	Fic C. fetus subesp. venerealis CCUG33872	Fic_04	642	Ι
CDF65967.1	Fic4 C. fetus subesp. venerealis 84-112	Fic_05	1314	III
CDF65920.1	Fic3 C. fetus subesp. venerealis 84-112	Fic_06	672	Ι
CDF65254.1	Fic1 C. fetus subesp. venerealis 84-112	Fic_07	837	II
CDF65253.1	Fic2 C. fetus subesp. venerealis 84-112	Fic_08	921	Ι
CAL11741.1	Fic Yersinia enterocolitica	Fic_09	588	Ι
KEA46573.1	Fic1 C. mucosalis	Fic_10	540	Ι
KEA45319.1	Fic2 C. mucosalis	Fic_11	771	II
AHK74499.1	Fic C. coli RM5611	Fic_12	1083	II
ETC95351.1	Fic C. coli K3	Fic_13	702	Ι
ERJ27453.1	Fic C. concisus ATCC 51561	Fic_14	378	II
EPH07314.1	Fic C. ureolyticus ACS-301-V-Sch3b	Fic_15	642	Ι
EEF14435.1	Fic C. rectus RM3267	Fic_16	1083	II
EJZ03953.1	Fic Streptococcus agalactiae STIR-CD-17	Fic_17	222	Ι
EFU71140.1	Fic C. upsaliensis JV21	Fic_18	702	II

Tabla 5.1: Proteínas Fic de referencia utilizadas en la búsqueda mediante BLASTp

esto la base de datos PLSDB [226]. En base a los resultados obtenidos se determinó qué plásmidos están representados en los genomas ensamblados, tomando como criterio de presencia de un plásmido si la longitud combinada de los *contigs* identificados como plasmídicos es igual o mayor al 70 % de la longitud del plásmido de referencia. Adicionalmente se calcularon las proporciones de bases en plásmidos por genoma.

Mapeo de datos de secuenciación masiva La segunda estrategia consistió en realizar mapeos de datos de secuenciación masiva con Bowtie2 v.2.3.4.1 [227], utilizando como referencia las secuencias plasmídicas identificadas mediante el abordaje anterior. La información asociada a los datos de secuenciación masiva utilizados se muestra en la Tabla 8.4. Los resultados obtenidos se procesaron con Samtools y Rsamtools [228, 229]. Para determinar los perfiles de presencia/ausencia, se consideró un plásmido "presente" en un genoma si al menos el 50 % de su longitud es mapeada por al menos un *read*. Se calculó el RPKM para cada plásmido presente según los resultados de cobertura. Paralelamente, se mapearon los datos de secuenciación sobre cada CDS obtenido de los plásmidos seleccionados para el mapeo como referencia. Los CDSs mapeados por cada grupo de *reads* se anotaron con eggNOG para obtener los perfiles funcionales de los genes plasmídicos identificados.

Identificación de elementos de inserción

Se buscaron secuencias de inserción en los genomas utilizando diferentes aproximaciones:

Islandviewer Para detectar la presencia de islas génicas en los genomas se utilizó Islandviewer, [230], disponible como servidor Web en https://www.pathogenomics.sfu.ca/islandviewer/.

Detección de ISCfe1 La secuencia de inserción ISCfe1 se identifcó en los genomas de *C. fetus* ST4 utilizando BLASTn con parámetros por defecto, y la secuencia de *accession* AM260752.1 como referencia. Los resultados se filtraron, reteniendo los *hits* con identidad superior a 90 % y cobertura de la referencia mayor a 70 %. Se extrayeron las secuencias de los *loci* ISCfe1 identificados, y se alinearon utilizando el paquete de R DECIPHER. Se construyó una filogenia a partir del alineamiento mediante el método ML. Otra estrategia para la identificación de ISCfe1 utilizó el mapeo de *reads* provenientes de 79 experimentos de secuenciación genómica masiva utilizando Bowtie2

Paralelamente se realizó la búsqueda de elementos integrativos utilizando BLASTn [146], con secuencias descargadas desde ICEdb como referencia. **Detección de sistemas CRISPR-Cas** Los sistemas de defensa contra elementos genéticos móviles CRISPR-Cas se detectaron utilizando el servidor web de CRISPRCasFinder [231], disponible en https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index, que utiliza el ya mencionado Macsyfinder para identificar los genes *cas* y el tipo y subtipo de sistema. Los resultados obtenidos fueron procesados en R, utilizando *scripts* elaborados específicamente para estos datos.

5.4. Resultados

5.4.1. Dataset

A partir del *dataset* presentado en el capítulo anterior, se retuvieron aquellos genomas que corresponden a aislamientos bovinos con genotipo ST4, pertenecientes al linaje L1, según el resultado de clusterización devuelto por hierBAPS. Así es que el *set* de datos comprende 103 genomas obtenidos a partir de cepas de *C. fetus* aisladas en el período comprendido entre los años 1952 y 2017 (Metadatos en la Tabla 8.2). Si bien se mantiene la representatividad de los cinco continentes, la mayoría de los genomas provienen de Europa y Sudamérica, siendo España y Argentina los países más representados, respectivamente. En la Figura 5.3 se muestra la distribución geográfica de los genomas de este *dataset*.





5.4.2. Estructura poblacional y distribución geográfica de *C. fetus* ST4

Con el fin de investigar la estructura poblacional y la evolución genómica de *C. fetus* ST4, se tomó como punto de partida el alineamiento resultante de todas las regiones comunes a todos los genomas, compuesto por 1.372.915 posiciones. Este alineamiento primario o genoma *core*, cuya longitud equivale al 67,93 % de la longitud total del genoma de la cepa cfvi84112, utilizada como referencia, fue filtrado con el fin de remover las regiones con evidencia de recombinación utilizando Gubbins. En este procedimiento también se removieron las posiciones no informativas, con lo que el alineamiento final o marco clonal está compuesto por 1763 posiciones polimórficas.

El alineamiento del marco clonal así obtenido fue el punto de partida tanto para la construcción de filogenias, como para identificar la presencia de sublinajes intra-L1 y observar diferencias en distancias pareadas, diversidad nucleotídica, entre otros.

La filogenia fue construida con el método ML y el modelo GTR, utilizando el programa raxm1 v.8.2.11, y el marco clonal como *input* (Ver Subsección 5.3.2). El árbol resultante, al que se le eliminó el grupo externo, se muestra en la Figura 5.4. Allí se observa la formación de 4 *clusters*, que se muestran encerrados por líneas punteadas, donde los *tips* se colorearon de acuerdo al origen geográfico de cada aislamiento, observándose una prevalencia de genomas sudamericanos en los *clusters* 1 (amarillo) y 2 (verde), mientras que en los *clusters* 3 y 4 (azul y rojo, respectivamente) predominan aislamientos de origen europeo. La Figura 5.5a muestra el número de genomas de cada linaje, y la Figura 5.5b la correspondencia entre linajes y el origen geográfico de los aislamientos. La topología del árbol demostró
correspondencia con las regiones geográficas que dieron origen a los aislamientos, según un análisis bayesiano de significancia de *tips* realizado con BaTS [209]. Según los resultados obtenidos, de las 6 regiones geográficas definidas, Europa, América del Sur, América del Norte y Australia dieron resultados de correspondencia significativos, con valores p de entre 0,01 y 0,02. Las dos regiones restantes, África y Reino Unido no dieron resultados de asociación significativos. A partir de la correspondencia entre el origen geográfico y la estructura poblacional obtenida se generaliza la asociación, por lo que los linajes SL1 y SL2 pasan a nombrarse como SA1 y SA2, y los linajes SL3 y SL4 como EU1 y EU2.

Las metodologías independientes de filogenia, dieron resultados congruentes con la topología del árbol filogenético mostrado en Figura 5.4, formándose 4 *clusters* completamente definidos. En los gráficos resultantes del MCA, mostrados en Figura 5.6a y Figura 5.6b, se observa la separación de todos los sublinajes en las proyecciones de las primeras tres dimensiones. En la primera dimensión (Figura 5.6a) resalta la separación de EU2 con respecto a los demás sublinajes, separación que se visibiliza también en el árbol filogenético (Figura 5.4). La segunda dimensión separa el clústeres EU1 de SA1 y SA2, mientras que en la tercera dimensión se observa la separación de SA2, muy cercano a SA1 tanto en la filogenia como en las dos primeras dimensiones del MCA.

Distancia intra y entre sublinajes Se determinaron las distancias pareadas en SNPs, tomando como punto de partida el alineamiento del marco clonal, encontrándose una media general de 180,5 SNPs de distancia (IQR=77-272). Al agrupar los genomas en sublinajes encontramos que el SA1 es el que presenta mayor rango de distancias pareadas, según se puede observar en la Tabla 5.2.





Antigüedad de los sublinajes

Con el fin de profundizar sobre la estructura poblacional y el origen geográfico de los aislamientos, se propuso evaluar la señal temporal de cada sublinaje,

con el fin de determinar el estado ancestral más probable en cada ancestro común.

Evaluación de la diversidad nucleotídica Primero, se evaluó la diversidad nucleotídica de cada sublinaje, determinada a partir de las posiciones polimórficas derivadas del alineamiento del marco clonal. Bajo los supuestos de ausencia de recombinación y tasas de mutación constante a lo largo de la topología de la filogenia, una mayor diversidad nucleotídica sería consecuencia de un mayor tiempo de evolución de un linaje. Así, se encontraron diferencias significativas (*t* de Welch, $p=2,2 \times 10^{-16}$), con una mayor diversidad nucleotídica en los linajes europeos que en los sudamericanos, tal como se observa en la Figura 5.7, lo cual, según el análisis realizado, se traduce en una mayor antigüedad de los linajes europeos.

Análisis temporal bayesiano Para determinar la antigüedad de los nodos que dan origen a cada linaje, se realizó un análisis bayesiano utilizando BEAST, a partir de un alineamiento construido con las secuencias nucleotídicas concatenadas de los genes *core* en los que no se detectó recombinación ni selección, según los resultados de PhiTest y D de Tajima, respectivamente. Estos genes provienen del pangenoma reconstruido con Roary a partir de set de genomas extendido (n=151), que incluye cepas no ST4, de origen animal y clínico, de los cuales se cuenta con la fecha de aislamiento.

	Mín.	Q1	Mediana	Media	Q3	Max.
intra-SA1	0	26	47	46,76	54	182
intra-SA2	4	24	47	83,50	142	151
intra-EU1	0	37	45	43,22	55	94
intra-EU2	0	8	12	18,83	36	48

Tabla 5.2: Medidas resumen de las distancias pareadas en SNP intra sublinajes
(Min=mínimo, Q=cuartil, Max=máximo)



Primero, se realizó la selección del modelo bayesiano mediante NS, favoreciendo el modelo BS (*likelihood*: -62327, SD:1,08). El mejor modelo de sustitución nucleotídica según MoldelFinder e IQ-TREE fue HKY+IA.

A partir de la distribución de posteriores, que pueden verse en la Figura 5.9, se determinaron los tMRCA de cada sublinaje (EU1, EU2, SA1, SA2), grupo de linajes (EU y SA), además del nodo que da origen a cada grupo (EU-SA), resultados que se describen en la Tabla 5.3 y en los nodos relevantes de la filogenia mostrada en la Figura 5.8. Para determinar si los linajes europeos son más antiguos que los sudamericanos, se realizó un test *t* de Welch a partir de las probabilidades posteriores de los tMRCA correspondientes, que dio como resultado una diferencia significativa con $p=8,34 \times 10^{-6}$. Este resultado coincide con el obtenido a partir del análisis de la diversidad nucleotídica, confirmando que los linajes europeos son más antiguos que los sudamericanos, son más antiguos que los europeos son más antiguos que los linajes europeos son más antiguos que los sudamericanos, situando la cladogénesis hace aproximadamente 639 años en promedio (mediana 458, HDP:118-1438, 95 % confianza).





	Media	Mediana	HPD(CI95) ¹
EU1	196	131	16-479
EU2	108	82	50-207
SA1	223	129	25-583
SA2	35	15	4-92
EU	388	265	74-926
SA	352	196	21-972
SAEU	639	458	118-1438

 Tabla 5.3: Medidas resumen de las probabilidades posteriores de tMRCA.

¹ Intervalos de mayor densidad posterior (*Highest Posterior Density intervals*), determinados a un nivel de confianza de 95 %

5.4.3. Rasgos genéticos

Pangenoma

El pangenoma, obtenido al ejecutar Pewit sobre el *dataset* de 102 genomas anotados, está compuesto por 3816 grupos de genes ortólogos. De estos, 1639 forman parte del *set core*, al estar presentes en al menos el 95 % de los genomas que forman parte de este análisis (Ver Tabla 5.4). A continuación se detallan las principales características del pangenoma obtenido.

El pangenoma se encuentra abierto ($\alpha = 0,90$), estimándose que se completaría con un total de 4.598 grupos de genes ortólogos, según el estimador de Chao [232], que representa la asíntota horizontal en el gráfico de la Figura 5.10.

El número de genes accesorios se extrajo de la matriz de presencia/ausencia de grupos de ortólogos para cada genoma, la panmatriz, con el fin de determinar si existen diferencias entre los diferentes sublinajes. Estos resultados se muestran en la Tabla 5.5. Puede observarse un leve aumento del número de genes accesorios en los genomas sudamericanos con respecto a los europeos, que resultó ser significativa con p=0,019, según el test de Wilcoxon con ajuste de Bonferroni. A nivel de sublinaje, las únicas diferencias

	Número de grupos	Distribución
Total	3816	
Core	1639	\geq 95 % de los genomas
Shell	1499	< 95 %
Cloud	678	presente en 1 a 3 genomas

Tabla 5.4:	Composición del	pangenoma.
------------	-----------------	------------



significativas encontradas fueron entre SA1 y EU2 ($p=1,59 \times 10^{-5}$) y entre SA1 y SA2 ($p=7,273 \times 10^{-5}$).

Tabla 5.5: Medidas resumen del número de genes accesorios por genoma paracada sublinaje. (Min=mínimo, Q=cuartil, Max=máximo)

	Min.	1er Q.	Mediana	Media	3er Q.	Max.
SA1	220	317,5	342,5	342,03	367,3	472
SA2	257	289,0	292,0	292,85	294,0	324
EU1	245	286,5	313,0	322,57	346,3	466
EU2	276	289,0	295,0	299,74	304,5	359
SA	220	294,0	330,0	329,49	358,0	472
EU	245	288,5	198,0	312,27	333,0	466



Análisis de componentes principales del genoma accesorio: Se realizaron análisis de componentes principales, PCA (Figura 5.12a) y de componentes principales discriminatorios, DAPC (Figura 5.12b) para determinar si el perfil de presencia/ausencia de genes en la panmatriz puede separar los sublinajes y determinar cuáles son los grupos de genes ortólogos que más contribuyen a esta separación. Ambos análisis logran separar totalmente los sublinajes, aunque en el DAPC los sublinajes aparecen más separados y compactos. Esto es debido a la naturaleza del análisis, que busca minimizar la distancia intra grupo, a la vez que maximiza la distancia entre grupos.

Análisis funcional del genoma accesorio: Para identificar funciones enriquecidas en cada sublinaje, se realizó un análisis funcional del pangenoma, anotando una secuencia extraída de cada grupo de genes ortólogos que forman parte del genoma accesorio (ocurrencia < 95 %), comprendiendo los *sets Cloud* y *Shell*, descritos en la Tabla 5.4. De los 2.167 grupos de genes accesorios, se lograron anotar 1.353 (62 %), obteniendo la categoría COG



en 1.095 de ellos. La distribución de las categorías COG encontradas para la totalidad del genoma accesorio se muestra en la Figura 5.13.

Dada la separación de los sublinajes en el DAPC (Figura 5.12b), se procedió a identificar los grupos de genes ortólogos más contribuyentes a esta separación, seleccionados utilizando un *threshold* de 0,0165, que corresponde al tercer cuartil de las contribuciones.

Se realizó la anotación funcional de 231 de los 545 grupos de genes ortólogos más contribuyentes. Su distribución se muestra en la Figura 5.14. **S** es la categoría COG más representada, que incluye los genes de función desconocida. Lo siguen las categorías**U** y **L**, correspondiente a tráfico intracelular, secreción y transporte vesicular y replicación, recombinación y reparación, respectivamente. Estas corresponden a las categorías más representadas en el genoma accesorio (Ver Figura 5.13).





Siguiendo con la caracterización funcional del genoma accesorio, se realizaron análisis de correspondencias multivariados (MCA), para determinar si los sublinajes tienen perfiles funcionales diferentes. En un primer análisis, cuyo resultado puede verse en la Figura 5.15a, se muestra como los genomas cfv10354 y 1273269.3, pertenecientes a los sublinajes EU1 y SA1 respectivamente están distanciados de sus respectivos grupos, razón por la cual se eliminan y se repite el MCA sin *outliers* (Figura 5.15b). En el último puede observarse la cercanía de los sublinajes SA1 y SA2, lo cual puede traducirse en perfiles funcionales similares, mientras que los sublinajes EU1 y EU2 se distinguen completamente entre sí y con el clústeres SA1-SA2.



Para determinar qué grupos de genes ortólogos son capaces de discriminar entre sublinajes se seleccionaron aquellos cuya presencia o ausencia es exclusiva de un sublinaje. De esta manera se retuvieron 102 grupos de ortólogos discriminantes, de los cuales 45 fueron anotados con eggNOGmapper. Los grupos discriminantes anotados se describen en la Tabla 5.6 y sus perfiles de presencia/ausencia se muestran en la Figura 5.16. Se identificaron dos



grupos de ortólogos exclusivos de los sublinajes sudamericanos, uno de función desconocida y otro que codifica la proteína FeoB. Esta proteína, que está involucrada en la adquisición de Fe(II), se encontró pseudogenizada en los genomas europeos. A diferencia de los sublinajes sudamericanos, en los europeos se encontraron 29 grupos discriminantes, de los cuales sólo 6 fueron anotados, predominando entre éstos proteínas fágicas (Ver Tabla 5.6).

Tabla 5.6: Grupos de genes ortólogos discriminantes, cuya presencia (o ausencia) es diferencial en uno o varios sublinajes. La columna *Presencia* indica en qué sublinajes fue identificado el clústeres.

clústeres	Presencia	COG	Descripción	Categoría general
group0007	SA1,SA2,EU1	NT	Proteína quimiotaxis	Movilidad y señalización
group0081	SA1,SA2,EU2	Е	Transportador ABC de aminoácidos	Metabolismo y transporte de aminoácidos
group0210	EU1,EU2	S	Proteína ribosomal	Poco caracterizada
group0233	SA1,SA2	S	Función desconocida	Poco caracterizada
group0239	SA1	Е	Aminoácidos ramificados	Metabolismo y transporte de aminoácidos
group0277	SA1,SA2,EU2	Р	Permeasa de transportador ABC	Metabolismo y transporte de iones inorgánicos
group0357	SA1	С	Citocromo	Metabolismo energético
group0419	EU2	С	Citocromo C oxidase subunidad II	Metabolismo energético
group0433	EU2	С	Reducción de nitrito a amonio	Metabolismo energético
group0525	SA1,SA2,EU1	S	Dominio de función desconocida (DUF1738)	Poco caracterizada
group0585	SA1,SA2,EU1	S	Proteína de función desconocida (DUF86)	Poco caracterizada
group0669	SA1,SA2	Р	Adquisición de Fe(II)	Metabolismo y transporte de iones inorgánicos
group1001	EU2	Κ	Regulador transcripcional	Transcripción
group1006	EU2	L	Dominio de unión a ADN trasposasa	Replicación y reparación
group1133	EU2	NT	Proteína quimiotaxis	Movilidad y señalización
group1143	SA2	F	Familia 5'-nucleotidasa family	Metabolismo y transporte de nucleótidos
group1144	SA1,EU1,EU2	F	Familia 5'-nucleotidasa family	Metabolismo y transporte de nucleótidos
group1299	EU1	С	Subunidad pequeña	Metabolismo energético
group1432	EU1,EU2	S	Proteína tape measure	Poco caracterizada
group1599	EU2	L	Dominio core integrasa	Replicación y reparación
group1732	EU1,EU2	S	Proteasa fágica familia HK97	Poco caracterizada
group2198	SA1,SA2,EU1	U	Proteína con dominio relaxasa	Transporte y secreción
group2353	EU2	L	Superfamilia RAMP	Replicación y reparación
group2354	EU2	С	Citocromo oxidasa subunidad II, dominio periplasmático	Metabolismo energético
group2355	EU2	Р	Proteína de unión a cobre periplasmática NosD	Metabolismo y transporte de iones inorgánicos
group2368	EU2	М	Dominio C-terminal TonB	Biogénesis de membrana o pared celular
group2623	SA2	S	Nucleotidil transferasa: AbiEii toxin, Sistema TA tipo IV	Poco caracterizada
group2847	EU1	С	Hidrogenasa hydA subunidad pequeña	Metabolismo energético
group2978	EU1,EU2	S	Proteína fágica conservada	Poco caracterizada
group3011	EU1,EU2	S	Proteína fágica tail	Poco caracterizada
group3039	SA1,SA2,EU1	S	Dominio nucleotidyltransferasa	Poco caracterizada
group3133	SA1,SA2,EU1	L	Superfamilia RAMP	Replicación y reparación
group3176	SA1,SA2,EU1	L	Resolvasa, dominio N-terminal	Replicación y reparación
group3288	EU1,EU2	L	Integrasa, dominio central	Replicación y reparación
group3290	SA1	L	Integrasa, dominio central	Replicación y reparación
group3316	SA1,SA2,EU2	ET	Proteína bacteriana de unión a soluto	Metabolismo de aminoácidos y señalización
group3434	SA1,SA2,EU1	NU	T4SS VirB11	Movilidad y transporte y secreción
group3478	SA1	S	Dominio terminasa tipo RNasaH	Poco caracterizada
group3528	SA1,SA2,EU1	S	Proteína de función desconocida (DUF3991)	Poco caracterizada
group3599	SA1,SA2,EU1	U	COG2948 T4SS VirB10	Transporte y secreción
group3619	SA1,SA2,EU1	U	TrbL/VirB6 proteína de transferencia de plásmido conjugativo	Transporte y secreción
group3642	SA1,SA2,EU1	Р	Canal de potasio	Metabolismo y transporte de iones inorgánicos
group3644	EU2	Р	Canal de potasio	Metabolismo y transporte de iones inorgánicos
group3692	SA1,SA2,EU1	U	T4SS, VirD4	Transporte y secreción
group3752	SA1,SA2,EU1	U	COG3736 T4SS, VirB8	Transporte y secreción

Pangenoma de cada sublinaje: Para estudiar la dinámica individual de cada sublinaje se estudiaron sus pangenomas por separado. Esto permitió reconocer al pangenoma del sublinaje SA1 como el más numeroso, pero también como el de mayor proporción de genes accesorios, observación compatible con un proceso de adaptación. En contraposición a éste, el mayor



número de genes *core* fue observado en el sublinaje SA2, aunque la mayor proporción *core*/accesorio se observó en EU2, con una proporción de 78 % de los genes de su pangenoma en el grupo *core*. La Figura 5.17 muestra cómo se distribuyen los grupos de genes ortólogos en el pangenoma de cada sublinaje.

5.4.4. Análisis de las regiones recombinantes

Las regiones recombinantes, filtradas del alineamiento del genoma *core* para la construcción de la filogenia, se analizaron para determinar su distribución en el genoma *core*, y los genes comprendidos en estas regiones. Se detectaron cinco regiones o bloques recombinantes, distribuidos en dos regiones del alineamiento del genoma *core*:

- La primera región comprende los bloques referidos como bloque 1 y bloque 2, de 15.124 y 2.463 pares de bases respectivamente, ubicados entre las posiciones 438.390 y 464.342 del genoma *core*. Estos bloques recombinantes, exclusivos de los linajes sudamericanos, contienen los genes *panB*, *ruvB*, *mlaF* y *sapA* en el bloque 1, y *epsL*, *radA* y *tlpA* en el bloque 2, tal como se muestra en la Figura 5.18.
- La segunda región, comprendida entre las posiciones 1.197.046 y 1.206.031, contiene los bloques recombinantes 3 (exclusivo de EU2), 4 (presente en un clado de SA2) y 5 (ausente en EU2), de 9.707, 8.711 y 4.224 pb de longitud respectivamente. En esta región se encontraron los genes *pseB*, *pseH*, *flaA*, *trmB*, *tsqA*, *ktrB* y *ktrA* (Ver Figura 5.18).

La distribución de los bloques recombinantes entre los diferentes linajes deja en evidencia una aparente barrera de recombinación que separa los genomas de los sublinajes SA1, SA2 y EU1 de los EU2, ya que se encontró un único bloque recombinante exclusivo en EU2.

Cuantitativamente también se encuentran diferencias en cuanto al número de bases en bloques recombinantes por base. En general, se observaron valores significativamente mayores en los sublinajes sudamericanos, como se observa en Figura 5.19b. Individualmente, se encontró mayor proporción de bases recombinantes en EU2 que en EU1 (Figura 5.19a), debido a la presencia del bloque recombinante exclusivo de EU2.





5.4.5. Identificación de factores de virulencia

Toxina citoletal distensora (CDT)

Como indicativo de producción de la toxina CDT se evaluó la presencia del gen *cdtB*, que codifica la subunidad B de la toxina. Para esto, se utilizaron dos estrategias:

abricate: La primera estrategia consistió en utilizar la herramienta abricate con la base de datos VFDB como referencia, ejecutada en una primera instancia con parámetros por defecto (Ver Subsubsección 5.3.5), sin obtener *hits* positivos para ningún genoma. Al disminuir los niveles de identidad de los *hits* reportados por la herramienta se obtuvieron resultados positivos, correspondientes a genes *cdtB*, con con niveles de cobertura e identidad superiores a 70 y 66 % respectivamente en todos los genomas estudiados. Si bien no se encontraron diferencias en el número de genes *cdtB* encontrados, se observaron menores niveles de cobertura en los *hits* de genomas del sublinaje EU2, sugiriendo la presencia en este sublinaje de una variante diferente a las encontradas en los demás sublinajes.

BLASTp: La segunda estrategia consistió en una búsqueda mediante BLASTp en la que se utilizaron como referencias las secuencias de cuatro proteínas CdtB de C. fetus extraídas de la base de datos InterPro (AOAAE6IYS4, A0AAE6IYV3, A0AAE6IX80 y W8VYQ8), y parámetros por defecto. Esta búsqueda dio como resultado la identificación de entre dos y cuatro hits por genoma. Los genomas del sublinaje EU2 tuvieron en su mayoría cuatro hits, dos de los cuales presentaban coberturas notablemente más bajas, de entre 30 y 50 %. Los genes correspondientes a los hits de baja cobertura se contrastaron con información obtenida a partir del análisis del pangenoma, encontrándose que ocupan posiciones contiguas en los genomas del sublinaje EU2, tal como se muestra en el mapa genético de la Figura 5.21. Este análisis, complementado con un análisis de secuencia permitió identificar la presencia de un SNP (G/T) en el gen cdtB de los genomas EU2 responsable de la incorporación de un codón STOP prematuro que generaría un producto de 182 aminoácidos en lugar de los 265 encontrados en la proteína CdtB codificada en los genomas de los otros sublinajes.



Proteínas Fic

A continuación se describen los resultados obtenidos mediante la búsqueda mediante BLASTp y mapeo de *reads* contra secuencias Fic de referencia.

BLASTp: No fue posible identificar proteínas Fic codificadas en los genomas 15031_S95 y 17059_2.36 mediante BLASTp. En el resto de los genomas se identificaron entre 1 y 6 CDSs que codifican proteínas Fic, con identidades y coberturas superiores a 70 y 90 % respectivamente. Al evaluar



el número de proteínas Fic codificadas por genoma (Ver Figura 5.22) se encontró un número significativamente mayor en los genomas europeos (Figura 5.22b), y dentro de estos, en los pertenecientes al sublinaje EU2 (Figura 5.22a) (Wilcoxon test con ajuste de Bonferroni: EU1 vs. EU2: $p=2,2 \times 10^{-4}$; EU vs. SA: $p=2,6 \times 10^{-8}$). Sin embargo, no se encontraron diferencias significativas entre los sublinajes sudamericanos entre sí. Las Fic de clase III se encontraron en los genomas de algunas cepas de SA1 y en un genoma EU1, mientras que se observó que las Fic de clase I y II se encuentran en todos los sublinajes con entre una y tres proteínas por genoma. Se observó también un mayor enriquecimiento en las Fic de clase II en el sublinaje EU2 (ver Figura 5.22c).

Mapeo: Con excepción del genoma 15031_S95, fue posible, mediante mapeo, identificar *reads* correspondientes a proteínas Fic en todos los genomas analizados. En general, se obtuvieron valores de RPKM considerablemente mayores en los genomas europeos que en los americanos (Wilcoxon test con ajuste de Bonferroni: EU vs. SA: $p=4,0 \times 10^{-9}$), y dentro de los europeos, se obtuvieron valores significativamente mayores para los genomas del sublinaje EU2 (EU1 vs. EU2: $p=2,3 \times 10^{-6}$; Ver Figura 5.23).

Si bien se encontraron diferencias entre los resultados obtenidos para cada estrategia, pueden encontrarse puntos en común en lo que refiere la ausencia de proteínas Fic en el genoma 1531_S95 y la mayor abundancia encontrada en los genomas europeos con respecto a los sudamericanos.

Sistemas de secreción

Para detectar sistemas de secreción se utilizó el programa Macsyfinder, tomando como referencias los modelos disponibles en MacsyModels. Este



programa se basa en la identificación de proteínas con homología a las presentes en los diferentes sistemas de secreción descritos, además de un conjunto de reglas relacionadas con su organización genómica, a partir de los cuales se generan modelos que establecen los requerimientos mínimos para considerar que un sistema de secreción está presente en un replicón o *contig*.



Fig. 5.23: Boxplots donde se representan los valores de RPKM obtenidos de los mapeos de *reads* provenientes de la secuenciación de 79 genomas de *C. fetus* bovinos contra las secuencias nucleotídicas de las proteínas Fic listadas en la Tabla 5.1, para cada sublinaje a la izquierda y entre EU y SA a la derecha. Se tomaron en cuenta sólo los resultados de los mapeos que contaran con una cobertura mayor o igual al 70 %, asignando valores de RPKM de 0 a todos los mapeos que no alcanzaran el mínimo de cobertura mencionado.

Así se obtuvieron elementos vinculados a los sistemas de secreción del tipo 1 (T1SS), 4 (T4SS), 4aP (T4aP) y 5 (T5SS), según puede observarse en los gráficos de cajas de las figuras Figura 5.26a y Figura 5.26b y en el *heatmap* de la Figura 5.25. A continuación se detallan los resultados obtenidos para cada tipo de sistema encontrado.

- **T1SS** La presencia de T1SS es una constante en todos los genomas estudiados, identificándose un sistema en la mayoría de los genomas, sin variaciones de ocurrencias entre linajes.
- **T4aP** El *pilus* de tipo IV es considerado un factor de virulencia en bacterias patógenas y permite el movimiento mediante *twitching*. Sólo fue encontrado en un par de genomas de SA1 y EU1.

T5aSS Para los T5SS del tipo **a** (T5aSS), que constituyen autotransportadores clásicos, se identificaron al menos dos sistemas por genoma en todos los genomas estudiados. En algunos genomas se encontraron hasta tres sistemas, pero la distribución de estos no es específica de





un sublinaje o grupo de linajes, sino que se encuentra en un número variable de genomas de cada linaje.

- **T5bSS** Para los T5bSS, transportadores de dos componentes, la identificación depende únicamente de la presencia de la proteína traslocadora, responsable de formar un poro a través de la membrana externa de la bacteria, a través de la cual se secreta una proteína que generalmente es codificada en la cercanía del traslocador. A diferencia que los T5SS de tipo **a** se encontró un único sistema en todos los linajes, menos en EU2. La ausencia de este sistema de secreción en este sublinaje se debe a que en el gen *fhaC*, que codifica el transportador y activador de la hemolisina, ocurrió una deleción que produjo un cambio en el marco de lectura que incorpora un codón *STOP* prematuro. Se observa además en el mapa genético mostrado en la Figura 5.27 la adición de dos ORFs adicionales que codifican proteínas hipotéticas conservadas, seguidos por la sección del gen *fhaC* faltante en el ORF anterior.
- **T4SSt** Los sistemas de secreción de tipo 4 t están implicados principalmente en la secreción de proteínas efectoras a través de la membrana externa, hacia el exterior celular o bien, hacia el citoplasma de otra



célula. MacsyFinder utiliza un modelo de identificación que requiere la presencia obligatoria del gen virB4, junto a la presencia de t4cp1 u otros genes vir opcionales (al menos 6), y la ausencia de mobB, con una distancia intergénica máxima de 30 genes. Se detectó pT4SSt en más del 90% de los genomas del dataset, con entre 1 y 5 ocurrencias por genoma. Se identificó al menos un pT4SSt en todos los genomas pertenecientes a linajes sudamericanos, mientras que en los europeos se identificó pT4SSt en el 80% de los genomas. Los genomas en los cuales no se identificó este sistema pertenecen a un clado del sublinaje EU2. A diferencia de este sublinaje, en EU1, no sólo se demuestra la prevalencia de este sistema, sino que también se observaron en su mayoría entre 3 y 4 pT4SSt por genoma. En los linajes sudamericanos se encuentran también diferencias en cuanto al número de sistemas por genoma, ya que en SA1 se observa un mayor número de copias que en SA2 (2,737 vs. 1,286 pT4SS por genoma en promedio). De esta manera, encontramos, a grandes rasgos, un mayor número de pT4SSt en genomas de linajes sudamericanos que en los europeos, observándose una explosión en el número de sistemas en los sublinajes EU1 y SA1, donde se encontraron hasta 5 sistemas por genoma, como se observa en la Figura 5.26. La ausencia de pT4SSt en los genomas EU2 se debe a que no se encontró el número mínimo de genes vir accesorios requeridos para identificar el sistema. La localización de estos sistemas en el extremo de los contigs puede ser una de las razones que dificulta su identificación en los genomas muy fragmentados.



Fig. 5.27: Contexto génico del gen *fhaC* en genomas representativos de los cuatro sublinajes. Una deleción en el gen que codifica la proteína FhaC, un transportador de hemolisina o una hemaglutinina filamentosa, explica la ausencia de T5bSS en los genomas de EU2 mostrados en rojo.



identificados por genoma. En (a) se muestra el *heatmap* que indica el número de veces que se identificaron T4SS en cada genoma. (b) y (c) muestran *boxplots* que representan en el eje y el número de T4SS identificados por genoma, distribuidos entre SA y EU y entre sublinajes, respectivamente.

5.4.6. Elementos móviles

Los elementos genéticos móviles cumplen un rol muy importante en la diferenciación de linajes, por lo cual se decidió identificar elementos conjugativos y/o movilizables, islas génicas y secuencias plasmídicas en los genomas.

Identificación de elementos conjugativos y movilizables con MacsyFinder

El modelo CONJScan, que forma parte de MacsyData, contiene información que permite identificar entre sistemas conjugativos (CONJ), conjugativos en decadencia (dCONJ) y movilizables (MOB), que se diferencian en el número y tipo de genes que presentan, considerándose un sistema CONJ como aquel que contiene toda la maquinaria para la formación del T4SS, además de la presencia obligatoria de los genes t4cp1/t4cp2 (factor acoplante conjugativo o *traD*) y un gen MOB. Un sistema que carece de los genes t4cp o *virB4* pero que conserva el gen MOB y los genes *virB*(1,2,3,5,6,7,8,9, 10, 11) constituye un sistema de conjugación en decadencia (dCONJ), mientras que un sistema que contiene sólo MOB y, opcionalmente, algunos genes *vir* constituye un sistema MOB. Debido a que se pueden considerar a los sistemas dCONJ y MOB como productos de la deleción de genes que constituyen el sistema de conjugación, estos se identifican y analizan de forma conjunta.

Presencia y número total de sistemas de conjugación y movilización:

Al analizar el número total de sistemas de conjugación reportados, se observa que este es significativamente mayor en los genomas sudamericanos que en los europeos (test de Wilcoxon EU vs. SA: p=0,0082). Al analizar los sublinajes por separado sólo se observaron diferencias entre SA2 y EU2 (test de Wilcoxon con corrección de Bonferroni EU2 vs. SA2: p=0,03).



utilizar el modulo CONJ de Macsydata en Macsyfinder. En el boxplot mostrado de la izquierda se muestra la distribución de los sistemas de conjugación por genoma en sublinajes, mientras que en el de la derecha se contrastan los genomas de los sublinajes sudamericanos con los europeos. Al aplicar el test de Wilcoxon con corrección de Bonferroni sólo se encontraron diferencias significativas al comparar SA2 con EU2 p=0,03. Al comparar los linajes sudamericanos con los europeos se encontró que la diferencia es significativa con p=0,0082.

Presencia y número de sistemas de conjugación y movilización por

tipo: Los resultados de la búsqueda de sistemas de conjugación y movilización por sistema se muestra en la Figura 5.30. En Figura 5.30a se muestra un *heatmap* adosado al árbol filogenético de los genomas bovinos ST4 de *C. fetus*, mientras que los *boxplots* de la Figura 5.30b se muestra la distribución de los sistemas identificados. En la Figura 5.30a se destaca la presencia de sistemas CONJ en todos los genomas de los sublinajes sudamericanos, mientras que en los sublinajes europeos no se identificaron estos sistemas en 9 genomas EU1 y 9 genomas EU2, lo que representa el 32 y 41 % respectivamente. Cabe destacar que en los genomas del sublinaje EU2 en los que no se encontró el sistema CONJ completo tampoco fue posible identificar T4SS t (ver Figura 5.28a), por lo que la ausencia de sistemas CONJ en estos genomas se debe a que no se detectó el gen *virB4*, o bien, a que no se encontrara el número suficiente de genes *vir* en su contexto génico. La diferencia entre genomas de los linajes sudamericanos y europeos fue significativa al aplicar el test de Wilcoxon (EU vs. SA: p=0,019, ver *boxplot* inferior izquierdo de la Figura 5.30b).

Por su parte, los sistemas dCONJ fueron significativamente menos frecuentes en genomas de los sublinajes SA1 y EU1, donde se encontraron con una frecuencia de 18 y 39 % respectivamente (test de Wilcoxon con corrección de Bonferroni: SA1: $p < 1 \times 10^{-6}$, EU1: p < 0,011).

Los sistemas MOB se encontraron en todos los genomas menos en un subgrupo del sublinaje EU2. Sin embargo, las diferencias entre sublinajes no resultaron ser significativas mediante el test estadístico utilizado, mas sí se encontró significativa la mayor cantidad de sistemas MOB en los genomas de los linajes sudamericanos, con p=0,0048 al aplicar el test de Wilcoxon.

Presencia y número de genes vir, t4cp y MOB: Los genes vir son parte de la maquinaria de los T4SS, involucrada en la secreción de proteínas y en la conjugación, por lo cual resulta interesante evaluar la distribución de estos tipos de proteínas entre sublinajes, y contrastar con lo observado para los sistemas T4SS y CONJ. La Figura 5.31 muestra un *heatmap* donde se observa la distribución de los genes *vir1-11*, t4cp1-2 y MOB identificados por Macsyf inder como parte de los sistemas de T4SS t o de conjugación. Se observa que la imposibilidad de identificar sistemas T4SS y CONJ completos en genomas del sublinaje EU2 se debe a que no se encontraron los genes *virB4* necesarios para completar la maquinaria de secreción.



del centro y de la derecha corresponden a los dCONJ y MOB.

5.4 Resultados



Búsqueda de secuencias plasmídicas

Para detectar *contigs* plasmídicos se utilizaron diferentes estrategias, que se describen en la Subsección 5.3.6. Las secuencias plasmídicas utilizadas como referencia se muestran en la Tabla 5.7.

Mapeo Esta estrategia consistió en alinear *reads* provenientes de 79 experimentos de secuenciación masiva públicos contra las secuencias plasmídicas de referencia mostradas en la Tabla 5.7, evaluándose tanto los perfiles de presencia/ausencia de plásmidos para cada uno de los experimentos de secuenciación masiva, como la profundidad de los mapeos.

Presencia Para determinar la presencia de un plásmido se tomó como valor de corte una cobertura de al menos el 75 % de la longitud de cada

Accession NCBI	Plásmido	Longitud (pb)
NC_010858.1	C. fetus subsp. venerealis pCFV108	3724
NZ_CP007000.1	C. fetus subsp. venerealis cfvi03/293 pCfviMP1	91400
NZ_CP007001.1	C. fetus subsp. venerealis cfvi03/293 pCfviMP2	35326
NZ_CP007002.1	C. fetus subsp. venerealis cfvi03/293 pCfviP3	3993
NZ_CP008809.1	C. fetus subsp. fetus 04/554 pCFF04554	25862
NZ_CP008812.1	C. fetus subsp. venerealis 97/608 pCFV97608-2	27124
NZ_CP014569.1	C. fetus subsp. venerealis 01/165 mp1	44765
NZ_CP014570.1	C. fetus subsp. venerealis 01/165 mp2	27417
NZ_CP043436.1	C. fetus subsp. venerealis NCTC 10354 p3226	27915
NZ_CP053829.1	C. hyointestinalis subsp. lawsonii CHY5 pCHY5	4413
NZ_CP059433.1	C. fetus CFViADRI1362 pCFViADRI1362_P1	40588
NZ_CP059434.1	C. fetus CFViADRI1362 pCFViADRI1362_P2	36566
NZ_CP059435.1	C. fetus CFViADRI1362 pCFViADRI1362_P3	35640
NZ_CP059436.1	C. fetus CFViADRI1362 pCFViADRI1362_P4	3993
NZ_CP059438.1	C. fetus CFViADRI545 pCFViADRI545_P1	48693
NZ_CP059440.1	C. fetus CFV08A1102-42A pCFV_08A1102_P1	37205
NZ_CP059442.1	C. fetus CFV08A948-2A pCFV_08A948_P1	38770
NZ_CP059446.1	C. fetus CFF09A980 pCFF_09A980_P1	52345

Tabla 5.7: Plásmidos utilizados como referencia

plásmido. Con este criterio, se obtuvo el perfil de presencia/ausencia de plásmidos mostrado en la Figura 5.32.

- Se visualiza la ausencia de los plásmidos NZ_CP053829.1, NZ_CP059436.1 y NZ_CP007002.1, muy similares entre sí, en los genomas del sublinaje EU1 (frecuencia: 4,5%). Estos plásmidos tienen la particularidad de portar el sistema toxina-antitoxina BrnT-BrnA, cuya función es detener el crecimiento bacteriano ante un estímulo de estrés.
- Por otra parte, los plásmidos NZ_CP043436.1 y NZ_CP008812.1 no se encontraron en ningún genoma del sublinaje SA2 mediante mapeo (frecuencia de 0%) pero sí en todos los genomas del sublinaje EU2. NZ_CP043436.1 y NZ_CP008812.1 tuvieron la misma distribución, ya que se trata de plásmidos muy similares entre sí, que difieren únicamente en que el primero contiene un gen que codifica una proteína hipotética que no está presente en

NZ_CP008812.1. La frecuencia de estos plásmidos en todos los sublinajes es superior al 77 %, por lo que se puede delimitar su ausencia al sublinaje SA2. Estos plásmidos codifican los sistemas toxina/antitoxina Fic y YafQ y poseen varios elementos de inserción y proteínas de origen viral, además de los elementos que constituyen la maquinaria de conjugación. YafQ es una endoribonucleasa que inhibe la traducción, al degradar el ARNm de forma secuencia específica [233].

- NZ_CP008809.1 y NZ_CP059433.1 son plásmidos que, como los anteriores, no fueron identificados en ninguno de los genomas del sublinaje SA2, pero sí en todos los de EU2. Estos plásmidos se diferencian en varios aspectos. En NZ_CP059433.1 se encuentran codificados los sistemas T/AT RelE, VapC, YafQ y Fic, y la toxina Zeta, mientras que en NZ_CP008809.1 sólo se encuentran VapC y YafQ parcialmente, mientras que Fic, junto a otros 5 genes se encuentran como pseudogenes. Las frecuencias de estos plásmidos en EU1 y SA1 son más bajas que los anteriores, alcanzando el 41 y 57 % para NZ_CP008809.1, y 41 y 76 % respectivamente para NZ_CP059433.1.
- Los plásmidos NZ_CP059442.1 y NZ_CP059440.1 se encontraron en baja frecuencia en los sublinajes sudamericanos (15 y 21 % respectivamente), pero fueron encontrados en casi todos los genomas europeos (73 y 90 % respectivamente).
- Profundidad Los mapeos también fueron utilizados para determinar si existen diferencias en el número de copias de los plásmidos con cobertura superior al 75 %, evaluando el índice RPKM (*reads* mapeados sobre el plásmido por millón de *reads* por kilobase). Los *boxplots* de la Figura 5.33 muestran las profundidades normalizadas (en RPKM) para cada plásmido entre las poblaciones sudamericanas y europeas. Se encontraron


Fig. 5.32: Perfiles de presencia/ausencia de plásmidos en genomas de *C. fetus*, identificados mediante mapeo de *reads*. Los rectángulos llenos denotan presencia y los vacíos ausencia del plásmido, considerando un plásmido presente si al menos el 75 % de su longitud está cubierta por al menos un *read*.

diferencias significativas en los niveles de profundidad de mapeo entre genomas sudamericanos y europeos, según un test de Wilcoxon con ajuste de Bonferroni, para 8 de 13 plásmidos analizados, observándose las mayores diferencias para el plásmido NZ_CP008809.1, para el que se encontraron mayores niveles de profundidad para los genomas europeos. Este plásmido se encontró ausente sólo en los genomas SA2, pero los niveles de RPKM en SA1 fueron significativamente inferiores a los de EU1 (p=0,023) y a los de EU2 (p=1,4 × 10⁻⁸).

PlaSquid, por otro lado, se utilizó para catalogar cada *contig* según sean de origen cromosómico o plasmídico, y evaluar diferencias en la proporción de bases plasmídicas y cromosomales por genoma. Esta herramienta, que utiliza la base de datos plsdb como referencia, identificó la presencia de 20



plásmidos diferentes en el *dataset* con coberturas iguales o mayores al 30 %. Sin embargo, no se encontraron diferencias en los perfiles de presencia/ausencia de estos plásmidos, con la excepción de NZ_CP059436.1, que forma parte del genoma de una cepa de origen argentino publicado en 2021, que se encontró únicamente en los linajes sudamericanos, con frecuencias de 45 y 92 % en SA1 y SA2 respectivamente, resultado que no concuerda con el obtenido mediante mapeo de *reads*, para el que observó la presencia del plásmido en la mayoría de los genomas sudamericanos y en EU2, pero en ninguno de los genomas del sublinaje EU1.

Búsqueda de secuencias de inserción

La secuencia de inserción ISCfe1 fue identificada mediante BLASTn en 94 de los 103 genomas analizados. En 87 de estos, ISCfe1 fue encontrado en copia única, y en los 7 genomas restantes se encontraron entre 2 y 8 secuencias. La clusterización de las 120 secuencias identificadas dio como resultado la formación de 3 *clusters*: El clústeres más numeroso abarca las secuencias encontradas en los genomas pertenecientes a los sublinajes sudamericanos y a EU1. El segundo clústeres está formado exclusivamente por secuencias del sublinaje EU2, mientras que el tercero está conformado por una única secuencia identificada en un genoma del sublinaje SA1. En la Figura 5.34 se representan estos resultados en un *heatmap*.

En una segunda estrategia de búsqueda de ISCfe1 se realizó el mapeo de *reads* contra la secuencia del ISCfe1 (AM260752.1). De esta manera se identificó la presencia de ISCfe1 en todos los *sets* de datos de secuenciación masiva analizados. Se determinaron los valores de RPKM como medida de abundancia del elemento de inserción, los cuales se muestran en la Figura 5.35. Allí puede observarse un aumento significativo en los genomas sudamericanos con respecto a los europeos (test de Wilcoxon con ajuste de Bonferroni: $p=6,88 \times 10^{-10}$).



clústeres obtenido con CD-HIT.

Identificación de sistemas CRISPR-Cas

Los sistemas CRISPR-Cas han evolucionado como mecanismos de defensa para limitar el impacto de los elementos genéticos extraños encontrados en el ambiente o liberados por fagos u otras bacterias. Se analizaron los genomas de C. fetus bovinos ST4 para buscar sistemas CRISPR-Cas, evaluando el número, el tipo y el subtipo de sistemas encontrados en cada genoma. En todos los genomas se encontró al menos un array CRISPR. El número de arrays CRISPR fue leve pero significativamente superior en los genomas sudamericanos que



en los europeos (Test de Wilcoxon con ajuste de Bonferroni: $4,09 \times 10^{-5}$, como se observa en el *boxplot* de la Figura 5.36. Sin embargo, el número de operones de genes *cas* entre los genomas sudamericanos y europeos no arrojó diferencias significativas. Si bien para la formación de un sistema CRISPR-Cas se requiere la presencia de un *array* CRISPR cercano a genes *cas*, es usual encontrar sólo un cluster de genes *cas* asociados a uno o dos CRISPRs con repetidos idénticos [234]. Es posible que por esta razón se haya encontrado un menor número de clústeres de genes *cas*. No se encontraron genes *cas* en 13 de los genomas analizados, de los cuales 10 pertenecen a aislamientos españoles (EU1), los cuales parecen pertenecer a un grupo monofilético, según la filogenia del linaje L1 mostrada junto al *heatmap* de la Figura 5.37. En todos los casos el tipo de clústeres Cas encontrado fue el IIID.





5.5. Discusión

En este capítulo se analizan los genomas de aislamientos de *C. fetus* de origen bovino pertenecientes al linaje L1, descrito en el capítulo anterior. Con el fin de homogeneizar el *dataset*, se retuvieron únicamente los genomas ST4, ya que este es el ST característico de los aislamientos bovinos [90]. Así es que el *dataset* final comprende 103 secuencias genómicas de *C. fetus*, provenientes mayoritariamente de Europa y América del Sur, siendo España y Argentina los países más representados, con lo que se mantiene el sesgo geográfico discutido anteriormente (Sección 4.5).

L1 presenta una estructura poblacional interna definida, donde los linajes están asociados al origen geográfico de los aislamientos

Si bien la estructura poblacional de *C. fetus* es abordada en diferentes estudios [154, 195, 235], la estructura interna del linaje bovino sólo fue analizada recientemente [236, 237]. En este trabajo se identificó una estructura poblacional definida, formada por cuatro grupos o linajes, respaldada mediante diferentes metodologías dependientes y no dependientes de filogenia. Además, se encontraron evidencias de asociación filogeográfica, sustentadas mediante un análisis bayesiano de significancia de *tips*, identificándose un grupo de clados enriquecidos en genomas de origen europeo (EU1 y EU2) y otro grupo con mayoría de genomas de origen sudamericano (SA1 y SA2). A su vez se observó una aparente asociación directa entre sublinajes y países, ya que la mayoría de los genomas de los sublinajes EU2, SA1 y SA2 provienen de España, Argentina y Uruguay respectivamente.

La correspondencia filogeográfica encontrada no constituye una novedad en el sentido de que se trata de un fenómeno común en procariotas. Por ejemplo, en *C. jejuni* el linaje ST-474 es responsable de la mayoría de los casos clínicos de gastroenteritis en Nueva Zelanda, sin embargo a nivel internacional se trata de un ST raro [238]. En cuanto a *C. fetus*, no se han reportado asociaciones filogeográficas a nivel de los grandes linajes L1 a L8. Sin embargo, en un trabajo reciente de colegas argentinos se declara la existencia de "ciertas asociaciones geográficas en el linaje bovino" [236], en consonancia con lo reportado aquí. En dicho trabajo se analiza la estructura poblacional de un *set* de 34 aislamientos de *C. fetus* provenientes de diferentes regiones de Argentina, incorporándolo a un *set* de datos global, situando la mayoría de los aislamientos en clados regionales, y atribuyendo esta observación al movimiento de animales dado por el comercio regional de ganado en pie.

C. fetus se disemina en América del Sur desde Europa, probablemente durante la colonización del continente

La estructura poblacional encontrada en el linaje bovino de *C. fetus*, sugiere una adaptación transcontinental, pero surge la duda sobre la situación temporal de esta divergencia. Los análisis bayesianos realizados para obtener una filogenia datada dieron como resultado que la divergencia entre los linajes europeos y sudamericanos ocurrió hace 650 años aproximadamente. Además, estos resultados sugieren que los linajes europeos son más antiguos, resultado respaldado por las distribuciones de los índices de diversidad nucleotídica pareados al contrastar EU y SA.

En el trabajo de Iraola et al. [154] se propone que el origen del linaje bovino ocurrió hace aproximadamente 10.500 años, época en la que el hombre comenzaba la domesticación del ganado. De forma análoga, podemos relacionar los resultados de los análisis filogenéticos del linaje bovino con la historia del ganado en América, el cual no es nativo, ya que fue introducido a este continente durante la colonización, a partir del siglo XVI.

Se ha reportado que la naturaleza clonal de *C. fetus* dificulta el datado de la filogenia con un grado de exactitud aceptable [90, 239]. En este trabajo la baja señal temporal es causante de grandes varianzas en los tMRCA obtenidos, con rangos de HPD de hasta 1300 años, como es el caso del nodo que da origen al linaje L1, que son consistentes con los obtenidos en el trabajo de Iraola et al. [154]. Sin embargo, considerando los resultados obtenidos en conjunto con la historia del ganado bovino en América resulta favorable un escenario en el que, con la incorporación del ganado al continente, los linajes sudamericanos de *C. fetus* comenzaran a evolucionar de forma divergente, en respuesta a un nuevo ambiente.

5.5.1. Rasgos genómicos diferenciales entre sublinajes sugieren procesos de adaptación

Campylobacter es un género que ha evolucionado para adaptarse a diferentes hospederos [240, 241]. Frente a los cambios genéticos producidos en respuesta a la adaptación a hospedero, la señal filogeográfica resulta muy débil, sin embargo, al remover la variable hospedero ha sido posible identificar alopatría en genomas de *Campylobacter* [242]. Este fenómeno resulta evidente en los genomas de *C. fetus* ST4 de origen bovino abordados en este capítulo.

El análisis del pangenoma permite diferenciar entre aquellos genes que cumplen funciones *housekeeping*, encontrados en todos o en la gran mayoría de los genomas de una población y los genes accesorios, que proveen funciones adaptativas. Su estudio detectó diferencias en los patrones de presencia/ausencia y en los perfiles funcionales entre los sublinajes individuales y entre los grupos de sublinajes europeos y sudamericanos:

- Se demostró que los sublinajes tienen perfiles funcionales diferentes, siendo U (transporte y secreción) y L (replicación, reparación, recombinación) las principales categorías discriminantes.
- Se identificaron genes particulares cuya presencia o ausencia discrimina entre sublinajes o grupos de sublinajes, resultando llamativa la alta frecuencia de genes derivados de fagos o relacionados al procesamiento de la información genética presentes exclusivamente en los genomas europeos.

Si bien se encontraron diferencias significativas en el número de genes accesorios por genoma para los grupos SA-EU, estas se dan debido a un gran número de genes accesorios presentes en los genomas del sublinaje SA1. Éste es un sublinaje compuesto casi exclusivamente por aislamientos argentinos. Teniendo en cuenta la extensión geográfica de este país y el hecho de que la ganadería es una actividad que se lleva a cabo en todo su territorio, es esperable la presencia de un mayor número de genes accesorios debido a la diversidad de ambientes en la que ocurre esta actividad. Por el otro lado, el sublinaje con el menor número de genes accesorios por genoma es SA2, compuesto mayoritariamente por aislamientos uruguayos. Una variable que podría aportar a la diferencia observada radica en las razas utilizadas en la actividad ganadera en ambos países, debido a que mientras que en Uruguay más del 90 % del ganado corresponde a las razas Hereford y Aberdeen Angus [243], en Argentina coexisten más de 10 razas diferentes [244].

Se evaluaron genes particulares por su relevancia en la interacción hospedero-patógeno Algunos genes relevantes por su rol en la interacción con el hospedero fueron abordados de forma particular. Esto permitió identificar la ausencia del sistema del secreción T5SSb y la toxina CdtB en todos o algunos de los genomas pertenecientes al sublinaje EU2, debido a la presencia de mutaciones sin sentido que producen proteínas truncas. El gen *feoB*, que codifica parte del sistema de adquisición de hierro FeoAB, con gran afinidad por el Fe (II), también fue encontrado como una variante pseudogenizada en los genomas europeos. Se ha reportado que una cepa de *C. jejuni* mutante en *feoB* presenta una menor capacidad para colonizar y sobrevivir en células intestinales de conejo [245]. En conjunto, estas variantes pueden indicar una menor virulencia en las cepas de los sublinajes europeos y en EU2 en particular, pero son necesarios experimentos *in vitro* o *in vivo* para confirmarlo.

Las proteínas Fic están presentes en todos los genomas analizados, pero se observa un enriquecimiento en los genomas europeos Las proteínas Fic actúan como sistemas toxina-antitoxina [203]. En estos el componente toxina es estable, mientras que la antitoxina es muy lábil, por lo que debe ser expresado de forma constitutiva. En ausencia de la antitoxina y ante una señal de estrés ambiental las toxinas actúan modificando las proteínas mediante la adición de AMP. Esta respuesta inhibe la replicación del ADN y la división celular, produciendo un fenotipo filamentoso característico que es revertido al finalizar la señal de estrés. Esta respuesta puede afectar a la misma bacteria fomentando un estado persistente, cuya importancia epidemiológica fue demostrada en *C. jejuni* [246], o bien, al ser secretadas mediante sistemas de secreción de tipo 3 o 4 pueden actuar sobre otras bacterias o sobre células del hospedero e inhibir su ciclo celular [247], por lo que constituyen importantes factores de virulencia de bacterias patógenas. Se describió la presencia de Fic como sistemas T-AT funcionales en *C. fetus* [203], y en este trabajo fue posible detectar proteínas Fic en todos los genomas estudiados, sin embargo se encontró un mayor número de proteínas Fic en los genomas europeos que en los sudamericanos, además de constatarse diferencias en las clases de Fic entre sublinajes. La redundancia de los sistemas T-AT está asociada a mayores tiempos de sobrevida de un patógeno en el hospedero [203], por lo que se puede inferir una mayor persistencia de las infecciones causadas por cepas de los sublinajes europeos, pero para profundizar esta hipótesis sería necesario evaluar y comparar la persistencia de la infección en animales de diferente origen.

Evidencias de mayor incidencia de recombinación en sublinajes sudamericanos y aislamientos de un sublinaje europeo Unos de los rasgos más notables al comparar los genomas pertenecientes a los sublinajes sudamericanos con los europeos ha sido la gran diferencia en el impacto de la recombinación en favor de los sublinajes sudamericanos, al encontrarse no sólo una mayor proporción de bases con evidencia de recombinación, sino también la presencia de dos bloques recombinantes exclusivos de estos sublinajes en una región que ha sido reportada como una isla génica [235], que contiene genes importantes en la interacción con el hospedero como *sap*, importante factor de virulencia debido a su función en la evasión del sistema inmune [248], y *panB*, un factor de especificidad de hospedero [249].

Por otro lado, la ausencia de bloques recombinantes compartidos entre EU2 y los otros sublinajes trazan una aparente barrera de recombinación que aislarían al sublinaje, cuyas causas aún no han sido estudiadas.

Incidencia de elementos móviles variable Los diferentes elementos móviles son importantes contribuyentes en la diferenciación de los linajes, ya

que constituyen mecanismos rápidos de evolución, que le permiten al organismo afrontar con éxito cambios ambientales desfavorables. Por esto, se estudiaron la distribución del elemento de inserción ISCfe1 y la presencia de diferentes plásmidos reportados en genomas de *C. fetus* y bacterias relacionadas, además de los sistemas CRISPR-Cas, que limitan la incidencia de los elementos génicos foráneos.

Las mayores profundidades de mapeo indican una mayor incidencia de ISCfe1 en los genomas de los sublinajes sudamericanos. Este elemento de inserción, compuesto por los genes *tnpA* y *tnpB*, fue descrito como blanco molecular para el diagnóstico de CGB [99], debido a su presencia exclusiva en cepas de Cfv y Cfvi. En este trabajo se identificó la presencia de ISCfe1 en 87 genomas al utilizarse BLASTn como herramienta de búsqueda, mientras que al utilizar una estrategia basada en mapeo de *reads* fue posible identificar este elemento de inserción en todos los genomas bovinos ST4 analizados. La imposibilidad de recuperar a ISCfe1 mediante BLASTn en todos los genomas se debe probablemente al grado de fragmentación de los genomas [186], debido a la dificultad que presentan las regiones de baja complejidad durante el ensamblado a partir de reads cortos. El análisis de secuencia permitió distinguir tres tipos diferentes de ISCfe1, uno de los cuales además de ser exclusivo, es el único presente en el sublinaje EU2. En cuanto al número de parálogos del elemento de inserción, según los resultados de mapeo se observa una mayor incidencia en los sublinajes sudamericanos, lo que podría traducirse en una mayor plasticidad de estos genomas, hipótesis que contrasta con la presencia de grupos de ortólogos de origen fágico exclusivos de sublinajes europeos.

Fic y otros sistemas T-AT encontrados en plásmidos identificados sólo en genomas europeos. Existen diferentes estrategias para la obtención de

secuencias plasmídicas, aplicables en diferentes fases de la obtención de los genomas. Por ejemplo, es posible inferir *contigs* plasmídicos a partir de grafos de ensamblado [250], pero para realizar este tipo de abordaje es necesario contar con los datos de secuenciación masiva, que no siempre se encuentran disponibles. Otras herramientas como PlasmidFinder [251] permiten clasificar los *contigs* según su origen más probable (cromosómico o plasmídico), basándose en la detección del replicón, región que contiene elementos que controlan la replicación [252]. Estas herramientas, sin embargo, tienen la limitación de que sólo pueden detectar los *contigs* que contienen replicones, por lo que su aplicabilidad en genomas fragmentados, que representan la inmensa mayoría de los datos genómicos disponibles en bases de datos públicas es limitada. Por esta razón, se adoptaron dos estrategias que se basan en la detección de secuencias plasmídicas de referencia extraídas de genomas completos de *C. fetus* y otras especies relacionadas, que se encuentran disponibles y públicas en bases de datos genómicos.

Este abordaje permitió inferir la presencia de varios plásmidos, identificándose a dos de ellos muy similares entre sí (NZ_CP059442.1 y NZ_CP059440.1, ver Figura 8.2) que parecen separar los sublinajes sudamericanos de los europeos. Estos plásmidos contienen, además de los genes que codifican la maquinaria de conjugación genes Fic y otros dos sistemas T-AT.

Es frecuente encontrar sistemas T-AT en elementos móviles como plásmidos, donde cumplen la función de asegurar que sólo las células hijas que contengan el plásmido sobrevivan luego de la división celular, mediante un proceso conocido como *Postsegregational killing* o PSK por sus siglas en inglés. Durante este proceso, la toxina estable presente en el citoplasma celular actúa sin la inhibición dada por la antitoxina, induciendo cambios a nivel celular que llevan a la inhibición del crecimiento y a la muerte celular [247]. El plásmido pCFV_08A948_P1 (accession NZ_CP059442.1), detectado en todos los genomas de los sublinajes europeos y sólo en 5 genomas SA1, se caracteriza por contener, además de Fic, los sistemas T-AT YAfQ y VapC (ver Figura 8.2). YafQ es una endonucleasa de ARNm que degrada los transcriptos, evitando así su traducción e induciendo un estado de detención del crecimiento [233], mientras que VapC actúa inhibiendo la traducción al romper el anticodón de iniciación, y se ha demostrado que contribuye en la colonización de las células del hospedero y que está involucrada en el mantenimiento de un plásmido de virulencia en *Salmonella enterica* [253]. La ausencia de este plásmido en los genomas europeos no implica que estos sistemas no se encuentren codificados en el cromosoma. Sin duda, los sistemas T-AT son importantes en la interacción con el hospedero, ya que su presencia puede significar una ventaja durante el proceso infectivo, como se ha reportado en otras especies. Sin embargo, es necesario un análisis dedicado a éstos para evaluar su relevancia y funcionalidad en *C. fetus*.

Mayor número de *arrays* CRISPR en sublinajes sudamericanos Los sistemas CRISPR-Cas actúan como defensas contra la incorporación desmedida de elementos genéticos móviles. La detección de estos sistemas se realiza mediante la identificación de los *arrays* CRISPR, tomando como punto de partida la búsqueda de secuencias repetidas, y paralelamente, la identificación de los genes *cas* mediante estrategias basadas homología de secuencia. Si bien no se encontraron diferencias en cuanto al número de genes *cas*, a pesar de la ausencia de estos genes en un clado de EU1, sí se observó una leve pero significativa diferencia en el número de *arrays* CRISPR en favor de los sublinajes sudamericanos. Adicionalmente, se detectaron menos *arrays* CRISPR por genoma en EU2. Estas observaciones podrían indicar una mayor predisposición en los genomas de los sublinajes europeos para adquirir material genético mediante THG, sin embargo, la menor incidencia de elementos de inserción ISCfe1 detectada en los sublinajes europeos no es compatible con esta hipótesis. Típicamente, los sistemas CRISPR-Cas contienen el cluster de genes *cas* próximos al *array* CRISPR, a una distancia no mayor a 10.000 pb de distancia, aunque se ha reportado que los sistemas CRISPR-Cas del tipo III pueden funcionar con lo genes *cas1-cas2* en *trans* [254]. Más allá de esto, es posible que CRISPR-Cas en *C. fetus* cumpla funciones de regulación de la expresión génica, debido que se han reportado en una cepa de *C. fetus* secuencias espaciadoras con identidad de secuencia significativa a un ORF propio [239]. Sería necesario comprobar la expresión de los ARNcr para determinar la funcionalidad de los sistemas hallados e identificar las secuencias espaciadoras para determinar el blanco de su interferencia. Por otro lado, para tener una visión más completa sobre el papel de los sistemas de defensa habría que comprobar el rol de los sistemas de restricción/modificación, que no fueron abordados en este trabajo.

EU2 es un sublinaje divergente

Tanto en los análisis filogenéticos como en los perfiles genéticos se observó a EU2 separado de los demás sublinajes, con distancias pareadas inter-linaje significativamente superiores (ver Figura 8.1), y una menor diversidad interna, a lo que se suma una aparente barrera de recombinación y una baja proporción de genes accesorios. Este sublinaje, formado casi exclusivamente por genomas de origen español, presenta un mayor número de proteínas Fic, que a menudo se encuentran asociados a elementos móviles y un menor número de sistemas de secreción dedicados a la exportación de proteínas efectoras y a la conjugación, además de variantes pseudogenizadas en los genes que codifican la toxina CdtB y el sistema de secreción T5bSS. EU2 a su vez tiene menos elementos de inserción ISCfe1 por genoma, y estos tienen una secuencia diferente a la de los demás sublinajes. Es posible que las diferencias que se están observando se deban a un proceso evolutivo que resulta en la selección de cepas menos virulentas, en el que los mecanismos de THG no serían los más contribuyentes. El artículo recientemente publicado por Pena-Fernández et al. describe un clado español de genomas pertenecientes al biovar intermedio Cfvi, reportando una proporción significativamente mayor de genes con funciones relacionadas a la infección, incluyendo tres alelos del gen *cdtB* [237]. Sin embargo, debido a que no se asegura la integridad de los mismos, no se considera la posibilidad de que estos se encuentren pseudogenizados, como se demostró en este trabajo.

5.6. Conclusiones

En este capítulo se describe la estructura poblacional del clado bovino de *C*. *fetus*, restringiendo el *set* de datos al ST predominante en este hospedero. Se encontró una estructura definida, compuesta por cuatro linajes asociados a diferentes regiones geográficas, con dos linajes de origen predominantemente europeo y dos asociados a América del Sur, cuya divergencia data de hace 650 años aproximadamente, época de introducción del ganado al continente Americano.

Los genomas de los sublinajes sudamericanos parecen ser más plásticos, debido a la mayor presencia de elementos de inserción y regiones con evidencia de recombinación, características compatibles con la expansión del genoma accesorio en respuesta a la adaptación al nuevo ambiente. Por otro lado, el sublinaje EU2, compuesto por genomas de origen español en su mayoría, se encontró aislado de los demás genomas estudiados por una aparente barrera de recombinación, además de ser tanto filogenética como funcionalmente distante.

6

Estudio de una serie de casos clínicos de *Campylobacter fetus* ocurridos en Montevideo entre 2013 y 2018

6.1. Introducción

Las infecciones por *C. fetus* en humanos causan mayoritariamente enfermedad sistémica, típicamente en individuos con otras enfermedades subyacentes. Algunos factores predisponentes van desde la inmunosupresión, causada por el virus de inmunodeficiencia humana o por cáncer hematológico, diabetes, cirrosis, embarazo o edad avanzada. La principal ruta de transmisión a humanos es mediante el consumo de productos de origen animal contaminados, o el contacto con animales de granja [73]. Sin embargo, la microbiota podría actuar como reservorio de cepas de *C. fetus* patobiontes en individuos sanos, facilitando la diseminación mediante transmisión entre personas [154]. Además, se ha reportado la transmisión de persona a persona en hombres que tienen sexo con hombres [72], y casos de meningitis en recién nacidos sugieren la transmisión desde adultos u otros recién nacidos [255, 256].

6.2. Antecedentes

En este capítulo se realiza el estudio retrospectivo de una serie de casos de infección por *C. fetus* ocurridos en 3 hospitales de la ciudad de Montevideo, en el período comprendido entre junio de 2013 y junio de 2018. A continuación, se describe brevemente cada uno de estos casos en orden cronológico.

6.2.1. Descripción de los casos

- HB1: En julio de 2013 en el hospital H1 se reportó el primer caso en un trabajador rural de 64 años, portador de un linfoma no Hodgkin en células del manto. El paciente recibía tratamiento oncológico consistente en una combinación de quimioterapia R-CHOP y Bendamustina-Rituximab.
- HC1: Esta cepa fue aislada en febrero de 2014 a partir del hemocultivo de un paciente de 58 años con desórdenes psiquiátricos que vivía en la indigencia y sufría de cáncer de próstata y fue ingresado al hospital H2 con fiebre.
- HC70-b y HC70-c: En junio de 2014 una mujer de 39 años que transcurría la semana 26 del embarazo fue admitida al hospital H2 con un diagnóstico de meningitis. Se aislaron las cepas HC70-b y HC70-c de sangre y líquido cerebroespinal respectivamente. Se le practicó una cesárea de urgencia al constatar sufrimiento fetal. El neonato fue diagnosticado con sepsis y sobrevivió luego del tratamiento con meropenem. Se había constatado antecedentes de consumo de carne cruda en la madre.

- HC71: El cuarto caso de esta serie se registró en setiembre de 2015 en el hospital H2. La cepa HC71 se aisló de un cultivo de fluido de ascitis obtenido a partir de un paciente de 48 años con cirrosis.
- HC73: Dos meses después del caso anterior se registró un nuevo caso en el mismo hospital (H2) en una paciente de 89 años portadora de un mieloma múltiple admitida por un cuadro febril. El aislamiento se realizó a partir del hemocultivo.
- **HB74:** Esta cepa se aisló a partir del hemocultivo de una mujer de 59 años en el hospital H1, también en diciembre de 2015.
- HM1: Luego de más de 2 años en los que no se reportaron casos de infección por *C. fetus*, en junio de 2018 se aisla la cepa HM1 a partir de hemocultivos de un paciente de 64 años admitido al hospital H3 con diagnóstico de aneurisma de aorta.

6.3. Metodología

6.3.1. Obtención de secuencias genómicas

La identificación de los aislamientos a nivel de género se realizó en laboratorios del Hospital Maciel y del Instituto de Higiene (Universidad de la República), utilizando los métodos de rutina implementados por cada nosocomio, como BactAlert, VITEK2 y MALDI-TOF. A nivel de especie, la identificación se hizo mediante secuenciación parcial del gen 16S y dos estrategias de PCR: la primera estrategia, implementada por Iraola et al. [100] consiste en una PCR multiplex, donde se utiliza el par de cebadores MG3F y MG4F para amplificar una región de 764 pb del gen *cstA* y otro par de cebadores diseñados por Moolhuijzen et al. [101] que amplifica 233 pb de *virB*, cuya presencia había sido reportada únicamente en la subespecie Cfv. La segunda estrategia empleada consiste en un método de PCR a tiempo real que amplifica una región del gen 16S y permite la identificación de aislamientos de *C. fetus* a nivel de especie [102].

La secuenciación y el ensamblado de los genomas de los aislamientos identificados se describe en el Capítulo 3.

6.3.2. Análisis bioinformáticos

MLST

Se realizó una tipificación de los aislamientos mediante MLST. Para esto se utilizó el esquema de alelos disponible en la base de datos PubMLST [92, 257] correspondiente a *C. fetus* [161] y MLSTar, un pipeline desarrollado en nuestro laboratorio, que devuelve los ST a los que pertenece cada genoma [162]. Para esto se utilizó la función MLSTar::doMLST, que extrae las secuencias de los alelos descriptos en la base de datos correspondiente mediante BLASTn.

Obtención del marco clonal

Se infirió a partir del alineamiento de genomas de los 8 aislamientos analizados en este capítulo. Para esto, en primer lugar se obtuvo el genoma *core* utilizando Harvest Suite [163]. Se obtuvo el marco clonal a partir del alineamiento resultante utilizando Gubbins [164].

Filogenias

Se realizaron utilizando el método de *Maximum Likelihood* y el modelo GTRCAT en RaxML [165], a partir del alineamiento de secuencias nucleotídicas correspondientes a los genes del MLST concatenados o al marco clonal. La visualización de los árboles filogenéticos resultantes se realizó utilizando el paquete ggtree [167] en R.

6.4. Resultados

6.4.1. Análisis de MLST de los aislamientos

A partir de las secuencias genómicas obtenidas de los 8 aislamientos descriptos se ejecutó MLSTar para su tipificación, encontrándose que ST4 es el genotipo mayoritario. En la Tabla 6.1 se detallan los resultados de tipificación obtenidos.

6.4.2. Predicción del genoma *core* y el marco clonal

El genoma *core*, deducido a partir del alineamiento genómico de los ocho aislamientos clínicos uruguayos consta de 1.634.667 posiciones, esto es 89 % del genoma más largo. La búsqueda de regiones recombinantes realizada mediante Gubbins dio como resultado la identificación de múltiples bloques recombinantes, que pueden verse en la Figura 6.1. Allí pueden identificarse bloques recombinantes ancestrales compartidos únicamente por los genomas de los aislamientos HB74 y HC71, y otros bloques recombinantes más

Aislamiento	aspA	glnA	gltA	glyA	pgm	tkt	uncA	ST
HB1	1	2	2	2	1	2	1	4
HC1	1	2	2	2	1	2	1	4
HC70-bc	1	2	2	2	1	2	1	4
HC71	1	5	4	2	1	7	9	20
HC73	1	2	2	2	1	2	1	4
HB74	1	1	1	1	1	1	1	1
HM1	1	2	2	2	1	2	4	6

Tabla 6.1: Resultados de la tipificación de los aislamientos mediante MLST

recientes en el genoma de la cepa M1. La escisión de los bloques recombinantes detectados del genoma *core* da lugar a la obtención del marco clonal, reteniéndose en total 5305 posiciones polimórficas a partir de las cuales se construyó la filogenia.

6.4.3. Filogenias y recombinación

La filogenia global mostrada en la Figura 4.7, obtenida a partir del marco clonal de un conjunto de datos compuesto por 219 aislamientos, nos permite obtener la posición de los genomas uruguayos en la estructura poblacional global de C. fetus no testudinum, la cual se muestra en la Fig. 1 de la pág. 2 de "Polyclonal Campylobacter fetus infections among unrelated patients Montevideo Uruguay 2013–2018" [258], donde se observa que los aislamientos clínicos no forman un grupo monofilético, sino que se distribuyen en tres linajes diferentes: HC71 y HB74 son los más alejados del grupo, encontrándose en los linajes L7 y L5 respectivamente, mientras que el resto de los aislamientos forma un grupo monofilético en el linaje L2. Paralelamente, en la filogenia construida a partir del alineamiento genómico de los aislamientos clínicos uruguayos, se observa que el grupo monofilético que se agrupó dentro del linaje L2 en la filogenia global se separan los genomas de acuerdo a su ST (Figura 6.1). Se encontraron bloques recombinantes exclusivos en los genomas de los aislamientos HM1 y HC71. A su vez, se encontró recombinación entre los genomas HB74 y HC71, que podría explicarse debido a la co-ocurrencia temporal de ambas cepas.

6.4.4. Distancias entre los aislamientos

Para determinar el grado de similitud entre los aislamientos clínicos uruguayos se extrajo el número de SNPs que separa cada aislamiento de forma



pareada. A partir del alineamiento del genoma *core* se obtuvieron las distancias genéticas mostradas en la parte B de la Fig. 1 de la pág. 2 en el artículo. Aquí se reflejan las filogenias obtenidas previamente, donde los genomas que forman el grupo monofilético en L2 en la filogenia contienen muy poca variabilidad (0-15 SNPs), y el aislamiento HM1 se encuentra algo más alejado. HB74 y HC71 están separados entre sí por una distancia mayor a la que separa a cada uno de ellos con los aislamientos del linaje L2.

	HC70-c	НС70-Ъ	HC1	HB1	HC73	HM1	HB74	HC71
HC70-c	0	0	0	14	15	388	2065	2531
HC70-b		0	0	14	15	388	2065	2531
HC1			0	14	15	388	2065	2531
HB1				0	13	386	2063	2529
HC73					0	387	2064	2530
HM1						0	2092	2561
HB74							0	2795
HC71								0

 Tabla 6.2: Distancias genómicas entre los aislamientos, expresadas en SNPs, obtenidas a partir del alineamiento del marco clonal.

6.5. Discusión

En este trabajo reportamos una serie de casos clínicos de *C. fetus* que sucedieron en un período de tiempo de 5 años, entre junio de 2013 y junio de 2018, en tres hospitales de la ciudad de Montevideo. En los 5 años anteriores a dicho período no se habían reportado casos vinculados en *C. fetus* en ninguno de estos hospitales, como tampoco se reportan cambios en las estrategias de diagnóstico utilizadas. Estos hechos sugirieron que los casos aquí reportados formaban parte de un brote de *C. fetus* que afectó a pacientes con antecedentes de inmunosupresión o condiciones patológicas preexistentes. Para esclarecer si estos casos fueron efectivamente parte de un brote, se secuenciaron los genomas y se realizaron estudios comparativos.

Se encontró que 5 de los 8 aislamientos forman un grupo perteneciente al linaje L2, donde todos presentan genotipo ST4, el cual se asocia mayoritariamente a bovinos. Las distancias genéticas entre los aislamientos de este grupo fueron muy bajas, menores a 15 SNPs, por lo cuál podría tratarse de un mismo clon. No se encontró evidencia que relacione entre sí a los pacientes infectados con este clon, ni que estuvieran expuestos a una fuente de contaminación en común. Esto, además del hecho de que los aislamientos ocurrieran en un período de tiempo prolongado, sugiere que estos pacientes fueron infectados por vías independientes. La fuente de infección podría haber sido de origen zoonótico, la vía más frecuentemente reportada [73], aunque no puede descartarse la transmisión humano – humano, como ya ha sido reportado anteriormente [72], y tratarse de un clon que circula en la población de forma asintomática. Esta hipótesis se ve favorecida por la identificación de *C. fetus* en microbiomas intestinales humanos pertenecientes a individuos sanos [154]. HM1, el aislamiento más cercano a este grupo, que también pertenece al linaje L2, está separado de los aislamientos del grupo por 390 SNPs aproximadamente, pero fue obtenido en un tercer hospital en el año 2018. Tampoco pudo identificarse un origen común con los aislamientos del grupo. Los aislamientos HC71 y HB74, muy divergentes entre sí y con los demás aislamientos, pertenecen a los linajes L5 y L7 respectivamente. HC71 pertenece a un linaje compuesto exclusivamente por aislamientos obtenidos a partir de sangre de humanos en Taiwan. Sin embargo, estos aislamientos 17144_S83, obtenido a partir de un feto ovino, aislado por el INIA La Estanzuela. Dada la pequeña diferencia que separa a HB74 y 17144_S83 (24 SNPs), podríamos estar hablando de un mismo clon. Sin embargo, al no contar con información que relacione ambos aislamientos, más allá de lo genético, no es posible trazar vías de transmisión y/o contaminación.

Cabe recordar que el ST4 es el predominante en aislamientos de *C. fetus* de origen bovino. La cercanía de estos aislamientos a los grupos de origen bovino podría estar evidenciando zoonosis reciente, en la que los individuos adquirieron la infección debido al contacto con animales infectados, aunque la vía de transmisión zoonótica sólo se confirmó en dos casos.

6.6. Conclusiones

La serie de casos reportada en este trabajo corresponde a aislamientos policlonales, sin evidencia de una fuente común de contaminación, o transmisión. Dada la despreciable distancia evolutiva que separa a los aislamientos del grupo compuesto por HB1, HC1, HC70 (c y b) y HC73 en contraposición a la temporalidad de los casos, se sugiere la hipótesis de que se originan a partir de un mismo clon que circula por la población de forma asintomática, que fuera recientemente adquirida mediante zoonosis. Este adquiere características de gravedad dada la presencia de patologías o condiciones preexistentes en los pacientes. Los aislamientos divergentes HB74 y HC71 tienen orígenes independientes.

El trabajo que se describe en este capítulo derivó en la publicación del artículo "Polyclonal Campylobacter fetus infections among unrelated patients, Montevideo, Uruguay, 2013–2018" [258], que se anexa a continuación.

BRIEF REPORT

Polyclonal *Campylobacter fetus* Infections Among Unrelated Patients, Montevideo, Uruguay, 2013–2018

Daniela Costa,^{1,2} Laura Betancor,³ Pilar Gadea,³⁴ Laura Cabezas,⁴ Leticia Caiata,⁴ Rosario Palacio,⁴ Verónica Seija,⁴ Antonio Galiana,⁵ Mariela Vieytes,⁵ Inés Cristophersen,⁵ Lucía Calleros,² and Gregorio Iraola^{1,6,7}

¹ Microbial Genomics Laboratory, Institut Pasteur de Montevideo, ²Sección Genética Evolutiva, Facultad de Ciencias, ³Instituto de Higiene, Facultad de Medicina, and ⁴Repartición Microbiología, Departamento de Laboratorio de Patología Clínica, Hospital de Clínicas, Facultad de Medicina, Universidad de la República, and ⁶Hospital Maciel, Montevideo, Uruguay, ⁶Center for Integrative Biology, Universidad Mayor, Santiago, Chile; ⁷Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, United Kingdom

In Montevideo (2013-2018), 8 Campylobacter fetus extraintestinal infections were reported. The polyclonal nature of strains revealed by whole-genome sequencing and the apparent lack of epidemiological links was incompatible with a single contamination source, supporting alternative routes of transmission.

Keywords. *Campylobacter fetus*; case series; unrelated patients; pathobiont; whole-genome sequencing.

Campylobacter fetus infections in humans mainly cause invasive illness, typically under the presence of defined underlying diseases. Predisposing factors range from immunosuppression (as caused by human immunodeficiency virus or hematological cancers), diabetes, cirrhosis, pregnancy or advanced age.

The main suspected route of *C. fetus* transmission to humans is the consumption of contaminated animal products or contact with farm animals [1]. However, recent evidence indicates that the gut microbiota could act as a reservoir for pathobiont *C. fetus* strains in healthy individuals, facilitating dissemination through asymptomatic carriage and person-to-person transmission [2]. In addition, person-to-person transmission has been recently proposed in men who have sex with men [3], and 2 previously reported cases of meningitis in newborns are also suggestive of transmission from an adult or another newborn [4, 5]. In the current study, we used whole-genome sequencing to characterize *C. fetus* epidemiologically, given an observed increase in *C. fetus* cases during 2013–2018 in Montevideo, Uruguay.

Clinical Infectious Diseases® 2020;70(6):1236–9

1236 • CID 2020:70 (15 March) • BRIEF REPORT



METHODS

A retrospective, genomic-based study was carried out to investigate the nature of 8 *C. fetus* cases reported by internal surveillance systems in 3 hospitals in Montevideo from June 2013 to June 2018. Remarkably, since the previous 5-year period (2008– 2013) when no *C. fetus* cases were reported, diagnostic tools and reporting procedures were essentially the same in these hospitals. Routinely, infections were initially detected from blood, cerebrospinal fluid and/or ascites fluid cultures using the BactAlert system (bioMérieux) and strains were subsequently identified using the VITEK 2 system (bioMérieux) and/or matrix-assisted laser desorption/ionization time-of-flight mass spectrometry at the Instituto de Higiene (Facultad de Medicina, Universidad de la República, Uruguay) and/or Hospital Maciel, Montevideo, Uruguay.

In all cases, the presence of Campylobacter was confirmed but some strains were not identified at the species level. Next, the Sección Genética Evolutiva (Facultad de Ciencias, Universidad de la República, Uruguay) identified the Campylobacter isolates to the species level using polymerase chain reaction methods, described elsewhere [6, 7], and partial 16S sequencing, confirming the presence of C. fetus in all cases. Subsequently, whole-genome sequencing was performed using an Illumina MiSeq platform at the Institut Pasteur de Montevideo, Uruguay. Sequencing data was processed, assembled, and annotated as described elsewhere [2]. Genome assemblies were deposited in the National Center for Biotechnology Information database (BioProject no. PRJNA543866). Multilocus sequence typing was performed from genomic sequences using MLSTar [8] or as described elsewhere [9]. A core genome phylogeny was constructed using the Harvest software (version 1.0) suite [10], by adding the 8 new genomes to a global data set representing the known genetic diversity of C. fetus [2, 11]. Strain membership in previously described lineages was based on phylogenetic placement within the core genome phylogeny.

RESULTS

From June 2013 to June 2018, there was an increase in patients infected with *C. fetus* (8 patients in all: 1 in 2013, then 3 each in 2014 and 2015, and 1 in 2018), compared with the previous 5-year period from June 2008 to May 2013, when no *C. fetus* infections were reported. The first case was reported in July 2013 (hospital 1). It occurred in a 64-year-old man who had received a combination of R-CHOP chemotherapy (rituximab with cyclophosphamide, doxorubicin, vincristine, and prednisone) and bendamustine-rituximab against a non-Hodgkin mantle cell lymphoma; the isolated strain HB1 (also referred

Received 23 May 2019; editorial decision 8 July 2019; accepted 12 July 2019; published online July 15, 2019.

Correspondence: G. Iraola, Institut Pasteur de Montevideo, Montevideo 11400, Uruguay (giraola@pasteur.edu.uy).

[©] The Author(s) 2019. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com. DOI: 10.1093/cid/ciz657

as H1-UY) was characterized by means of multilocus sequence typing as genotype sequence type (ST) 4 [12].

Then in February 2014, a 58-year-old homeless man with psychiatric disorders and prostate cancer was seen for fever and admitted to hospital 2. Blood cultures were positive for *C. fetus*, and the isolated strain HC1 belonged to the same ST-4 genotype. In June 2014, a 39-year-old woman, who was 26 weeks pregnant and had a history of raw meat consumption, was also admitted to hospital 2 and received a diagnosis of meningitis. *C. fetus* ST-4 was isolated from blood (strain HC70-b) and cerebrospinal fluid (strain HC70-c) cultures. Cesarean delivery was required owing to fetal bradycardia, and neonatal sepsis was diagnosed in the newborn, caused by vertical transmission of the same ST-4 *C. fetus* strain (HC70-nw). The newborn survived after treatment with meropenem.

In September 2015, another case was reported in hospital 2, isolated from ascites fluid in a 48-year-old cirrhotic male patient carrying a ST-20 *C. fetus* strain (HC71). In December



Figure 1. Chronology and genomic relatedness of *Campylobacter fetus* strains. *A*, Chronology of *C. fetus* cases reported in 3 hospitals in Montevideo, Uruguay. Colors represent the multilocus sequence typing (MLST) genotype (blue, sequence type [ST] 4; yellow, ST-20; green, ST-1). *B*, Phylogenetic relationships between strains based on the core genome. MLST genotypes are represented as colored circles. The tabulation on the right shows numbers of single-nucleotide polymorphisms shared between strains; colors indicate whether strains belong to the same or a different phylogenetic lineage. *C*, Phylogenetic relationships of local cluster strains incorporated into a global collection of *C. fetus* strains (phylogenetic lineages as defined by Iraola et al [2]); strains are identified by color, and the specific positions of cluster strains are highlighted with red dots. A detailed phylogenetic tree with tip labels is presented in Supplementary Figure 1.

BRIEF REPORT • CID 2020:70 (15 March) • 1237

2015, yet another case was reported in hospital 2, in a 89-yearold female patient. In this case, a ST-4 *C. fetus* strain (HC73) was isolated from blood cultures. Also in December 2015, but in hospital 1 (where the first case was reported), *C. fetus* ST-1 (strain HB74) was isolated from the blood of a 59-year-old female patient. No cases were reported in 2016 or 2017. In June 2018, *C. fetus* was again isolated in hospital 3 from blood cultures of a 64-year-old male patient with a diagnosis of aortic aneurysm. This strain (HM1) was genotyped as ST-4. The chronology of cases is summarized in Figure 1A, and further details are provided in Table 1.

Except for the mother and her newborn, no epidemiological link was evident between any of the patients. Given this and the observed dominance of the ST-4 but also the presence of other *C. fetus* clones, whole-genome sequences were used to reconstruct the precise phylogenetic relationship among strains. Figure 1B shows the close genetic relationship between ST-4 strains isolated in different periods of time and also in different hospitals. Furthermore, the number of single-nucleotide polymorphisms (SNPs) in the core genome evidenced that strain HM1 is divergent with respect to the other ST-4 strains (Figure 1B). Strains belonging to ST-20 and ST-1 are phylogenetically divergent, as also evidenced by a much higher number of SNPs compared with ST-4 strains.

C. fetus strains have been previously classified into 8 phylogenetic lineages (L1–L8) [2], so we added our strains to the global *C. fetus* phylogeny, observing that they all belonged to 3 previously described lineages frequently reported in humans (Figure 1C). In particular, all ST-4 strains belonged to L2, a worldwide distributed, human-adapted lineage, but were divided into 2 sublineages. Strain HM1 belonged to the major sublineage, and the rest (HB1, HC73, HC1, HC70-b, and HC70-c) clustered in a second, more restricted sublineage. Strains HB1, HC73, HC1, and HC70 (both isolates HC70-b and HC70-c from the same patient) probably represent the same clone circulating in the Uruguayan population, given their extremely close genetic relatedness (0–15 SNPs) (Figure 1B). Interestingly, strain HC71 (ST-20) belonged to lineage 5 (L5) which to date was composed only by strains from Taiwan isolated from humans between 2002 and 2014. Strain HB74 (ST-1) belonged to lineage 8 and was closely related to JP10, a strain isolated from a patient in Japan in 2001 (Supplementary Figure 1). Together, the genomic and phylogenetic analysis of *C. fetus* strains isolated from Uruguayan patients revealed epidemiological complexity, characterized by both very similar and very divergent strains circulating among unrelated patients.

DISCUSSION

Campylobacter fetus infections in humans are sporadically reported, and longitudinal studies are even more rarely described. Indeed, ours is the first *C. fetus* case series reported in humans in South America. Specifically in Uruguay, *C. fetus* had not been formally reported in humans until the appearance of these cases, even though routine detection methods in the 3 hospitals where *C. fetus* was detected have been similar in past years. Accordingly, and beyond our study's retrospective nature, the absence of control groups, and the fragmented clinical data available from our patients, the emergence of these cases is unlikely to be caused by any improvement in diagnostic procedures.

The apparent lack of epidemiological link between patients and the absence of an identified common food source at least suggest independent routes of infection. Indeed, both the presence of almost identical strains circulating in unrelated patients and different hospitals (HC1, HC73, HC70, and HB1), and the appearance of different genotypes co-occurring in time (HB74 and HC73 in December 2015) may be evidence that most patients were independently contaminated from animals or animal products or colonized with strains circulating among other humans.

Interestingly, we have recently suggested that asymptomatic carriage in the gut microbiota of healthy people could constitute a way for *C. fetus* to disseminate among humans [2]. In

Table 1. Description of Campylobacter fetus Strains Isolated From Uruguayan Patients

Strain	Comorbid Condition or Status	Reason for Admission	Source	MLST	Lineage	Hospital	Month and Year of Admission
HC71	Cirrhosis	Peritonitis ^a	Ascites	ST-20	L5	2	September 2015
HB74	None	NA	Blood	ST-1	L8	1	December 2015
HM1	None	Aortic aneurysm	Blood	ST-4	L2.2	3	June 2018
HC73	Multiple myeloma	Febrile illness	Blood	ST-4	L2.1	2	December 2015
HB1	Non-Hodgkin lymphoma	Cellulitis ^a	Blood	ST-4	L2.1	1	July 2013
HC1	Prostate cancer	Pneumonia	Blood	ST-4	L2.1	2	February 2014
HC70-b ^b	Pregnancy	Meningitis ^a	Blood	ST-4	L2.1	2	June 2014
HC70-c ^b	Pregnancy	Meningitis ^a	CSF	ST-4	L2.1	2	June 2014
HC70-nwb	Newborn	Perinatal sensis ^a	Blood	ST-4	ΝΔ	2	July 2014

Abbreviations: CSF, cerebrospinal fluid; MLST, multilocus sequence typing; NA, not available; ST, sequence type.

^aThe reason for admission was confirmed to be caused by *Campylocbacter fetus* infection.

^bThe same strain isolated from the mother and her newborn (HC70-nw was not sequenced).

1238 • CID 2020:70 (15 March) • BRIEF REPORT

this situation, *C. fetus* could act as a pathobiont residing in the gut that arises and causes invasive infections when the immune function is compromised by underlying diseases. In favor of this hypothesis, our previous work [2] identified a strain present in the gut microbiota of healthy individuals, which was then responsible of a *C. fetus* cluster with human-to-human transmission in Canada [3], supporting the idea that *C. fetus* strains may be asymptomatically carried pathobionts that emerge and cause disease in particular conditions.

Our genomic approach uncovered similarities among certain patients that were not apparent in the epidemiological investigation of this case series; however, the drivers for the observed increase in *C. fetus* infections remain uncertain. Our results support the notion that whole-genome sequencing is currently the only approach that allows to precisely characterize *C. fetus* strains, and its incorporation into clinical surveillance is needed to better characterize the nature of these infections. In addition, an assessment of asymptomatic carriage of *C. fetus* in healthy individuals would improve our understanding of its epidemiological characteristics in humans.

Notes

Acknowledgments. The authors thank Arací Martínez for providing assistance with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for strain identification at Instituto de Higiene, Gabriela Algorta for providing some of the *Campylobacter fetus* strains, and Hugo Naya from the Bioinformatics Unit (Institut Pasteur de Montevideo, Uruguay) for providing laboratory space and computational resources during the development of this work.

Financial support. This work was supported by the Agencia Nacional de Investigación e Innovación, Uruguay (grant ANII-FSSA-2014-1-105252 to D. C.).

Potential conflicts of interest. The authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

- 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; 58:1579–86.
- Iraola G, Forster SC, Kumar N, et al. Distinct *Campylobacter fetus* lineages adapted as livestock pathogens and human pathobionts in the intestinal microbiota. Nat Commun 2017; 8:1367.
- Marchand-Senécal X, Bekal S, Pilon PA, Sylvestre JL, Gaudreau C. Campylobacter fetus cluster among men who have sex with men, Montreal, Quebec, Canada, 2014–2016. Clin Infect Dis 2017; 65:1751–53.
- Rennie RP, Strong D, Taylor DE, Salama SM, Davidson C, Tabor H. Campylobacter fetus diarrhea in a Hutterite colony: epidemiological observations and typing of the causative organism. J Clin Microbiol 1994; 32:721–4.
- Morooka T, Umeda A, Fujita M, et al. Epidemiologic application of pulsed-field gel electrophoresis to an outbreak of *Campylobacter fetus* meningitis in a neonatal intensive care unit. Scand J Infect Dis 1996; 28:269–70.
- Iraola G, Hernández M, Calleros L, et al. Application of a multiplex PCR assay for *Campylobacter fetus* detection and subspecies differentiation in uncultured samples of aborted bovine fetuses. J Vet Sci 2012; 13:371–6.
- Iraola G, Pérez R, Betancor L, et al. A novel real-time PCR assay for quantitative detection of *Campylobacter fetus* based on ribosomal sequences. BMC Vet Res 2016; 12:286.
- Ferrés I, Iraola G. MLSTar: automatic multilocus sequence typing of bacterial genomes in R. PeerJ 2018; 6:e5098.
- Calleros L, Betancor L, Iraola G, et al. Assessing the intra-species genetic variability in the clonal pathogen *Campylobacter fetus*: CRISPRs are highly polymorphic DNA markers. J Microbiol Methods 2017; 132:86–94.
- Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid coregenome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol 2014; 15:524.
- van der Graaf-van Bloois L, Duim B, Miller WG, Forbes KJ, Wagenaar JA, Zomer A. Whole genome sequence analysis indicates recent diversification of mammal-associated *Campylobacter fetus* and implicates a genetic factor associated with H2S production. BMC Genomics 2018; 17:713.
- Iraola G, Betancor L, Calleros L, et al. A rural worker infected with a bovineprevalent genotype of *Campylobacter fetus* subsp. *fetus* supports zoonotic transmission and inconsistency of MLST and whole-genome typing. Eur J Clin Microbiol Infect Dis 2015; 34:1593–6.

BRIEF REPORT • CID 2020:70 (15 March) • 1239

Consideraciones finales

Uruguay es un país en el que la ganadería es una de las actividades económicas más importantes, sin embargo, las tasas de procreo han sido bajas. Se considera que las causas de la baja producción son multifactoriales, y *Campylobacter fetus* es uno de los factores infecciosos más prevalentes. *C. fetus* se presenta con baja frecuencia en humanos, con cuadros gastrointestinales, aunque frecuentemente causa infecciones sistémicas, sobre todo en individuos inmunocomprometidos o con condiciones predisponentes.

Actualmente *C. fetus* se diagnostica mediante aislamiento en medios selectivos que inhiben el crecimiento de otras bacterias, y la identificación a nivel de subespecie se realiza mediante métodos bioquímicos.

Históricamente, las subespecies definidas mediante estos métodos están asociadas a dos variantes ecológicas. Se considera que Cfv y Cfvi son los causantes de la CGB, al colonizar el tracto genital de bovinos y generar aborto endémico y baja fertilidad. Por otro lado, Cff es una subespecie más generalista, que infecta, además de humanos, a bovinos y ovinos, en los que produce abortos esporádicos y tiene tropismo por el tracto gastrointestinal. Existe solapamiento a nivel genómico entre ambas subespecies, y hasta ahora no se han encontrado asociaciones entre sus respectivas preferencias ecológicas y las pruebas bioquímicas que las diferencian.

Con la hipótesis de que la baja especificidad y/o sensibilidad de las herramientas moleculares propuestas se debe a que han sido formuladas en base a un bajo número de muestras o a muestras poco diversas, en este trabajo se propuso reunir una gran diversidad genética de *C. fetus*, incluyendo cepas circulantes en Uruguay, para identificar, mediante estrategias basadas en genómica comparativa, características relevantes en la ecología y patogenicidad de esta especie, y diseñar una metodología de diagnóstico de CGB basada en PCR.

El punto de partida para los estudios poblacionales realizados fue la compilación de una base de datos con secuencias genómicas de buena calidad y orígenes geográficos y de hospederos diversos. Para tener en cuenta el factor local se incluyeron genomas provenientes de aislamientos obtenidos en Uruguay, a partir de muestras clínicas y de animales de producción. La disponibilidad de datos genómicos de buena calidad es indispensable para el estudio de bacterias como *C. fetus*, de la cual se dispone de un número limitado de secuencias y la mayoría de ellas se encuentran altamente fragmentadas. Así es que, como parte de este trabajo, se incluyó la obtención de los genomas de 18 aislamientos uruguayos de diferentes hospederos, según se detalla en el Capítulo 3.

Estudios poblacionales de C. fetus

Los estudios poblacionales se plantearon primero desde una perspectiva general, al incluir genomas de diferentes hospederos, para luego focalizar el estudio en el linaje bovino. Por último, se analizaron genomas obtenidos a partir de aislamientos de pacientes tratados en 3 hospitales de Montevideo, en un estudio retrospectivo que abarcó 5 años y fue publicado en 2019 [137].

La metodología utilizada se basó en flujos de trabajo y herramientas bioinformáticas ampliamente utilizadas. A nivel general, los genomas de las cepas de *C. fetus* forman una estructura poblacional en la que se distinguen
diferentes linajes con asociación a hospedero. Los resultados obtenidos son similares a los reportados en otros estudios [154, 186], validando entonces la metodología empleada. Es común encontrar en diferentes especies de *Campylobacter* evidencias de procesos de adaptación a hospedero [242]. Por ejemplo, *C. hyointestinalis*, la especie más cercana a *C. fetus* se divide en dos subespecies, una de ellas adaptada a cerdos y la otra más generalista, que ha sido encontrada en bovinos y otros hospederos.

El linaje bovino como objetivo para diagnóstico de la campylobacteriosis genital bovina

El linaje bovino L1 sobresalió en comparación con otros linajes donde predominan los hospederos humanos, ya que se identificaron diversas características que indican un proceso de adaptación, como el aumento observado tanto en el número de genes accesorios como en la incidencia de la recombinación. En este linaje co-ocurren las dos subespecies de *C. fetus* asociadas a mamíferos y el *biovar* intermedio, mientras que en los linajes humanos la ocurrencia de Cfv y Cfvi es marginal.

En base a estas observaciones y a que no existe un vínculo observable entre los marcadores fenotípicos y las características ecológicas de las subespecies que éstas definen, se puede sugerir una reinterpretación de las subespecies según su descripción clásica, para dar paso a la adopción de metodologías que reflejen las características de los ecotipos a nivel genómico.

La búsqueda de blancos moleculares para el diagnóstico de la CGB pasaría entonces, de la identificación clásica de subespecies a la distinción de los linajes evidenciados en la estructura poblacional obtenida, con el linaje L1 como principal objetivo. Esto derivó en el diseño de cebadores específicos para la identificación de L1, que tienen como blanco una región del primer ORF del elemento de inserción ISCfe1. Si bien la validación *in vitro* de estos cebadores no se completó, los altos niveles de sensibilidad y especificidad logrados en las pruebas *in silico*, junto a la probada eficiencia de herramientas basadas en el mismo blanco, se anticipa su potencial como método de diagnóstico.

Se encontraron evidencias de alopatría transcontinental y aislamiento genético en el linaje bovino

Al focalizar el estudio sobre el linaje bovino se identificó una estructura poblacional con sublinajes asociados a diferentes regiones geográficas. La divergencia entre los sublinajes sudamericanos y los europeos probablemente se originó, según las estimaciones realizadas, hace aproximadamente 650 años, con la introducción del ganado al continente americano. Sin embargo, no pueden descartarse múltiples introducciones a América, así como tampoco la posibilidad de que este continente actúe como fuente de *C. fetus*, al ser movilizada al resto del mundo debido a fallas en los controles que permiten el ingreso de ganado en pie o semen infectado al circuito de comercio internacional.

Los resultados sugieren que los sublinajes sudamericanos serían más plásticos que los europeos, en base a las mayores incidencias observadas en la recombinación e incorporación de elementos móviles, así como una expansión en el número de genes accesorios. A su vez, estas observaciones sugieren un proceso de alopatría transcontinental. Por otro lado, se observó en un sublinaje europeo, otro proceso de aislamiento, evidenciado por la ausencia de bloques recombinantes compartidos con genomas de otros sublinajes, aún cuando estos provienen de la misma región geográfica. El origen del aislamiento de este sublinaje no ha sido estudiado.

Caracterización de casos clínicos de C. fetus

Las actividades agrícolas y ganaderas generan contactos estrechos con animales, favoreciendo la transmisión zoonótica o antropogénica de especies bacterianas potencialmente patógenas para el humano o para los animales, respectivamente. El consumo de carnes crudas o productos lácteos no pasteurizados contaminados es la principal vía de transmisión de *C. fetus* a humanos [259], pero también se ha reportado la trasmisión la directa de humano a humano [73, 72].

La presencia de *C. fetus* en el microbioma intestinal de individuos sanos condujo a clasificar a esta especie como un patobionte. Estas personas actuarían como portadores asintomáticos, facilitando la diseminación de *C. fetus* a otros individuos. *C. fetus* es detectada por los sistemas de salud sólo al afectar a personas inmunocomprometidas, debido a la gravedad asociada al alcance sistémico de la infección en estos pacientes.

Los casos reportados en el Capítulo 6 corresponden a individuos vulnerables a la infección por este patógeno oportunista. Sólo se puede deducir la adquisición zoonótica en dos de los casos que forman parte de un cluster de aislamientos muy cercanos entre sí, pertenecientes al linaje L2 y con secuenciotipo ST4. La cercanía entre los genomas de este cluster sugiere que podría tratarse de un mismo clon, que circula en la población de forma asintomática, pero la ausencia de vínculos epidemiológicos entre los pacientes y la distribución temporal de los casos sugieren eventos de transmisión independientes.

Conclusión general

Este trabajo proporciona una visión general sobre *Campylobacter fetus* asociado a mamíferos, donde, desde la perspectiva genómica, se estudiaron aislamientos de origen clínico y de animales de producción, destacando las diferencias que los distinguen entre sí y la diversidad encontrada dentro de cada uno de estos grupos.

Un estudio poblacional a nivel general permitió reconocer una estructura de linajes con asociación a diferentes hospederos. Esta estructura se utilizó como base para el diseño de una metodología, proponiendo una actualización al clásico de diagnóstico basado en pruebas bioquímicas, a partir de evidencias a nivel genómico.

La disponibilidad de metodologías de diagnóstico confiables y fácilmente implementables tiene una importancia pivotal tanto a nivel veterinario como clínico:

- **En la práctica veterinaria**, la adopción de un método de diagnóstico de estas características permitirá montar un sistema de vigilancia epidemiológica para reconocer la verdadera prevalencia de la CGB en Uruguay, con lo qué será posible implementar medidas de prevención y contención ante la presencia de nuevos casos.
- **En la clínica,** se producirían consecuencias directas, a partir de la aplicación de la metodología en casos sospechosos, e indirectas, debido a que se limitaría la ocurrencia de casos de zoonosis, al tomar medidas de control en el ámbito veterinario.

Si bien en este trabajo no fue posible completar la validación necesaria para proponer la PCR diseñada como metodología de diagnóstico, esta se presenta como la base para una futura aplicación. Considero que la información genómica generada, la caracterización poblacional y el método de diagnóstico descritos en este trabajo pueden ser un insumo sobre el cual se puedan generar nuevos conocimientos dirigidos a la minimización del impacto de la CGB en Uruguay y sus consecuencias, desde la perspectiva de una sola salud.

8

Apéndices

Abreviaturas y siglas

ADN: Ácido desoxirribonucleico
ARNr: Ácido ribonucleico ribosomal
ARNt: Ácido ribonucleico de transferencia
AFLP: Amplified fragment length polymorphism o polimorfismo de longitud de frag-
mentos amplificados
BLAST: Basic local alignment search tool
CDS: Coding sequence o secuencia codificante
Cff: Campylobacter fetus subsp. fetus
Cfv: Campylobacter fetus subsp. venerealis
Cfvi: Campylobacter fetus subsp. venerealis biovar intermedius
Cft: Campylobacter fetus subsp.testudinum
CGB: Campylobacteriosis genital bovina
DAPC: Análisis de componentes principales discriminatorios
DILAVE: División de Laboratorios Veterinarios de la Dirección General de Servicios
Ganaderos del Ministerio de Ganadería, Agricultura y Pesca (MGAP)
dNTPs o NTPs: Nucleótidos
ddNTPs: didesoxinucleótidos
ELISA: Enzyme-Linked ImmunoSorbent Assay o ensayo por inmunoadsorción ligado
a enzimas
FCIEN: Facultad de Ciencias, Universidad de la República
GC: Contenido de Guaninas + Citosinas
IA: Inseminación artificial
IFD: Inmunofluorescencia directa

IH: Instituto de Higiene, Facultad de Medicina, Universidad de la República

INaC: Instituto Nacional de Carnes

- INIA: Instituto Nacional de Investigación Agropecuaria
- INTA: Instituto Nacional de Tecnología Agropecuaria, Ministerio de Agricultura y Pesca de la Nación, Argentina

IQR: Rango intercuartil

Kb: Kilobases. Equivalente a text 1×10^3 pb

KEGG: Kyoto Encyclopedia of Genes and Genomes

LPSN: Listado de Nombres Procariotas con Posición en Nomenclatura

MALDI-TOF: Matrix-assisted Laser Desorption/Ionization-Time Of Flight

Mb: Megabases. Equivalente a 1×10^6 pb

MCA: Multiple correspondence analysis o anñalisis de múltiples componentes

MCM: Mauve contig mover

MGAP: Ministerio de Ganadería, Agricultura y Pesca

ML: Maximum likelihood o máxima verosimilitud

MLST: Multi locus sequence typing o tipificación multilocus de secuencias

NCBI: National Center for Biotechnology Information

- NGS: Next Generation Sequencing o Secuenciación de próxima generación
- **OIE:** Oficina Internacional de Epizootias, actualmente Organización Mundial de Sanidad Animal
- **ORF:** Open reading frame o marco de lectura abierto

PATRIC: Pathosystems Resource Integration Center

pb: Pares de bases

PCA: Principal components analysis o análisis de componentes principales

PCR: Polymerase chain reaction o reacción en cadena de la polimerasa

PIB: Producto interno bruto

PE: Paired end

PFGE: Pulse field gel electrophoresis o electroforesis en gel de campo pulsado

R-CHOP: Rituximab ciclofosfamida H - doxorubicina

SE: Single end

SLD: Spotty Liver Disease

- SNP: Single nucleotide polymorphism o polimorfismo de un nucleótido
- **ST:** Sequence type
- **UFC:** Unidades formadoras de colonias
- UV: Ultravioleta
- WGA: Whole-genome alignment o alineamiento de genoma entero

Principales herramientas bioinformáticas utilizadas

- adegenet v.2.1.5: Paquete de R para el análisis exploratorio de datos genéticos
 [179]
- BEAST v.2.6.3: Paquete de análisis para estimar parámetros evolutivos y probar hipótesis. [260, 213]

BLASTn: BLAST local de secuencias nucleotídicas [145]

Bowtie2 v.2.3.5: Mapeo de reads [227]

eggNOG-Mapper v.2: Anotación de genes [173]

FactoMineR v.2.4: Paquete de R para cálculo multivariado [178]

FastPCR: Simulación in silico de PCRs [182]

ggtree v.1.14.6: Visualización de árboles filogenéticos en R [167]

Gubbins v.2.4.1: Detección de recombinación en WGA [164]

hierBAPS: Clusterización jerárquica del WGA [170]

KEGG mapper v.4.3: Anotación de genes [175]

MLSTar: Obtención del ST a partir de la secuencia genómica [162]

Pagoo v.0.3.13: Exploración del pangenoma [172]

parsnp v.1.2: Alineamiento de genomas completos [163]

Pewit v.1.1.0: Construcción del pangenoma [261]

Phandango: Visualización de árboles filogenéticos, metadatos y alineamientos [169]

Primer3: Diseño de cebadores

RaxML v8.2.11: Construcción de filogenias mediante el método ML [165]

RhierBAPS: Implementación en R de hierBAPS [171]

```
RUCS v.1.0.3: Búsqueda de blancos moleculares del genomacore [180]
```

Samtools v.1.9 y RSamtools v.2.2.1: Manipulación y visualización de mapeos en línea de comandos y su implementación en R [228, 229]

snpdist v.0.7.0: Cálculo de distancias en SNPs [166]

Tablet v.1.21.02.08: Visualización de mapeos [262]

Unipro UGENE v.43.0: Suite de herramientas para análisis y visualización de datos biológicos [183]

Tablas suplementarias

Tabla 8.1: Categorías funcionales	de grupos de or	rtólogos COG. J	Extraída de Ecoli-
Wiki.			

Categoría	Función
А	Procesamiento y modificación del ARN
В	Estructura y dinámica de la cromatina
С	Producción y conversión de energía
D	Control de ciclo celular y mitosis
Е	Transporte y metabolismo de aminoácidos
F	Transporte y metabolismo de nucleótidos
G	Transporte y metabolismo de carbohidratos
Н	Metabolismo de coenzimas
Ι	Metabolismo lipídico
J	Traducción
K	Transcripción
L	Replicación y reparación
Μ	Biogénesis de pared celular/membrana/cubierta
Ν	Movilidad celular
Ο	Modificación post-traduccional, reutilización de proteínas, funciones chaperonas
Р	Transporte y metabolismo de iones inorgánicos
Q	Estructura secundaria
Т	Transducción de señales
U	Transporte y secreción intracelular
Y	Estructura nuclear
Z	Citoesqueleto
R	Sólo predicción de la función general
S	Función desconocida

Genoma	Сера	Año	País	Hospedero	Fuente	Sub-	Sub-
						especie	linaje
1265739.3	LMG_6570	-	Bélgica	Bovino	Feto	Cfv	EU1
1273127.3	9825	2004	Argentina	Bovino	Feto	Cfvi	SA1
1273269.3	02/298	2004	Argentina	Bovino	Feto	Cfvi	SA1
1273270.3	97/532	2004	Argentina	Bovino	Vagina	Cfvi	SA1
1273272.3	92/203	2004	Argentina	Bovino	Placenta	Cfvi	SA1
1273274.4	03/596	2004	Argentina	Bovino	Feto	Cfvi	SA1
1273276.3	CCUG_33872	1995	Chequia	Bovino	-	Cfv	EU2
1273279.3	B10	2011	EEUU	Bovino		Cfv	EU1
15031_S95	15031	-	-	-	-	-	SA1
15671_S81	15671	2015	Uruguay	Bovino	Feto	Cfv	SA2
	21-C0091-					- 4 -	
16244_6.33	10-14_2	2014	Reino Unido	Bovino	Prepucio	Cfvi	SA1
16473_4.64	CAN_66Y	2012	Canadá	Bovino	Prepucio	Cfv	EU1
16473_4.65	CAN_TD	2011	Canadá	Bovino	Prepucio	Cfv	EU1
17059_2.16	99/801	1999	Argentina	Bovino	Prepucio	Cff	SA1
17059_2.17	00/398	2000	Argentina	Bovino	Feto	Cff	SA1
17059_2.18	00/564	2000	Argentina	Bovino	Prepucio	Cff	SA1
17059_2.19	01/320	2001	Argentina	Bovino	Feto	Cff	SA1
17059 2.20	01/210	2001	Argentina	Bovino	Vagina	Cff	SA1
17059 2.21	02/146	2002	Argentina	Bovino	Feto	Cfvi	SA1
17059 2.22	03/596	2003	Argentina	Bovino	Vagina	Cff	SA1
17059 2.23	04/875	2004	Argentina	Bovino	Feto	Cff	SA1
17059 2.24	05/394	2005	Argentina	Bovino	Feto	Cff	SA1
17059 2.25	05/434	2005	Argentina	Bovino	Vagina	Cff	SA1
17059 2.26	06/340	2006	Argentina	Bovino	Prepucio	Cff	SA1
17059 2.27	06/195	2006	Argentina	Bovino	Vagina	Cfvi	SA2
17059 2.28	07/379	2007	Argentina	Bovino	Feto	Cfvi	SA1
17059 2.29	07/485	2007	Argentina	Bovino	Vagina	Cff	SA1
17059 2.30	08/362	2008	Argentina	Bovino	Feto	Cff	SA1
17059 2.31	10/247	2010	Argentina	Bovino	Prepucio	Cff	SA1
17059 2.32	10/445	2010	Argentina	Bovino	Prepucio	Cff	SA1
17059 2.33	11/360	2011	Argentina	Bovino	Feto	Cff	SA1
17059 2.36	15/301	2015	Argentina	Bovino	Vagina	Cff	SA1
17059 2.56	C1	2009	España	Bovino	Prepucio	Cfv	EU1
17059 2.57	C2	2007	España	Bovino	Prepucio	Cfv	EU1
17059 2.58	C3	2008	España	Bovino	Prepucio	Cff	EU2
$17059^{-}2.59$	C4	2009	España	Bovino	Prepucio	Cff	EU2
17059 2.61	C6	2010	España	Bovino	Prepucio	Cfv	EU2
17059 2.62	C7	2007	España	Bovino	Prepucio	Cff	EU1
17059 2.63	C8	2002	España	Bovino	Prepucio	Cff	EU2
17059 2.66	C11	2014	España	Bovino	Prepucio	Cff	EU2
17059 2.67	C12	2004	España	Bovino	Prepucio	Cfvi	EU2
17059 2.69	C14	2002	España	Bovino	Prepucio	Cff	EU2
17059 2.70	C15	2008	España	Bovino	Prepucio	Cff	EU2
17059 2.71	C16	2008	España	Bovino	Prepucio	Cff	EU2
17059 2.72	C17	2002	España	Bovino	Prepucio	Cff	EU2
17059_2.74	C19	2006	España	Bovino	Prepucio	Cfv	EU1

Tabla 8.2: Metadatos de aislamientos bovinos ST4 analizados en el Capítulo 5

Tabla 8.2 – Continúa en la siguiente página

$ \begin{array}{c c c c c c c c c c c c c c c c c c c $							Sub-	Sub-
17059_2.75 C20 2010 España Bovino Prepucio Cff EU2 17059_2.76 C21 2010 España Bovino Prepucio Cff EU2 17059_2.78 C23 2007 España Bovino Prepucio Cfv EU1 17059_2.79 C24 2010 España Bovino Prepucio Cfv EU1 17059_2.81 C26 2021 España Bovino Prepucio Cfv EU1 17059_2.82 C27 2014 España Bovino Prepucio Cfv EU1 17059_2.84 C29 2014 España Bovino Prepucio Cfv EU1 17059_2.86 C33 2002 España Bovino Prepucio Cfvi EU2 17059_2.89 C34 2002 España Bovino Prepucio Cfvi EU2 17059_2.89 C34 2002 España Bovino Prepucio Cfvi	Genoma	Сера	Año	País	Hospedero	Fuente	osposio	linaio
17059 2.75 C.20 2010 España Bovino Prepucio Cff EU2 17059 2.76 C.21 2008 España Bovino Prepucio Cfv EU1 17059 2.78 C.23 2007 España Bovino Prepucio Cfv EU1 17059 2.79 C.24 2010 España Bovino Prepucio Cfv EU1 17059 2.83 C.25 2011 España Bovino Prepucio Cfv EU1 17059 2.83 C.28 2010 España Bovino Prepucio Cfv EU2 17059 2.84 C.29 2014 España Bovino Prepucio Cfv EU2 17059 2.85 C.30 2014 España Bovino Prepucio Cfv EU2 17059 2.86 C.31 2002 España Bovino Prepucio Cfv EU2 17059 2.88 C.33 2002 España Bovino Prepucio Cfv EU2 17059 2.89 C34 2002 Alemania Bovino							especie	iiiaje
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	17059_2.75	C20	2010	España	Bovino	Prepucio	Cff	EU2
	17059_2.76	C21	2010	España	Bovino	Prepucio	Cff	EU2
	17059_2.77	C22	2008	España	Bovino	Prepucio	Cfv	EU1
17059_2.79 C24 2010 España Bovino Prepucio Cfr EU1 17059_2.80 C25 2011 España Bovino Prepucio Cfr EU1 17059_2.82 C27 2011 España Bovino Prepucio Cfr EU1 17059_2.83 C28 2010 España Bovino Prepucio Cfr EU2 17059_2.84 C29 2014 España Bovino Prepucio Cfr EU2 17059_2.86 C31 2002 España Bovino Prepucio Cfr EU2 17059_2.88 C33 2002 España Bovino Prepucio Cfr EU2 17150_1.54 BS 76/04 2002 Alemania Bovino Prepucio Cfr EU1 17150_1.55 BS 38/06 2006 Alemania Bovino Prepucio Cfr EU1 17150_1.54 BS 76/04 2004 Alemania Bovino Prepucio Cfr EU1 17150_1.54 BS 76/04 2006 Alemania	17059_2.78	C23	2007	España	Bovino	Prepucio	Cfv	EU1
	17059_2.79	C24	2010	España	Bovino	Prepucio	Cfv	EU1
17059_2.81 C26 2002 España Bovino Prepucio Cfv EU1 17059_2.82 C27 2011 España Bovino Prepucio Cfv EU1 17059_2.84 C28 2010 España Bovino Prepucio Cfv EU1 17059_2.85 C30 2014 España Bovino Prepucio Cfv EU1 17059_2.86 C31 2002 España Bovino Prepucio Cfv EU2 17059_2.88 C33 2002 España Bovino Prepucio Cfv EU2 17059_2.88 C33 2002 España Bovino Prepucio Cfv EU1 17150_1.54 BS 76/04 2004 Alemania Bovino Prepucio Cfv EU1 17150_1.55 BS 8201/02 2007 Alemania Bovino Prepucio Cfv EU1 17150_1.55 BS 83/06 2006 Alemania Bovino Prepucio Cfv EU1 17150_1.55 BS 38/06 2007 Alemania <t< td=""><td>17059_2.80</td><td>C25</td><td>2011</td><td>España</td><td>Bovino</td><td>Prepucio</td><td>Cfv</td><td>EU1</td></t<>	17059_2.80	C25	2011	España	Bovino	Prepucio	Cfv	EU1
17059_2.82 C27 2011 España Bovino Prepucio Cfv EU1 17059_2.83 C28 2010 España Bovino Prepucio Cfvi EU2 17059_2.84 C29 2014 España Bovino Prepucio Cfvi EU1 17059_2.85 C30 2014 España Bovino Prepucio Cfvi EU2 17059_2.87 C32 2004 España Bovino Prepucio Cfvi EU2 17059_2.88 C33 2002 España Bovino Prepucio Cfvi EU2 17150_1.53 BS 201/02 2002 Alemania Bovino Prepucio Cfv EU1 17150_1.55 BS 38/06 2006 Alemania Bovino Prepucio Cfv EU1 17150_1.55 BS 38/06 2006 Alemania Bovino Prepucio Cfv EU1 17150_1.59 11CS0190 2011 Alemania Bovino Prepucio Cfv EU1 17150_1.61 11CS0191 2011 Alemania <td>17059_2.81</td> <td>C26</td> <td>2002</td> <td>España</td> <td>Bovino</td> <td>Prepucio</td> <td>Cfvi</td> <td>EU2</td>	17059_2.81	C26	2002	España	Bovino	Prepucio	Cfvi	EU2
17059 2.83 C28 2010 España Bovino Prepucio Cfri EU2 17059 2.84 C29 2014 España Bovino Prepucio Cfri EU2 17059 2.85 C30 2014 España Bovino Prepucio Cfri EU2 17059 2.86 C31 2002 España Bovino Prepucio Cfri EU2 17059 2.88 C33 2002 España Bovino Prepucio Cfri EU2 17059 2.89 C34 2002 Alemania Bovino Prepucio Cfri EU2 17150 1.53 BS 201/02 2002 Alemania Bovino Prepucio Cfr EU1 17150 1.54 BS 76/04 2004 Alemania Bovino Prepucio Cfr EU1 17150 1.55 BS 38/06 2006 Alemania Bovino Prepucio Cfr EU1 17150 1.59 B12C0120 2007 Alemania Bovino Prepucio Cfr EU1 17150 1.61 13CS0183 2013 Alemania	17059_2.82	C27	2011	España	Bovino	Prepucio	Cfv	EU1
17059 2.84 C29 2014 España Bovino Prepucio Cfr EU2 17059 2.85 C30 2014 España Bovino Prepucio Cfv EU1 17059 2.86 C31 2002 España Bovino Prepucio Cfvi EU2 17059 2.87 C32 2004 España Bovino Prepucio Cfvi EU2 17153 2.89 C34 2002 España Bovino Prepucio Cfv EU1 17150 1.53 BS 201/02 2002 Alemania Bovino Prepucio Cfv EU1 17150 1.55 BS 38/06 2006 Alemania Bovino Prepucio Cfv EU1 17150 1.56 0785020 2007 Alemania Bovino Prepucio Cfv EU1 17150 1.61 13CS0183 2013 Alemania Bovino Prepucio Cfv EU1 17150 1.61 13CS0183 2014 Alemania Bovino Prepucio<	17059_2.83	C28	2010	España	Bovino	Prepucio	Cfvi	EU2
	17059_2.84	C29	2014	España	Bovino	Prepucio	Cff	EU2
	17059_2.85	C30	2014	España	Bovino	Prepucio	Cfv	EU1
	17059_2.86	C31	2002	España	Bovino	Prepucio	Cfvi	EU2
	17059_2.87	C32	2004	España	Bovino	Prepucio	Cfvi	EU2
17059_2.89 C34 2002 España Bovino Prepucio Cfv EU2 17133_S82 17133 2017 Uruguay Bovino Feto Cfv SA2 17150_1.53 BS 201/02 2002 Alemania Bovino Prepucio Cfv EU1 17150_1.55 BS 38/06 2006 Alemania Bovino Prepucio Cfv EU1 17150_1.58 0PS0030 2007 Alemania Bovino Prepucio Cfv EU1 17150_1.59 11CS0190 2011 Alemania Bovino Prepucio Cfv EU1 17150_1.60 11CS0191 2011 Alemania Bovino Prepucio Cfv EU1 17150_1.61 13CS0183 2013 Alemania Bovino Prepucio Cfv EU1 17150_1.62 14CS001 2014 Alemania Bovino Prepucio Cfv EU1 17150_1.62 14CS001 2014 Alemania Bovino Prepucio Cfv SA1 18048_2.15 6 2000	17059_2.88	C33	2002	España	Bovino	Prepucio	Cfvi	EU2
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17059_2.89	C34	2002	España	Bovino	Prepucio	Cfv	EU2
17150_1.53 BS 201/02 2002 Alemania Bovino Prepucio Cfv EU1 17150_1.54 BS 76/04 2004 Alemania Bovino Prepucio Cfv EU1 17150_1.55 BS 38/06 2006 Alemania Bovino Prepucio Cfv EU1 17150_1.56 07BS020 2007 Alemania Bovino Prepucio Cfv EU1 17150_1.58 09CS0030 2009 Alemania Bovino Prepucio Cfv EU1 17150_1.61 13CS0191 2011 Alemania Bovino Prepucio Cfv EU1 17150_1.62 14CS0001 2014 Alemania Bovino Prepucio Cfv EU1 17150_1.62 14CS0001 2014 Alemania Bovino Prepucio Cff SA1 18048_2.15 6 2000 Reino Unido Bovino Prepucio Cff SA2 18048_2.23 52/13_5 2013 Brasil Bovino Prepucio Cfv SA2 2740_S96 2740 -	17133_S82	17133	2017	Uruguay	Bovino	Feto	Cfv	SA2
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17150_1.53	BS 201/02	2002	Alemania	Bovino	Prepucio	Cfv	EU1
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17150_1.54	BS 76/04	2004	Alemania	Bovino	Feto	Cfv	EU1
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	17150_1.55	BS 38/06	2006	Alemania	Bovino	Prepucio	Cfv	EU1
17150_1.58 09CS0030 2009 Alemania Bovino Prepucio Cfv EU1 17150_1.59 11CS0190 2011 Alemania Bovino Prepucio Cfv EU1 17150_1.60 11CS0191 2011 Alemania Bovino Prepucio Cfv EU1 17150_1.61 13CS0183 2013 Alemania Bovino Prepucio Cfv EU1 17150_1.61 13CS0183 2013 Alemania Bovino Prepucio Cfv EU1 18048_2.11 2 1997 Reino Unido Bovino Prepucio Cff SA1 18048_2.27 564/98 1998 Brasil Bovino Prepucio Cff SA2 2740_896 2740 - Uruguay - - SA2 32020.1 ADRIS13 - Australia - Cfv SA1 32020.6 WBT011/09 - Reino Unido Bovino - Cfv SA1 32020.7 ADRI1362 1989 Argentina Bovino - -	17150_1.56	07BS020	2007	Alemania	Bovino	Prepucio	Cfv	EU1
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17150_1.58	09CS0030	2009	Alemania	Bovino	Prepucio	Cfv	EU1
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17150_1.59	11CS0190	2011	Alemania	Bovino	Prepucio	Cfv	EU1
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17150_1.60	11CS0191	2011	Alemania	Bovino	Prepucio	Cfv	EU1
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17150_1.61	13CS0183	2013	Alemania	Bovino	Prepucio	Cfv	EU1
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	17150_1.62	14CS0001	2014	Alemania	Bovino	Prepucio	Cfv	EU1
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	18048_2.11	2	1997	Reino Unido	Bovino	Prepucio	Cff	SA1
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	18048_2.15	6	2000	Reino Unido	Bovino	Prepucio	Cff	SA1
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	18048_2.27	564/98	1998	Brasil	Bovino	Prepucio	Cff	SA2
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$18048_{2.32}$	52/13_5	2013	Brasil	Bovino	Prepucio	Cfv	SA2
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	2740_S96	2740	-	Uruguay	-	-	-	SA2
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	32020.1	ADRI513	-	Australia	-	-	Cfv	EU2
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	32020.12	zaf65	2007	Sudáfrica	Bovino	-	Cfv	SA1
32020.7ADRI13621989ArgentinaBovino-CfvSA132020.9zaf32006SudáfricaBovino-CfvEU295258_S94952581995ArgentinaSA1ATCC_S93ATCC-1962Reino UnidoBovinoVaginaCfvEU1H1_S84H12016UruguayBovinoPrepucioCffSA2M8_S86M82017UruguayBovinoVaginaCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2K11_S9112/182001ArgentinaBovinoPrepucioCfvSA2K29_S92X292017UruguayBovinoPrepucioCfvSA2K29_S92X292017UruguayBovinoPrepucioCfvSA2K29_S92X292017UruguayBovinoPrepucioCfvSA2K29_S92X292017UruguayBovinoPrepucioCfvSA2K29_S92X292017UruguayBovinoPrepucio	32020.6	WBT011/09	-	Reino Unido	Bovino	-	Cfv	SA1
32020.9zaf32006SudáfricaBovino-CfvEU295258_S94952581995ArgentinaSA1ATCC_S93ArCC- 194381962Reino UnidoBovinoVaginaCfvEU1H1_S84H12016UruguayBovinoPrepucioCffSA2M8_S86M82017UruguayBovinoVaginaCfvSA2R18_S87R182017UruguayBovinoPrepucioCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2K160116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	32020.7	ADRI1362	1989	Argentina	Bovino	-	Cfv	SA1
95258_S94952581995ArgentinaSA1ATCC_S93ATCC- 194381962Reino UnidoBovinoVaginaCfvEU1H1_S84H12016UruguayBovinoPrepucioCffSA2M8_S86M82017UruguayBovinoVaginaCfvSA2R18_S87R182017UruguayBovinoPrepucioCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2K11_S9112/2182001ArgentinaBovinoVaginaCfvSA2K29_S92X292017UruguayBovinoPrepucioCfvSA2K29_S92X292017UruguayBovinoPrepucioCfvSA1cff0116501/1652001ArgentinaBovinoFetoCffSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	32020.9	zaf3	2006	Sudáfrica	Bovino	-	Cfv	EU2
ATCC_S93ATCC- 194381962Reino UnidoBovinoVaginaCfvEU1H1_S84H12016UruguayBovinoPrepucioCffSA2M8_S86M82017UruguayBovinoVaginaCfvSA2R18_S87R182017UruguayBovinoPrepucioCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	95258_S94	95258	1995	Argentina	-	-	-	SA1
AICC_S931962Reino UnidoBovinoVaginaCrvEU119438194382016UruguayBovinoPrepucioCffSA2M8_S86M82017UruguayBovinoVaginaCfvSA2R18_S87R182017UruguayBovinoPrepucioCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1		ATCC-	10(0	D' 11'1			<u> </u>	DI 11
H1_S84H12016UruguayBovinoPrepucioCffSA2M8_S86M82017UruguayBovinoVaginaCfvSA2R18_S87R182017UruguayBovinoPrepucioCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	AICC_893	19438	1962	Reino Unido	Bovino	vagina	CIV	EUI
M8_S86M82017UruguayBovinoVaginaCfvSA2R18_S87R182017UruguayBovinoPrepucioCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCfiSA1	H1 S84	H1	2016	Uruguay	Bovino	Prepucio	Cff	SA2
R18_S87R182017UruguayBovinoPrepucioCfvSA2U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	M8 S86	M8	2017	Uruguav	Bovino	Vagina	Cfv	SA2
U10_S88U102017UruguayBovinoPrepucioCfvSA2V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	R18 S87	R18	2017	Uruguav	Bovino	Prepucio	Cfv	SA2
V28_S89V282017UruguayBovinoPrepucioCfvSA2W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	U10 S88	U10	2017	Uruguay	Bovino	Prepucio	Cfv	SA2
W11_S91W112017UruguayBovinoPrepucioCfvSA2W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	V28 S89	V28	2017	Uruguay	Bovino	Prepucio	Cfv	SA2
W8_S90W82017UruguayBovinoPrepucioCfvSA2X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	W11 S91	W11	2017	Uruguav	Bovino	Prepucio	Cfv	SA2
X29_S92X292017UruguayBovinoPrepucioCfvSA2cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	W8 S90	W8	2017	Uruguav	Bovino	Prepucio	Cfv	SA2
cff0116501/1652001ArgentinaBovinoVaginaCfviSA1cff1221812/2182012ArgentinaBovinoFetoCffSA1	X29 S92	X29	2017	Uruguay	Bovino	Prepucio	Cfv	SA2
cff12218 12/218 2012 Argentina Bovino Feto Cff SA1	cff01165	01/165	2001	Argentina	Bovino	Vagina	Cfvi	SA1
	cff12218	12/218	2012	Argentina	Bovino	Feto	Cff	SA1
ctt90189 90/189 1990 Argentina Bovino Feto Cff SA1	cff90189	90/189	1990	Argentina	Bovino	Feto	Cff	SA1

Tabla 8.2 – Continúa en la siguiente página

Tabla 8.2,	Continuación	ł
------------	--------------	---

Genoma	Сера	Año	País	Hospedero	Fuente	Sub- especie	Sub- linaje
cfv05355	05/355	2005	Argentina	Bovino	Feto	Cfv	SA1
cfv10354	NCTC- 10354	1952	Reino Unido	Bovino	Vagina	Cfv	EU1
cfv84112	84/112	1984	EEUU	Bovino	Vagina	Cfv	EU1
cfv90264	90/264	1990	Argentina	Bovino	Feto	Cfv	SA1
cfv97608	97/608	1997	Argentina	Bovino	Placenta	Cfv	EU1
cfvB6	B6	1964	Australia	Bovino	Vagina	Cfv	EU1
cfvi03293	03/293	2003	Argentina	Bovino	Feto	Cfvi	SA1
cfvi64221	64-221	1967	Australia	Bovino	Vagina	Cfvi	EU2
cfvi99541	99/541	1999	Argentina	Bovino	Prepucio	Cfvi	SA1

Genoma	Сера	Año	País	Hospedero	Fuente	Subes-	Linaje
						pecie	
1273127.3	9825	2004	Argentina	Bovino	Feto	Cfvi	L1
1273269.3	02/298	2004	Argentina	Bovino	Feto	Cfvi	L1
1273270.3	97/532	2004	Argentina	Bovino	Vagina	Cfvi	L1
1273272.3	92/203	2004	Argentina	Bovino	Placenta	Cfvi	L1
1273274.4	03/596	2004	Argentina	Bovino	Feto	Cfvi	L1
1273276.3	CCUG_33872	1995	Chequia	Bovino	-	Cfv	L1
1273279.3	B10	2011	EEUU	Bovino	-	Cfv	L1
15671_S81	15671	2015	Uruguay	Bovino	Feto	Cfv	L1
16244_6.33	21-C0091- 10-14 2	2014	Reino Unido	Bovino	Prepucio	Cfvi	L1
16473 4.64	CAN 66Y	2012	Canadá	Bovino	Prepucio	Cfv	L1
16473 4.65	CAN TD	2011	Canadá	Bovino	Prepucio	Cfv	L1
17059 2.16	99/801	1999	Argentina	Bovino	Prepucio	Cff	L1
17059 2.17	00/398	2000	Argentina	Bovino	Feto	Cff	L1
17059 2.18	00/564	2000	Argentina	Bovino	Prepucio	Cff	L1
17059 2.19	01/320	2001	Argentina	Bovino	Feto	Cff	L1
17059 2.20	01/210	2001	Argentina	Bovino	Vagina	Cff	L1
17059 2.21	02/146	2002	Argentina	Bovino	Feto	Cfvi	L1
17059 2.22	03/596	2003	Argentina	Bovino	Vagina	Cff	L1
17059 2.23	04/875	2004	Argentina	Bovino	Feto	Cff	L1
17059 2.24	05/394	2005	Argentina	Bovino	Feto	Cff	L1
17059 2.25	05/434	2005	Argentina	Bovino	Vagina	Cff	L1
17059 2.26	06/340	2006	Argentina	Bovino	Prepucio	Cff	L1
17059 2.27	06/195	2006	Argentina	Bovino	Vagina	Cfvi	L1
17059 2.28	07/379	2007	Argentina	Bovino	Feto	Cfvi	L1
17059 2.29	07/485	2007	Argentina	Bovino	Vagina	Cff	L1
17059 2.30	08/362	2008	Argentina	Bovino	Feto	Cff	L1
17059 2.31	10/247	2010	Argentina	Bovino	Prepucio	Cff	L1
17059 2.32	10/445	2010	Argentina	Bovino	Prepucio	Cff	L1
17059 2.33	11/360	2011	Argentina	Bovino	Feto	Cff	L1
17059 2.34	11/427	2011	Argentina	Bovino	Vagina	Cff	L1
17059 2.35	14/270	2014	Argentina	Bovino	Feto	Cff	L1
17059 2.36	15/301	2015	Argentina	Bovino	Vagina	Cff	L1
17059_2.56	C1	2009	España	Bovino	Prepucio	Cfv	L1
17059 2.57	C2	2007	España	Bovino	Prepucio	Cfv	L1
17059 2.58	C3	2008	España	Bovino	Prepucio	Cff	L1
17059_2.59	C4	2009	España	Bovino	Prepucio	Cff	L1
17059_2.61	C6	2010	España	Bovino	Prepucio	Cfv	L1
17059_2.62	C7	2007	España	Bovino	Prepucio	Cff	L1
17059 2.63	C8	2002	España	Bovino	Prepucio	Cff	L1
	C11	2014	España	Bovino	Prepucio	Cff	L1
17059_2.67	C12	2004	España	Bovino	Prepucio	Cfvi	L1
17059_2.69	C14	2002	España	Bovino	Prepucio	Cff	L1
17059 2.70	C15	2008	España	Bovino	Prepucio	Cff	L1
17059_2.71	C16	2008	España	Bovino	Prepucio	Cff	L1
17059_2.72	C17	2002	España	Bovino	Prepucio	Cff	L1

Tabla 8.3: Metadatos de genomas utilizados para reconstruir la filogenia datada en
el Capítulo 5

Tabla 8.3 – Continúa en la siguiente página

abia 0.0, commuteton

Genoma	Сера	Año	País	Hospedero	Fuente	Subes- pecie	Linaje
17059 2.74	C19	2006	España	Bovino	Prepucio	Cfv	L1
17059 2.75	C20	2010	España	Bovino	Prepucio	Cff	L1
17059 2.76	C21	2010	España	Bovino	Prepucio	Cff	L1
17059 2.77	C22	2008	España	Bovino	Prepucio	Cfv	L1
17059 2.78	C23	2007	España	Bovino	Prepucio	Cfv	L1
17059 2.79	C24	2010	España	Bovino	Prepucio	Cfv	L1
17059 2.80	C25	2011	España	Bovino	Prepucio	Cfv	L1
17059 2.81	C26	2002	España	Bovino	Prepucio	Cfvi	L1
17059 2.82	C27	2011	España	Bovino	Prepucio	Cfv	L1
17059 2.83	C28	2010	España	Bovino	Prepucio	Cfvi	L1
17059 2.84	C29	2014	España	Bovino	Prepucio	Cff	L1
17059 2.85	C30	2014	España	Bovino	Prepucio	Cfv	L1
17059 2.86	C31	2002	España	Bovino	Prepucio	Cfvi	L1
17059 2.87	C32	2004	España	Bovino	Prepucio	Cfvi	11
17059 2.88	C33	2002	España	Bovino	Prepucio	Cfvi	L1
17059 2.80	C34	2002	España	Bovino	Prepucio	Cfv	II I1
17133 \$82	17133	2002	Uruguay	Bovino	Feto	Cfv	II I1
17150 1 52	BS 201/02	2017	Alemania	Bovino	Prepucio	Cfv	II I1
17150_1.53	BS 76/04	2002	Alemania	Bovino	Fiepucio	Cfv	II I1
17150_1.54	DS 70/04	2004	Alemania	Bovino	Dropucio	Cfv	II I1
17150_1.55	0725020	2000	Alemania	Bovino	Prepucio	Cfv	L1 I 1
17150_1.50	0783020	2007	Alemania	Bovino	Prepucio	CIV	L1 11
17150_1.56	11660100	2009	Alemania	Bovino	Prepucio	CIV	LI 11
17150_1.59	11C50190	2011	Alemania	DOVIIIO Dessine	Prepucio	CIV	LI 11
17150_1.60	11050191	2011	Alemania	Bovino	Prepucio	CIV	LI I 1
17150_1.01	13050183	2013	Alemania	Bovino	Prepucio	CIV	LI I 1
1/150_1.62	14050001	2014	Alemania Deine Unite	Bovino	Prepucio	CIV	LI I 1
18048_2.11	2	1997	Reino Unido	Bovino	Prepucio	CII	LI I 1
18048_2.15	ь 14	2000	Reino Unido	Bovino	Prepucio	CII	LI I 1
18048_2.21	14	2007	Reino Unido	Ovino	Placenta	CII	
18048_2.26	515/98	1998	Brasil	Bovino	Prepucio	Cfv	
18048_2.27	564/98	1998	Brasil	Bovino	Prepucio	Cff	
18048_2.32	52/13_5	2013	Brasil	Bovino	Prepucio	Cfv	L1
32020.12	zaf65	2007	Sudáfrica	Bovino	-	Cfv	L1
32020.7	ADRI1362	1989	Argentina	Bovino	-	Cfv	L1
32020.9	zaf3	2006	Sudáfrica	Bovino	-	Cfv	L1
H1_S84	H1	2016	Uruguay	Bovino	Prepucio	Cff	L1
M8_S86	M8	2017	Uruguay	Bovino	Vagina	Cfv	L1
R18_S87	R18	2017	Uruguay	Bovino	Prepucio	Cfv	L1
V28_S89	V28	2017	Uruguay	Bovino	Prepucio	Cfv	L1
W11_S91	W11	2017	Uruguay	Bovino	Prepucio	Cfv	L1
W8_S90	W8	2017	Uruguay	Bovino	Prepucio	Cfv	L1
X29_S92	X29	2017	Uruguay	Bovino	Prepucio	Cfv	L1
cff01165	01/165	2001	Argentina	Bovino	Vagina	Cfvi	L1
cff12218	12/218	2012	Argentina	Bovino	Feto	Cff	L1
cff90189	90/189	1990	Argentina	Bovino	Feto	Cff	L1
cfv05355	05/355	2005	Argentina	Bovino	Feto	Cfv	L1
cfv10354	NCTC_10354	1952	Reino Unido	Bovino	Vagina	Cfv	L1
cfv84112	84/112	1984	EEUU	Bovino	Vagina	Cfv	L1

Tabla 8.3 – Continúa en la siguiente página

Tabla	8.3,	Continu	ación
-------	------	---------	-------

Genoma	Сера	Año	País	Hospedero	Fuente	Subes- pecie	Linaje
cfv90264	90/264	1990	Argentina	Bovino	Feto	Cfv	L1
cfv97608	97/608	1997	Argentina	Bovino	Placenta	Cfv	L1
cfvB6	B6	1964	Australia	Bovino	Vagina	Cfv	L1
cfvi03293	03/293	2003	Argentina	Bovino	Feto	Cfvi	L1
cfvi64221	64-221	1967	Australia	Bovino	Vagina	Cfvi	L1
cfvi99541	99/541	1999	Argentina	Bovino	Prepucio	Cfvi	L1
16244 6.11	2008/898h	2008	Francia	Humano	Sangre	Cff	L2
16244_6.12	2010/41h	2010	Francia	Humano	Heces	Cff	L2
16244_6.15	2010/1119h	2010	Francia	Humano	Heces	Cff	L2
16244_6.16	2010/1180h	2010	Francia	Humano	Sangre	Cff	L2
16244_6.20	2012/331h	2012	Francia	Humano	Sangre	Cff	L2
16244_6.24	2014/52h	2014	Francia	Humano	CSF	Cff	L2
16244_6.3	2006/588h	2006	Francia	Humano	CSF	Cff	L2
16244_6.30	2007/123h	2007	Francia	Humano	CSF	Cff	L2
16244_6.31	2009/56h	2009	Francia	Humano	CSF	Cff	L2
16473_4.92	2004/605h	2004	Francia	Humano	Heces	Cff	L2
17059_2.12	ID136551	2014	Canadá	Humano	Sangre	Cff	L2
17150_1.57	08CS0024	2008	Alemania	Bovino	Prepucio	Cfv	L2
17150_1.74	001A-0648	2007	Canadá	Humano	Sangre	Cff	L2
18048_2.25	161/97	1997	Brasil	Bovino	Prepucio	Cfv	L2
HC	H2-UY	2014	Uruguay	Humano	Sangre	Cff	L2
HC73	H6-UY	2015	Uruguay	Humano	Sangre	Cff	L2
KCH09208	H1-JP	2015	Japón	Humano	Sangre	NA	L2
cff8240	82-40	2006	EEUU	Humano	Sangre	Cff	L2
cffH1UY	H1-UY	2014	Uruguay	Humano	Sangre	Cff	L2
cffHC1	HC1	2014	Uruguay	Humano	Sangre	Cff	L2
cffHC2csf	HC2	2014	Uruguay	Humano	CSF	Cff	L2
17059_2.13	ID136656	2014	Canadá	Humano	Sangre	Cff	L3
17059_2.44	0003304-2	2009	Taiwan	Humano	Sangre	Cff	L3
17059_2.45	2115	2010	Taiwan	Humano	Sangre	Cff	L3
17150 1.65	BS 03/04	2004	Alemania	Bovino	Feto	Cff	L3
17150 1.67	08CS0027	2008	Alemania	Bovino	Prepucio	Cff	L3
1273268.3	BT 10/98	1999	Reino Unido	Ovino	-	Cff	L4
16244_6.25	2014/602h	2014	Francia	Humano	Sangre	Cff	L4
17150 1.72	13CS0373	2013	Alemania	Monkey	Heces	Cff	L4
18048 2.12	3	1997	Reino Unido	Ovino	Placenta	Cff	L4
18048 2.13	4	1998	Reino Unido	Ovino	Placenta	Cff	L4
17059 2.40	My5726	2006	Taiwan	Humano	Sangre	Cff	L5
17059 2.42	1830	2008	Taiwan	Humano	Sangre	Cff	L5
17059 2.51	9502	2011	Taiwan	Humano	Sangre	Cff	L5
17059 2.53	8031708	2012	Taiwan	Humano	Sangre	Cff	L5
17059 2.55	3069482	2014	Taiwan	Humano	Sangre	Cff	L5
HC71	H5-UY	2015	Uruguay	Humano	Ascitos	Cff	L5
16244 6.19	2012/286h	2012	Francia	Humano	Sangre	Cff	L6
17059 2.14	ID136706	2014	Canadá	Humano	Sangre	Cff	L6
17059 2.15	ID132939	2014	Canadá	Humano	Sangre	Cff	L6
17059_2.68	C13	2011	España	Bovino	Prepucio	Cff	L6

Tabla 8.3 – Continúa en la siguiente página

Tabla 8.3,	Continuaci	ón
------------	------------	----

Genoma	Сера	Año	País	Hospedero	Fuente	Subes- pecie	Linaje
17150_1.63	BS 456/99	1999	Alemania	Ovino	Feto	Cff	L6
17150_1.69	11CS0098	2011	Alemania	Ovino	Placenta	Cff	L6
18048_2.23	17	2008	Reino Unido	Ovino	Feto	Cff	L6
16244_6.2	2006/479h	2006	Francia	Humano	Heces	Cff	L7
17144_S83	17144	2017	Uruguay	Ovino	Feto	Cfv	L7
HB74	H7-UY	2015	Uruguay	Humano	Sangre	Cff	L7
16244_6.32	CF156	2013	Turquía	Humano	Sangre	Cff	L8
17059_2.8	ID129038	2013	Canadá	Humano	Sangre	Cff	L8
17150_1.71	13CS0001	2013	Alemania	Bovino	Prepucio	Cff	L8
18048_2.10	1	1996	Reino Unido	Bovino	Prepucio	Cff	L8
cff04554	04/554	2004	Argentina	Bovino	Feto	Cff	L8

Nombre-	Genoma	Сера	Año	País	Fuente	Sub-	Sub-
Accession	Genoma	Gepu	Thio	T uits	Tuente	especie	linaje
ERR987447	16473 4.64	CAN 66Y	2012	Canadá	Prepucio	Cfv	EU1
ERR987448	16473 4.65	CAN TD	2011	Canadá	Prepucio	Cfv	EU1
ERR1046025	17059 2.56	C1	2009	España	Prepucio	Cfv	EU1
ERR1046026	17059 2.57	C2	2007	España	Prepucio	Cfv	EU1
ERR1046031	17059 2.62	C7	2007	España	Prepucio	Cff	EU1
ERR1046043	17059 2.74	C19	2006	España	Prepucio	Cfv	EU1
ERR1046046	17059 2.77	C22	2008	España	Prepucio	Cfv	EU1
ERR1046047	17059 2.78	C23	2007	España	Prepucio	Cfv	EU1
ERR1046048	17059 2.79	C24	2010	España	Prepucio	Cfv	EU1
ERR1046049	17059 2.80	C25	2011	España	Prepucio	Cfv	EU1
ERR1046051	17059 2.82	C27	2011	España	Prepucio	Cfv	EU1
ERR1046054	17059 2.85	C30	2014	España	Prepucio	Cfv	EU1
ERR1069013	17150 1.53	BS 201/02	2002	Alemania	Prepucio	Cfv	EU1
ERR1069014	17150 1.54	BS 76/04	2004	Alemania	Feto	Cfv	EU1
ERR1069015	17150 1.55	BS 38/06	2006	Alemania	Prepucio	Cfv	EU1
ERR1069016	17150 1.56	07BS020	2007	Alemania	Prepucio	Cfv	EU1
ERR1069018	17150 1.58	09CS0030	2009	Alemania	Prepucio	Cfv	EU1
ERR1069019	17150 1.59	11CS0190	2011	Alemania	Prepucio	Cfv	EU1
ERR1069020	17150 1.60	11CS0191	2011	Alemania	Prepucio	Cfv	EU1
ERR1069021	17150 1.61	13CS0183	2013	Alemania	Prepucio	Cfv	EU1
ERR1069022	$17150_{-1.62}$	14CS0001	2010	Alemania	Prepucio	Cfv	EU1
21007022	1,100_1102	ATCC-		, nonnanna	Tropuelo		201
NC	ATCC_S93	19438	1965	Reino Unido	Vagina	Cfv	EU1
ERR1046027	17059_2.58	C3	2008	España	Prepucio	Cff	EU2
ERR1046028	17059_2.59	C4	2009	España	Prepucio	Cff	EU2
ERR1046030	17059_2.61	C6	2010	España	Prepucio	Cfv	EU2
ERR1046032	17059_2.63	C8	2002	España	Prepucio	Cff	EU2
ERR1046035	17059_2.66	C11	2014	España	Prepucio	Cff	EU2
ERR1046036	17059_2.67	C12	2004	España	Prepucio	Cfvi	EU2
ERR1046038	17059_2.69	C14	2002	España	Prepucio	Cff	EU2
ERR1046039	17059 2.70	C15	2008	España	Prepucio	Cff	EU2
ERR1046040	17059 2.71	C16	2008	España	Prepucio	Cff	EU2
ERR1046041	17059 2.72	C17	2002	España	Prepucio	Cff	EU2
ERR1046044	17059 2.75	C20	2010	España	Prepucio	Cff	EU2
ERR1046045	17059 2.76	C21	2010	España	Prepucio	Cff	EU2
ERR1046050	17059 2.81	C26	2002	España	Prepucio	Cfvi	EU2
ERR1046052	17059 2.83	C28	2010	España	Prepucio	Cfvi	EU2
ERR1046053	17059 2.84	C29	2014	España	Prepucio	Cff	EU2
ERR1046055	17059 2.86	C31	2002	España	Prepucio	Cfvi	EU2
ERR1046056	17059 2.87	C32	2004	España	Prepucio	Cfvi	EU2
ERR1046057	17059 2.88	C33	2002	España	Prepucio	Cfvi	EU2
ERR1046058	17059 2.89	C34	2002	España	Prepucio	Cfv	EU2
NC	15031 \$95	15031		-	-	-	SA1
EDD07(000	16044 6 00	21-C0091-	2014	Daina TT-14	Duen	Cf!	C 4 1
ERR976390	16244_6.33	10-14_2	2014	Reino Unido	Prepucio	Cfvi	SAI
ERR1045985	17059_2.16	99/801	1999	Argentina	Prepucio	Cff	SA1
ERR1045986	17059_2.17	00/398	2000	Argentina	Feto	Cff	SA1
ERR1045987	17059_2.18	00/564	2000	Argentina	Prepucio	Cff	SA1

Tabla 8.4: Datos de secuenciación masiva utilizados para los mapeos

Tabla 8.4 – Continúa en la siguiente página

Tabla 8.4, Continuación

Nombre- Accession	Genoma	Сера	Año	País	Fuente	Sub- especie	Sub- linaje
EDD1045088	17050 2 10	01/320	2001	Argenting	Feto	Cff	SA1
ERR1045080	17059_2.19	01/320	2001	Argentina	Vagina	Cff	SA1
ERR1045000	17059_2.20	02/146	2001	Argentina	Feto	Cfvi	SA1
ERR1045001	17059_2.21	03/596	2002	Argentina	Vagina	Cff	SA1
ERR1045002	17059_2.22	04/875	2003	Argentina	Feto	Cff	SA1
ERR1045003	$17059_{2.23}$	05/304	2004	Argentina	Feto	Cff	SA1
ERR1045004	17059_2.21	05/434	2005	Argentina	Vagina	Cff	SA1
ERR1045005	17059_2.25	06/340	2005	Argentina	Prepucio	Cff	SA1
FRR1045997	17059_2.20	07/379	2000	Argentina	Feto	Cfvi	SA1
ERR1045008	17059_2.20	07/485	2007	Argentina	Vagina	Cff	SA1
ERR1045000	17059_2.29	08/362	2007	Argentina	Feto	Cff	SA1
ERR1046000	$17059_{2.30}$ 17059_2.31	10/247	2000	Argentina	Prepucio	Cff	SA1
FRR1046001	17059_2.31	10/445	2010	Argentina	Prepucio	Cff	SA1
ERR1046002	17059_2.32	11/360	2010	Argentina	Feto	Cff	SA1
ERR1203002	18048 2 15	6	2011	Reino Unido	Prepucio	Cff	SA1
NC	95258 594	95258	1995	Argentina	-	-	SA1
NC	15671 \$81	15671	2015	Uruguay	Feto	Cfv	542
FRR1045996	17059 2 27	06/195	2015	Argentina	Vagina	Cfvi	SA2
NC	17133 \$82	17133	2000	Uruguay	Feto	Cfv	SA2
NC	2740 \$96	2740	-	Uruguay	-	GIV	SA2
NC	H1 S84	27 10 H1	2016	Uruguay	Prenucio	Cff	SA2
NC	M8_S86	M8	2017	Uruguay	Vagina	Cfv	SA2
NC	R18_S87	R18	2017	Uruguay	Prepucio	Cfv	SA2
NC	U10_S88	U10	2017	Uruguay	Prepucio	Cfv	SA2
NC	V28_S89	V28	2017	Uruguay	Prepucio	Cfv	SA2
NC	W11_S91	W11	2017	Uruguay	Prepucio	Cfv	SA2
NC	W8 S90	W8	2017	Uruguay	Prepucio	Cfv	SA2
NC	X29 S92	X29	2017	Uruguay	Prepucio	Cfv	SA2
ERR1046003	17059 2.34	11/427	2011	Argentina	Vagina	Cff	-

NC: Datos de secuenciación masiva no publicados

Figuras suplementarias





Bibliografía

- [1]America Mederos. *Diagnóstico y control de Campylobacteriosis genital bovina*. 23/10/2015. INIA Tacuarembó (vid. págs. 1, 28, 29).
- [2]Rafael Delpiazzo. «AISLAMIENTO E IDENTIFICACIÓN DE CAMPYLOBACTER FETUS Y TRITRICHOMONA FOETUS EN BOVINOS DEL URUGUAY». Tesis doct. Uruguay, 2018 (vid. págs. 2, 83).
- [3]Rafael Delpiazzo, Maila Barcellos, Sofía Barros et al. «Accurate and fast identification of Campylobacter fetus in bulls by real-time PCR targeting a 16S rRNA gene sequence». en. En: *Veterinary and Animal Science* 11 (mar. de 2021), pág. 100163 (vid. págs. 2, 24).
- [4]Ellyn P. Marder, Paul R. Cieslak, Alicia B. Cronquist et al. «Incidence and Trends of Infections with Pathogens Transmitted Commonly Through Food and the Effect of Increasing Use of Culture-Independent Diagnostic Tests on Surveillance — Foodborne Diseases Active Surveillance Network, 10 U.S. Sites, 2013–2016». En: MMWR. Morbidity and Mortality Weekly Report 66.15 (abr. de 2017), págs. 397-403 (vid. pág. 6).
- [5]Collette Fitzgerald. «Campylobacter». En: *Clinics in Laboratory Medicine* 35.2 (jun. de 2015). ISBN: 9780323388948 Publisher: Elsevier BV, págs. 289-298 (vid. págs. 6, 17).
- [6]Aidan C. Parte. «LPSN—list of prokaryotic names with standing in nomenclature». En: *Nucleic Acids Research* 42.D1 (ene. de 2014). Publisher: Oxford University Press, págs. D613-D616 (vid. pág. 7).
- [7] Jan P. Meier-Kolthoff y Markus Göker. «TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy». En: *Nature Communications* 10.1 (mayo de 2019), pág. 2182 (vid. págs. 7, 9, 15).
- [8] Jan P Meier-Kolthoff, Joaquim Sardà Carbasse, Rosa L Peinado-Olarte y Markus Göker. «TYGS and LPSN: a database tandem for fast and reliable genomebased classification and nomenclature of prokaryotes». En: *Nucleic Acids Research* 50.D1 (ene. de 2022), págs. D801-D807 (vid. págs. 7, 9, 15).

- [9]Fuat Aydin, Secil Abay, Tuba Kayman et al. «Campylobacter anatolicus sp. nov., a novel member of the genus Campylobacter isolated from feces of Anatolian Ground Squirrel (Spermophilus xanthoprymnus) in Turkey». en. En: Systematic and Applied Microbiology 44.6 (nov. de 2021), pág. 126265 (vid. pág. 7).
- [10]W. Miller y E. Yee. «Complete Genome Sequences of the Campylobacter fetus subsp. Venerealis, Campylobacter lari subsp. Concheus, Campylobacter sputorum bv. Sputorum, and Campylobacter volucris Type Strains». En: *Microbiology Resource Announcements* 8.45 (2019). Publisher: American Society for Microbiology (vid. pág. 7).
- [11]Erin Bryant, Zeli Shen, Anthony Mannion et al. «Campylobacter taeniopygiae sp. nov., Campylobacter aviculae sp. nov., and Campylobacter estrildidarum sp. nov., Novel Species Isolated from Laboratory-Maintained Zebra Finches». En: Avian Diseases 64.4 (jun. de 2020) (vid. págs. 7, 8).
- [12]Mirko Rossi, Lies Debruyne, Renato Giulio Zanoni et al. «Campylobacter avium sp. nov., a hippurate-positive species isolated from poultry». En: *International Journal of Systematic and Evolutionary Microbiology* 59.9 (sep. de 2009). Publisher: Microbiology Society, págs. 2364-2369 (vid. pág. 7).
- [13]Canh Phung, Peter C. Scott, Chaitali Dekiwadia, Robert J. Moore y Thi Thu Hao Van. «Campylobacter bilis sp. nov., isolated from chickens with spotty liver disease». En: *International Journal of Systematic and Evolutionary Microbiology* 72.4 (abr. de 2022), pág. 5314 (vid. pág. 7).
- [14]Maarten J Gilbert, Aldert L Zomer, Arjen J Timmerman et al. «Campylobacter blaseri sp. nov., isolated from common seals (Phoca vitulina).» En: *International journal of systematic and evolutionary microbiology* 68.5 (mayo de 2018), págs. 1787-1794 (vid. pág. 7).
- [15]G. Douglas Inglis, Bryanne M. Hoar, Douglas P. Whiteside y Douglas W. Morck.
 «Campylobacter canadensis sp. nov., from captive whooping cranes in Canada».
 En: *International Journal of Systematic and Evolutionary Microbiology* 57.11
 (nov. de 2007). Publisher: Microbiology Society, págs. 2636-2644 (vid. pág. 7).
- [16]LP Doyle. «The etiology of swine dysentery». En: Am. J. Vet. Res 9 (1948), págs. 50-51 (vid. págs. 7, 11).
- [17]M. Veron y R. Chatelain. «Taxonomic Study of the Genus Campylobacter Sebald and Veron and Designation of the Neotype Strain for the Type Species, Campylobacter fetus (Smith and Taylor) Sebald and Veron». En: *International Journal of Systematic Bacteriology* 23.2 (abr. de 1973). ISBN: doi:10.1099/00207713-23-2-122 Publisher: Microbiology Society, págs. 122-134 (vid. págs. 7, 8, 16, 20, 100).

- [18]Monika Koziel, Pat O'Doherty, Peter Vandamme et al. «Campylobacter corcagiensis sp. nov., isolated from faeces of captive lion-tailed macaques (Macaca silenus)». En: International Journal of Systematic and Evolutionary Microbiology 64.PART 8 (ago. de 2014). Publisher: Society for General Microbiology, págs. 2878-2883 (vid. pág. 7).
- [19]R. G. Zanoni, Lies Debruyne, Mirko Rossi, Joana Revez y Peter Vandamme. «Campylobacter cuniculorum sp. nov., from rabbits». En: *International Journal* of Systematic and Evolutionary Microbiology 59.7 (jul. de 2009). Publisher: Microbiology Society, págs. 1666-1671 (vid. pág. 7).
- [20]Peter Vandamme, E. FALSEN, R. ROSSAU et al. «Revision of Campylobacter, Helicobacter, and Wolinella Taxonomy: Emendation of Generic Descriptions and Proposal of Arcobacter gen. nov.» En: *International Journal of Systematic Bacteriology* 41.1 (ene. de 1991), págs. 88-103 (vid. pág. 7).
- [21]T. Smith y M. S. Taylor. «SOME MORPHOLOGICAL AND BIOLOGICAL CHA-RACTERS OF THE SPIRILLA (VIBRIO FETUS, N. SP.) ASSOCIATED WITH DISEASE OF THE FETAL MEMBRANES IN CATTLE.» En: *The Journal of experimental medicine* 30.4 (oct. de 1919). Publisher: The Rockefeller University Press, págs. 299-311 (vid. págs. 7, 10, 14).
- [22]M Sebald y M Veron. «[BASE DNA CONTENT AND CLASSIFICATION OF VIBRIOS].» En: Annales de l'Institut Pasteur 105 (nov. de 1963), págs. 897-910 (vid. págs. 7, 10, 13).
- [23]Alessandra Piccirillo, Giulia Niero, Lucía Calleros et al. «Campylobacter geochelonis sp. nov. isolated from the western Hermann's tortoise (Testudo hermanni hermanni)». En: International Journal of Systematic and Evolutionary Microbiology 66.9 (sep. de 2016). Publisher: Microbiology Society, págs. 3468-3476 (vid. pág. 7).
- [24]Peter Vandamme, M. I. Daneshvar, F. E. Dewhirst et al. «Chemotaxonomic analyses of Bacteroides gracilis and Bacteroides ureolyticus and reclassification of B. gracilis as Campylobacter gracilis comb. nov.» En: *International Journal of Systematic Bacteriology* 45.1 (ene. de 1995). Publisher: Microbiology Society, págs. 145-152 (vid. pág. 7).
- [25]J. Stanley, A. P. Burnens, D. Linton et al. «Campylobacter helveticus sp. nov., a new thermophilic species from domestic animals: Characterization, and cloning of a species-specific DNA probe». En: *Journal of General Microbiology* 138.11 (nov. de 1992). Publisher: Microbiology Society, págs. 2293-2303 (vid. pág. 7).

- [26] Thi Thu Hao Van, Eltaher Elshagmani, Mian Chee Gor, Peter C. Scott y Robert J. Moore. «Campylobacter hepaticus sp. nov., isolated from chickens with spotty liver disease». En: *International Journal of Systematic and Evolutionary Microbiology* 66.11 (nov. de 2016). Publisher: Microbiology Society, págs. 4518-4524 (vid. págs. 7, 12, 144, 149).
- [27] A J Lawson, S L On, J M Logan y J Stanley. «Campylobacter hominis sp. nov., from the human gastrointestinal tract.» En: *International Journal of Systematic and Evolutionary Microbiology* 51.2 (mar. de 2001), págs. 651-660 (vid. pág. 7).
- [28]C. J. Gebhart, P. Edmonds, G. E. Ward, H. J. Kurtz y D. J. Brenner. «'Campylobacter hyointestinalis' sp. nov.: a new species of Campylobacter found in the intestines of pigs and other animals». En: *Journal of Clinical Microbiology* 21.5 (mayo de 1985). ISBN: 0095-1137 Publisher: American Society for Microbiology Journals, págs. 715-720 (vid. págs. 7, 14).
- [29]Maarten J. Gilbert, Marja Kik, William G. Miller, Birgitta Duim y Jaap A. Wagenaar. «Campylobacter iguaniorum sp. nov., isolated from reptiles». En: *International Journal of Systematic and Evolutionary Microbiology* 65.3 (2015). ISBN: 3130253319 (vid. pág. 7).
- [30]Geoffrey Foster, Barry Holmes, Arnold G. Steigerwalt et al. «Campylobacter insulaenigrae sp. nov., isolated from marine mammals». En: *International Journal of Systematic and Evolutionary Microbiology* 54.6 (nov. de 2004). Publisher: Microbiology Society, págs. 2369-2373 (vid. pág. 7).
- [31]F S Jones, M Orcutt y R B Little. «VIBRIOS (VIBRIO JEJUNI, N.SP.) ASSO-CIATED WITH INTESTINAL DISORDERS OF COWS AND CALVES.» En: *The Journal of experimental medicine* 53.6 (mayo de 1931), págs. 853-63 (vid. págs. 7, 11).
- [32]J. M. Logan, A. Burnens, D. Linton, A. J. Lawson y J. Stanley. «Campylobacter lanienae sp. nov., a new species isolated from workers in an abattoir.» En: *International Journal of Systematic and Evolutionary Microbiology* 50.2 (mar. de 2000). Publisher: Microbiology Society, págs. 865-872 (vid. pág. 7).
- [33]J. Benjamin, S. Leaper, R. J. Owen y M. B. Skirrow. «Description of Campylobacter laridis, a new species comprising the nalidixic acid resistant thermophilicCampylobacter (NARTC) group». En: *Current Microbiology* 8.4 (jul. de 1983). Publisher: Springer-Verlag, págs. 231-238 (vid. pág. 8).
- [34]Caoimhe Lynch, Charlotte Peeters, Niamh Walsh et al. «Campylobacter majalis sp. nov. and Campylobacter suis sp. nov., novel Campylobacter species isolated from porcine gastrointestinal mucosa». En: *International Journal of Systematic and Evolutionary Microbiology* 72.12 (dic. de 2022) (vid. pág. 8).

- [35]A Antezack, M Boxberger, C Rolland et al. «Isolation and characterization of Campylobacter massiliensis sp. nov., a novel Campylobacter species detected in a gingivitis subject.» En: 71.10 (), pág. 5039 (vid. pág. 8).
- [36]G. H. K. LAWSON, J. L. LEAVER, G. W. PETTIGREW y A. C. ROWLAND. «Some Features of Campylobacter sputorum subsp. mucosalis subsp. nov., nom. rev. and Their Taxonomic Significance». En: *International Journal of Systematic Bacteriology* 31.4 (oct. de 1981), págs. 385-391 (vid. pág. 8).
- [37]R. M. Roop, R. M. Smibert, J. L. Johnson y N. R. Krieg. «Campylobacter mucosalis (Lawson, Leaver, Pettigrew, and Rowland 1981) comb. nov.: Emended Description». En: *International Journal of Systematic Bacteriology* 35.2 (abr. de 1985), págs. 189-192 (vid. pág. 8).
- [38]Samuel Bloomfield, David Wilkinson, Lynn Rogers et al. «Campylobacter novaezeelandiae sp. nov., isolated from birds and water in New Zealand». En: *International Journal of Systematic and Evolutionary Microbiology* 70.6 (jun. de 2020), págs. 3775-3784 (vid. pág. 8).
- [39]Alberto Caceres, Ivo Muñoz, Gregorio Iraola, Florencia Díaz-Viraqué y Luis Collado. «Campylobacter ornithocola sp. nov., a new member of the Campylobacter lari group isolated from wild bird faecal samples». En: *International Journal of Systematic and Evolutionary Microbiology* (ene. de 2017) (vid. pág. 8).
- [40]Lies Debruyne, Stephen L. W. On, Evie De Brandt y Peter Vandamme. «Novel Campylobacter lari-like bacteria from humans and molluscs: description of Campylobacter peloridis sp. nov., Campylobacter lari subsp. concheus subsp. nov. and Campylobacter lari subsp. lari subsp. nov.» En: *International Journal* of Systematic and Evolutionary Microbiology 59.Pt 5 (mayo de 2009). ISBN: 1466-5026 (Print)\r1466-5026 (Linking) Publisher: Microbiology Society, págs. 1126-32 (vid. pág. 8).
- [41]Maarten J Gilbert, William G Miller, Judy St Leger et al. «Campylobacter pinnipediorum sp. nov., isolated from pinnipeds, comprising Campylobacter pinnipediorum subsp. pinnipediorum subsp. nov. and Campylobacter pinnipediorum subsp. caledonicus subsp. nov.» En: *International journal of systematic and evolutionary microbiology* 67.6 (jun. de 2017), págs. 1961-1968 (vid. pág. 8).
- [42]Marta F. Silva, Ana L. Pereira, Maria J. Fraqueza et al. «Genomic and phenotypic characterization of campylobacter fetus subsp. Venerealis strains». En: *Microorganisms* 9.2 (2021), págs. 1-14 (vid. pág. 8).

- [43]Peter Vandamme, Floyd E. Dewhirst, Bruce J. Paster y Stephen L.W. On. «Campylobacter». En: Bergey's Manual of Systematics of Archaea and Bacteria. Chichester, UK: John Wiley & Sons, Ltd, sep. de 2015, págs. 1-27 (vid. págs. 8, 11).
- [44]Y. ETOH, F. E. DEWHIRST, B. J. PASTER, A. YAMAMOTO y N. GOTO. «Campylobacter showae sp. nov., Isolated from the Human Oral Cavity». En: *International Journal of Systematic Bacteriology* 43.4 (oct. de 1993), págs. 631-639 (vid. pág. 8).
- [45]A.R. Prévot. Manual de classification et de Determination des Bacteries Aanaerobies, 1st edition. Paris: Masson y Co., 1940 (vid. págs. 8, 11).
- [46]Lies Debruyne, Tina Broman, Sven Bergström et al. «Campylobacter volucris sp. nov., isolated from black-headed gulls (Larus ridibundus)». En: International Journal of Systematic and Evolutionary Microbiology 60.8 (ago. de 2010). Publisher: Microbiology Society, págs. 1870-1875 (vid. pág. 8).
- [47]Karin Sandstedt y Jan Ursing. «Description of Campylobacter upsaliensis sp. nov. Previously Known as the CNW Group». En: Systematic and Applied Microbiology 14.1 (ene. de 1991). Publisher: Urban & Fischer, págs. 39-45 (vid. pág. 8).
- [48]F. L. JACKSON e Y. E. GOODMAN. «Bacteroides ureolyticus, a New Species to Accommodate Strains Previously Identified as "Bacteroides corrodens, Anaerobic"». En: *International Journal of Systematic Bacteriology* 28.2 (abr. de 1978), págs. 197-200 (vid. pág. 8).
- [49]Peter Vandamme, L. Debruyne, E. De Brandt y E. Falsen. «Reclassification of Bacteroides ureolyticus as Campylobacter ureolyticus comb. nov., and emended description of the genus Campylobacter». En: *International Journal* of Systematic and Evolutionary Microbiology 60.9 (sep. de 2010). Publisher: Microbiology Society, págs. 2016-2022 (vid. pág. 8).
- [50]Antonio Parisi, Matteo Chiara, Monica Caffara et al. «Campylobacter vulpis sp. nov. isolated from wild red foxes». En: *Systematic and Applied Microbiology* 44.3 (mayo de 2021), pág. 126204 (vid. pág. 8).
- [51]M A Karmali e Y. C. Tan. «Neonatal campylobacter enteritis». En: Canadian Medical Association Journal 122.2 (jun. de 1980). Publisher: Canadian Medical Association ISBN: 0008-4409 (Print) 0008-4409 (Linking), págs. 192-193 (vid. pág. 10).
- [52]Steffen Backert, ed. *Fighting Campylobacter Infections: Towards a One Health Approach.* en. Vol. 431. Current Topics in Microbiology and Immunology. Cham: Springer International Publishing, 2021 (vid. pág. 10).

- [53]J.-P. Butzler. «Campylobacter, from obscurity to celebrity». en. En: *Clinical Microbiology and Infection* 10.10 (oct. de 2004), págs. 868-876 (vid. pág. 10).
- [54]Sati Samuel Ngulukun. «Chapter 3 Taxonomy and physiological characteristics of Campylobacter spp.» En: *Campylobacter*. 2017, págs. 41-60 (vid. pág. 10).
- [55]A. J. LEVY. «A gastro-enteritis cutbreak probably due to a bovine strain of vibrio.» En: *The Yale journal of biology and medicine* 18.4 (mar. de 1946). ISBN: 0044-0086 Publisher: Yale Journal of Biology and Medicine, págs. 243-58 (vid. pág. 11).
- [56]Albert J. Lastovica, Stephen L. W. On y Li Zhang. «The Family Campylobacteraceae». En: *The Prokaryotes*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, págs. 307-335 (vid. pág. 11).
- [57]Si Ming Man. «The clinical importance of emerging Campylobacter species». En: *Nature Reviews Gastroenterology & Hepatology* 8.12 (dic. de 2011). ISBN: 1759-5053 (Electronic)\r1759-5045 (Linking) Publisher: Nature Publishing Group, págs. 669-685 (vid. págs. 12, 17).
- [58]Daniela Costa y Gregorio Iraola. «Pathogenomics of Emerging Campylobacter Species». En: *Clinical Microbiology Reviews* 32.4 (sep. de 2019). Publisher: American Society for Microbiology Journals, e00072-18 (vid. págs. 13, 15, 144, 151).
- [59] Jérôme Pacanowski, Valérie Lalande, Karine Lacombe et al. «Campylobacter Bacteremia: Clinical Features and Factors Associated with Fatal Outcome». en. En: Clinical Infectious Diseases 47.6 (sep. de 2008), págs. 790-796 (vid. pág. 13).
- [60]M. B. Skirrow. «John McFadyean and the Centenary of the First Isolation of Campylobacter Species». En: *Clinical Infectious Diseases* 43.9 (nov. de 2006).
 Publisher: Oxford University Press, págs. 1213-1217 (vid. pág. 14).
- [61]A Florent. «Les deux vibriosis genitals: la vibriose due a». En: V. fetus venerealis (1959), págs. 1-60 (vid. págs. 14, 20, 99).
- [62] «Campylobacteriosis genital bovina». En: *Manual terrestre de la OIE*. Capítulo 3.4.4. OIE, 2021 (vid. págs. 14, 20, 21, 28, 99, 101).
- [63]Daniela Costa, Simon Lévesque, Nitin Kumar et al. «Pangenome analysis reveals genetic isolation in Campylobacter hyointestinalis subspecies adapted to different mammalian hosts». En: *Scientific Reports* 11.1 (dic. de 2021). Publisher: Nature Publishing Group ISBN: 0123456789, pág. 3431 (vid. pág. 14).

- [64]Collette Fitzgerald, Zheng Chao Tu, Mary Patrick et al. «Campylobacter fetus subsp. testudinum subsp. nov., isolated from humans and reptiles». En: *International Journal of Systematic and Evolutionary Microbiology* 64.Pt_9 (sep. de 2014). Publisher: Microbiology Society, págs. 2944-2948 (vid. pág. 16).
- [65]Kate E. Dingle, Martin J. Blaser, Zheng Chao Tu et al. «Genetic relationships among reptilian and mammalian Campylobacter fetus strains determined by multilocus sequence typing». En: *Journal of Clinical Microbiology* 48.3 (2010). ISBN: 0095-1137, págs. 977-980 (vid. pág. 16).
- [66]Z. C. Tu, F. E. Dewhirst y M. J. Blaser. «Evidence that the Campylobacter fetus sap locus is an ancient genomic constituent with origins before mammals and reptiles diverged». En: *Infection and Immunity* 69.4 (abr. de 2001). Publisher: American Society for Microbiology, págs. 2237-2244 (vid. pág. 16).
- [67]S Harvey y J R Greenwood. «Isolation of Campylobacter fetus from a pet turtle». en. En: *Journal of Clinical Microbiology* 21.2 (feb. de 1985), págs. 260-261 (vid. pág. 16).
- [68]Zheng-Chao Tu, John Hui y Martin J. Blaser. «Conservation and Diversity of *sap* Homologues and Their Organization among *Campylobacter fetus* Isolates». En: *Infection and Immunity* 72.3 (mar. de 2004), págs. 1715-1724 (vid. pág. 16).
- [69]Nicodemus M. Masila, Kirstin E. Ross, Michael G. Gardner y Harriet Whiley. «Zoonotic and Public Health Implications of Campylobacter Species and Squamates (Lizards, Snakes and Amphisbaenians)». en. En: *Pathogens* 9.10 (sep. de 2020), pág. 799 (vid. pág. 16).
- [70]H. Sprenger, Ellen L. Zechner y Gregor Gorkiewicz. «So close and yet so far — Molecular microbiology of Campylobacter fetus subspecies». En: *European Journal of Microbiology and Immunology* 2.1 (mar. de 2012). Publisher: Akademiai Kiado Zrt., págs. 66-75 (vid. pág. 17).
- [71]Caroline da Silva Silveira, Martin Fraga, Federico Giannitti, Melissa Macías-Rioseco y Franklin Riet-Correa. «Diagnosis of Bovine Genital Campylobacteriosis in South America». En: Frontiers in Veterinary Science 5 (dic. de 2018), pág. 321 (vid. pág. 17).
- [72]Xavier Marchand-Senécal, Sadjia Bekal, Pierre A. Pilon, Jean-Loup Sylvestre y Christiane Gaudreau. «Campylobacter fetus Cluster Among Men Who Have Sex With Men, Montreal, Quebec, Canada, 2014–2016». En: *Clinical Infectious Diseases* 65.10 (oct. de 2017). Publisher: Oxford University Press, págs. 1751-1753 (vid. págs. 17, 227, 236, 247).

- [73]Jaap A. Wagenaar, M. A. P. van Bergen, Martin J. Blaser et al. «Campylobacter fetus infections in humans: Exposure and disease». En: *Clinical Infectious Diseases* 58.11 (jun. de 2014). Publisher: Oxford University Press, págs. 1579-86 (vid. págs. 17, 227, 236, 247).
- [74]Aubrey N. Michi, Pedro H. Favetto, John Kastelic y Eduardo R. Cobo. «A review of sexually transmitted bovine trichomoniasis and campylobacteriosis affecting cattle reproductive health». en. En: *Theriogenology* 85.5 (mar. de 2016), págs. 781-791 (vid. pág. 18).
- [75]G.D. Schurig, C.E. Hall, K. Burda et al. «Infection patterns in heifers following cervicovaginal or intrauterine instillation of Campylobacter (Vibrio) fetus venerealis.» En: *The Cornell veterinarian* 64.4 (1974), págs. 533-548 (vid. pág. 18).
- [76]I. W. Moynihan y P. L. Stovell. «Vibriosis In Cattle». eng. En: Canadian Journal of Comparative Medicine and Veterinary Science 19.4 (abr. de 1955), págs. 105-112 (vid. pág. 18).
- [77] Jessica B. Rush y Misty A. Edmondson. «Infectious Agents: Campylobacter». en. En: *Bovine Reproduction*. Ed. por Richard M. Hopper. 1.^a ed. Wiley, ago. de 2021, págs. 717-724 (vid. pág. 18).
- [78]A. Cipolla, A. Casaro, H. Terzolo et al. «Persistence of Campylobacter fetus subspecies venerealis in experimentally infected heifers». en. En: *Veterinary Record* 134.24 (jun. de 1994), págs. 628-628 (vid. pág. 18).
- [79]Patrick C.-Y. Woo, Kit-Wah Leung, Hoi-Wah Tsoi et al. «Thermo-tolerant Campylobacter fetus bacteraemia identified by 16S ribosomal RNA gene sequencing: an emerging pathogen in immunocompromised patients». en. En: *Journal of Medical Microbiology* 51.9 (sep. de 2002), págs. 740-746 (vid. pág. 20).
- [80] Josely Ferreira Figueiredo, Aiesca Oliveira Pellegrin, Cid Bastos Fóscolo et al. «Evaluation of direct fluorescent antibody test for the diagnosis of bovine genital campylobacteriosis». eng. En: *Revista Latinoamericana De Microbiologia* 44.3-4 (2002), págs. 118-123 (vid. pág. 21).
- [81]C. M. Campero, M. L. Anderson, R. L. Walker et al. «Immunohistochemical Identification of Campylobacter fetus in Natural Cases of Bovine and Ovine Abortions». En: *Journal of Veterinary Medicine Series B* 52.3 (abr. de 2005), págs. 138-141 (vid. pág. 21).
- [82]G. D. Mshelia, J. D. Amin, Z. Woldehiwet, R. D. Murray y G. O. Egwu. «Epidemiology of Bovine Venereal Campylobacteriosis: Geographic Distribution and Recent Advances in Molecular Diagnostic Techniques». En: *Reproduction in Domestic Animals* 45.5 (nov. de 2009), e221-30 (vid. pág. 21).

- [83]Gregor Gorkiewicz, Gebhard Feierl, Caroline Schober et al. «Species-Specific Identification of Campylobacters by Partial 16S rRNA Gene Sequencing». En: *Journal of Clinical Microbiology* 41.6 (jun. de 2003), págs. 2537-2546 (vid. pág. 21).
- [84]Birgitta Duim, Peter A. R Vandamme, Alan Rigter et al. «Differentiation of Campylobacter species by AFLP fingerprinting». en. En: *Microbiology* 147.10 (oct. de 2001), págs. 2729-2737 (vid. pág. 22).
- [85]Linda Van der Graaf-Van Bloois, Marcel a P. van Bergen, Fimme J. van der Wal et al. «Evaluation of molecular assays for identification Campylobacter fetus species and subspecies and development of a C. fetus specific real-time PCR assay». En: *Journal of Microbiological Methods* 95.1 (2013). ISBN: 0167-7012 Publisher: Elsevier B.V., págs. 93-97 (vid. págs. 22, 26, 148, 149).
- [86]S. M. Salama, H. Tabor, M. Richter y D. E. Taylor. «Pulsed-field gel electrophoresis for epidemiologic studies of Campylobacter hyointestinalis isolates». En: *Journal of Clinical Microbiology* 30.8 (ago. de 1992). Publisher: American Society for Microbiology Journals, págs. 1982-1984 (vid. pág. 22).
- [87]S. Hum, K. Quinn, J. Brunner y Slw On. «Evaluation of a PCR assay for identification and differentiation of *Campylobacter fetus* subspecies». en. En: *Australian Veterinary Journal* 75.11 (nov. de 1997), págs. 827-831 (vid. págs. 22, 24-26).
- [88]S.L.W. On y C.S. Harrington. «Evaluation of numerical analysis of PFGE-DNA profiles for differentiating Campylobacter fetus subspecies by comparison with phenotypic, PCR and 16S rDNA sequencing methods». en. En: *Journal of Applied Microbiology* 90.2 (feb. de 2001), págs. 285-293 (vid. pág. 22).
- [89]M. C. J. Maiden, Jane A. Bygraves, Edward Feil et al. «Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms». En: *Proceedings of the National Academy of Sciences* 95.6 (mar. de 1998). Publisher: National Academy of Sciences, págs. 3140-3145 (vid. pág. 23).
- [90]Marcel A. P. Van Bergen, Kate E. Dingle, MartinC C. J. Maiden et al. «Clonal Nature of Campylobacter fetus as Defined by Multilocus Sequence Typing». En: *Journal of Clinical Microbiology* 43.12 (dic. de 2005), págs. 5888-5898 (vid. págs. 23, 99, 100, 215, 217).
- [91]K. E. Dingle, F. M. Colles, D. R. A. Wareing et al. «Multilocus Sequence Typing System for *Campylobacter jejuni*». en. En: *Journal of Clinical Microbiology* 39.1 (ene. de 2001), págs. 14-23 (vid. pág. 23).

- [92]Keith A. Jolley, James E. Bray y Martin C. J. Maiden. «Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications». En: *Wellcome Open Research* 3 (sep. de 2018). Publisher: F1000 Research Limited, pág. 124 (vid. págs. 23, 106, 230).
- [93]K Blom, C M Patton, M A Nicholson y B Swaminathan. «Identification of Campylobacter fetus by PCR-DNA probe method». en. En: *Journal of Clinical Microbiology* 33.5 (mayo de 1995), págs. 1360-1362 (vid. pág. 23).
- [94]O.A. Oyarzabal, I.V. Wesley, K.M. Harmon et al. «Specific identification of Campylobacter fetus by PCR targeting variable regions of the 16S rDNA». en. En: Veterinary Microbiology 58.1 (oct. de 1997), págs. 61-71 (vid. pág. 23).
- [95]R. P. Spence, I. R. Bruce, A. M. J. McFadden et al. «Cross-reaction of a *Campy-lobacter fetus* subspecies *venerealis* real-time PCR». en. En: *Veterinary Record* 168.5 (feb. de 2011), págs. 131-131 (vid. págs. 24, 25).
- [96]G. Wang, C. G. Clark, T. M. Taylor et al. «Colony Multiplex PCR Assay for Identification and Differentiation of Campylobacter jejuni, C. coli, C. lari, C. upsaliensis, and C. fetus subsp. fetus». En: *Journal of Clinical Microbiology* 40.12 (dic. de 2002), págs. 4744-4747 (vid. págs. 24, 25).
- [97]M. A. Van Bergen, Guus Simmons, Linda Van der Graaf-Van Bloois et al. «Amplified fragment length polymorphism based identification of genetic markers and novel PCR assay for differentiation of Campylobacter fetus subspecies». en. En: *Journal of Medical Microbiology* 54.12 (dic. de 2005), págs. 1217-1224 (vid. págs. 24, 25).
- [98]Lyle McMillen, Geoffry Fordyce, Vivienne J. Doogan y Ala E. Lew. «Comparison of culture and a novel 5' Taq nuclease assay for direct detection of Campylobacter fetus subsp. venerealis in clinical specimens from cattle.» En: *Journal of clinical microbiology* 44.3 (mar. de 2006), págs. 938-45 (vid. pág. 25).
- [99]Carlos Abril, E. M. M. Vilei, I. Brodard et al. «Discovery of insertion element ISCfe1: a new tool for Campylobacter fetus subspecies differentiation». En: *Clinical Microbiology and Infection* 13.10 (oct. de 2007). Publisher: Elsevier, págs. 993-1000 (vid. págs. 25, 26, 99, 148, 149, 221).
- [100]Gregorio Iraola, Martín Hernández, Lucía Calleros et al. «Application of a multiplex PCR assay for *Campylobacter fetus* detection and subspecies differentiation in uncultured samples of aborted bovine fetuses». en. En: *Journal* of Veterinary Science 13.4 (2012), pág. 371 (vid. págs. 26, 99, 230).

- [101]Paula M. Moolhuijzen, Ala E. Lew-Tabor, Bartosz M. Wlodek et al. «Genomic analysis of Campylobacter fetus subspecies: identification of candidate virulence determinants and diagnostic assay targets.» En: *BMC Microbiology* 9.1 (ene. de 2009). Publisher: BioMed Central, pág. 86 (vid. págs. 26, 99, 230).
- [102]Gregorio Iraola, Ruben Pérez, Laura Betancor et al. «A novel real-time PCR assay for quantitative detection of Campylobacter fetus based on ribosomal sequences». En: *BMC Veterinary Research* 12.1 (dic. de 2016), pág. 286 (vid. págs. 26, 27, 230).
- [103]World Organisation for Animal Health. *Terrestrial animal health code*. Paris, 2008 (vid. pág. 27).
- [104]M. V. Repiso, A. Gil, P. Bañales et al. «Prevalencia de las principales enfermedades infecciosas que afectan el comportamiento reproductivo en la ganadería de carne y caracterización de los establecimientos de cría del Uruguay». En: *Veterinaria* 40.157 (2005), págs. 5-28 (vid. pág. 28).
- [105]*Informe sector Ganadero*. Inf. téc. Uruguay: Uruguay XXI, oct. de 2022 (vid. pág. 29).
- [106]Alejandra Suanes, Valentina Macchi, Federico Fernandez et al. «Reproductive, health and management characteristics in dairy herds in Uruguay». En: *Veterinaria (Montevideo)* 47.215 (abr. de 2021) (vid. pág. 30).
- [107]Robert D. Fleischmann, Mark D. Adams, Owen White et al. «Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd». en. En: *Science* 269.5223 (jul. de 1995), págs. 496-512 (vid. pág. 54).
- [108]Duccio Medini, Davide Serruto, Julian Parkhill et al. «Microbiology in the post-genomic era». en. En: *Nature Reviews Microbiology* 6.6 (jun. de 2008), págs. 419-430 (vid. págs. 54, 97).
- [109]Martin Kircher y Janet Kelso. «High-throughput DNA sequencing concepts and limitations». en. En: *BioEssays* 32.6 (2010). _eprint: https://onlinelibrary.wiley.com/doi/ págs. 524-536 (vid. pág. 56).
- [110]Oxford Nanopore Technologies (vid. pág. 59).
- [111]Clara Delahaye y Jacques Nicolas. «Sequencing DNA with nanopores: Troubles and biases». eng. En: *PloS One* 16.10 (2021), e0257521 (vid. pág. 59).
- [112]PacBio Sequence with confidence. en-US (vid. pág. 60).
- [113]Simon C Andrews, Lindenbaum, Pierre, Howard, Brian y Ewels, Phil. *FastQC*. 2010 (vid. págs. 61, 72).

- [114]Niranjan Nagarajan y Mihai Pop. «Sequence assembly demystified». en. En: *Nature Reviews Genetics* 14.3 (mar. de 2013). Number: 3 Publisher: Nature Publishing Group, págs. 157-167 (vid. pág. 61).
- [115]Jang-il Sohn y Jin-Wu Nam. «The present and future of de novo whole-genome assembly». En: *Briefings in Bioinformatics* 19.1 (ene. de 2018), págs. 23-40 (vid. págs. 62, 64).
- [116]Anton Bankevich, Sergey Nurk, Dmitry Antipov et al. «SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing». En: *Journal of Computational Biology* 19.5 (mayo de 2012). Publisher: Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, págs. 455-477 (vid. págs. 62, 73).
- [117]Nicola De Maio, Liam P. Shaw, Alasdair Hubbard et al. «Comparison of longread sequencing technologies in the hybrid assembly of complex bacterial genomes». En: *Microbial Genomics* 5.9 (2019). Publisher: Microbiology Society, e000294 (vid. pág. 62).
- [118]Ryan R. Wick, Louise M. Judd, Claire L. Gorrie y Kathryn E. Holt. «Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads». En: *PLOS Computational Biology* 13.6 (jun. de 2017). Ed. por Adam M. Phillippy. Publisher: Public Library of Science, e1005595 (vid. págs. 63, 73).
- [119]Ryan R. Wick, Louise M. Judd y Kathryn E. Holt. «Assembling the perfect bacterial genome using Oxford Nanopore and Illumina sequencing». en. En: *PLOS Computational Biology* 19.3 (mar. de 2023). Publisher: Public Library of Science, e1010905 (vid. pág. 64).
- [120]Andrew J. Page, Nishadi De Silva, Martin Hunt et al. «Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data». En: *Microbial Genomics* 2.8 (ago. de 2016). Publisher: Microbiology Society, e000083 (vid. págs. 65, 73).
- [121]Marten Boetzer, Christiaan V. Henkel, Hans J. Jansen, Derek Butler y Walter Pirovano. «Scaffolding pre-assembled contigs using SSPACE». En: *Bioinformatics* 27.4 (feb. de 2011). Publisher: Oxford University Press, págs. 578-579 (vid. págs. 65, 73).
- [122]Marten Boetzer y Walter Pirovano. «Toward almost closed genomes with GapFiller». En: *Genome Biology* 13.6 (jun. de 2012). Publisher: BioMed Central, R56 (vid. págs. 65, 73).
- [123] Theo H. M. Smits. «The importance of genome sequence quality to microbial comparative genomics». En: *BMC Genomics* 20.1 (ago. de 2019), pág. 662 (vid. págs. 65, 66, 90).

- [124] Patricia Sieber, Matthias Platzer y Stefan Schuster. «The Definition of Open Reading Frame Revisited». English. En: *Trends in Genetics* 34.3 (mar. de 2018).
 Publisher: Elsevier, págs. 167-170 (vid. pág. 66).
- [125]Torsten Seemann. «Prokka: Rapid prokaryotic genome annotation». En: *Bio-informatics* 30.14 (jul. de 2014). ISBN: 1367-4811 (Electronic) Publisher: Oxford University Press, págs. 2068-2069 (vid. págs. 66, 74).
- [126]Wenjun Li, Kathleen R O'Neill, Daniel H Haft et al. «RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation». En: *Nucleic Acids Research* 49.D1 (ene. de 2021), págs. D1020-D1028 (vid. pág. 66).
- [127]Oliver Schwengers, Lukas Jelonek, Marius Alfred Dieckmann et al. «Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification». eng. En: *Microbial Genomics* 7.11 (nov. de 2021), pág. 000685 (vid. pág. 66).
- [128]Doug Hyatt, Gwo Liang Chen, Philip F. LoCascio et al. «Prodigal: Prokaryotic gene recognition and translation initiation site identification». En: *BMC Bioinformatics* 11.1 (mar. de 2010). Publisher: BioMed Central, págs. 1-11 (vid. pág. 66).
- [129]Karin Lagesen, Peter Hallin, Einar Andreas Rødland et al. «RNAmmer: consistent and rapid annotation of ribosomal RNA genes». En: *Nucleic Acids Research* 35.9 (mayo de 2007), págs. 3100-3108 (vid. pág. 66).
- [130]Dean Laslett y Bjorn Canback. «ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences». En: *Nucleic Acids Research* 32.1 (2004), págs. 11-16 (vid. pág. 66).
- [131]Eric P. Nawrocki y Sean R. Eddy. «Infernal 1.1: 100-fold faster RNA homology searches». En: *Bioinformatics* 29.22 (nov. de 2013), págs. 2933-2935 (vid. pág. 66).
- [132]E. J. Richardson y M. Watson. «The automatic annotation of bacterial genomes». en. En: *Briefings in Bioinformatics* 14.1 (ene. de 2013), págs. 1-12 (vid. pág. 66).
- [133]Steven L. Salzberg. «Next-generation genome annotation: we still struggle to get it right». En: *Genome Biology* 20.1 (mayo de 2019), pág. 92 (vid. pág. 66).
- [134]R. W. A. Park. «Observations on the ability of two biochemical types of Vibrio fetus to pro-liferate in the genital tract of cattle and their importance with respect to infertility». En: *British Veterinary Journal* 118 (1962), págs. 411-420 (vid. pág. 69).

- [135]A. C. Chávez, S. Barrera, A. Leon y G. Trueba. «Campylobacter fetus Bacteremia in a Healthy Patient Returning from a Trip to the Ecuadorian Amazonia». En: *Zoonoses and Public Health* 64.5 (2017), págs. 391-393 (vid. págs. 69, 76, 88).
- [136]Sachiko Kitamura, Noriomi Matsumura, Noriko Ohtake, Masato Kita e Ikuo Konishi. «Tubo-ovarian abscess with endometrial cyst probably infected by Campylobacter fetus: Two cases». En: *Journal of Obstetrics and Gynaecology Research* 42.8 (2016), págs. 1052-1057 (vid. págs. 69, 76, 88).
- [137]Daniela Costa, Virginia Aráoz, Maila Barcellos et al. «Complete Genome Sequence of Campylobacter fetus Isolated from a Sheep». En: *Microbiology Resource Announcements* 9.45 (nov. de 2020). Ed. por Irene L G Newton. Publisher: American Society for Microbiology, e01008-20 (vid. págs. 70, 92, 93, 111, 244).
- [138]Daehwan Kim, Li Song, Florian P. Breitwieser y Steven L. Salzberg. «Centrifuge: rapid and sensitive classification of metagenomic sequences». En: *Genome Research* 26.12 (dic. de 2016). Publisher: Cold Spring Harbor Laboratory Press, págs. 1721-1729 (vid. pág. 72).
- [139]Anthony M. Bolger, Marc Lohse y Bjoern Usadel. Trimmomatic. 2014 (vid. pág. 72).
- [140]Alexey Gurevich, Vladislav Saveliev, Nikolay Vyahhi y Glenn Tesler. «QUAST: quality assessment tool for genome assemblies». En: *Bioinformatics* 29.8 (abr. de 2013). Publisher: Oxford University Press, págs. 1072-1075 (vid. pág. 74).
- [141]Donovan H. Parks, Michael Imelfort, Connor T. Skennerton, Philip Hugenholtz y Gene W. Tyson. «CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes». En: *Genome Research* 25.7 (jul. de 2015). Publisher: Cold Spring Harbor Laboratory Press, págs. 1043-1055 (vid. págs. 74, 91).
- [142]A. C. Frank y J. R. Lobry. «Oriloc: Prediction of replication boundaries in unannotated bacterial chromosomes». En: *Bioinformatics* 16.6 (jun. de 2000).
 Publisher: Oxford Academic, págs. 560-561 (vid. págs. 74, 82).
- [143]Delphine Charif y Jean R. Lobry. «SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis». En: *Structural Approaches to Sequence Evolution*. Ed. por Elias Greenbaum, Ugo Bastolla, Markus Porto, H. Eduardo Roman y Michele Vendruscolo. Series Title: Biological and Medical Physics, Biomedical Engineering. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, págs. 207-232 (vid. pág. 74).
- [144]Anna I. Rissman, Bob Mau, Bryan S. Biehl et al. «Reordering contigs of draft genomes using the Mauve Aligner». En: *Bioinformatics* 25.16 (ago. de 2009). Publisher: Oxford Academic, págs. 2071-2073 (vid. pág. 74).
- [145] Thomas Madden y George Coulouris. «BLAST Command Line Applications User Manual BLAST Command Line Applications User Manual - BLAST ®» En: Md (2008), págs. 1-28 (vid. págs. 75, 253).
- [146]Christiam Camacho, George Coulouris, Vahram Avagyan et al. «BLAST+: architecture and applications». En: *BMC Bioinformatics* 10.1 (dic. de 2009), pág. 421 (vid. págs. 75, 108, 168).
- [147] Tim Carver, Matthew Berriman, Adrian Tivey et al. «Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database». en. En: *Bioinformatics* 24.23 (dic. de 2008), págs. 2672-2676 (vid. pág. 75).
- [148]Tim Carver, Simon R. Harris, Matthew Berriman, Julian Parkhill y Jacqueline A. McQuillan. «Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data». en. En: *Bioinformatics* 28.4 (feb. de 2012), págs. 464-469 (vid. pág. 75).
- [149]Matías A. Dorsch, María L. Casaux, Lucía Calleros et al. «Placentitis and abortion caused by a multidrug resistant strain of Campylobacter fetus subspecies fetus in a sheep in Uruguay». en. En: *Revista Argentina de Microbiología* 54.1 (ene. de 2022), págs. 25-30 (vid. pág. 81).
- [150]Magdalena Rajewska, Katarzyna Wegrzyn e Igor Konieczny. «AT-rich region and repeated sequences – the essential elements of replication origins of bacterial replicons». En: *FEMS Microbiology Reviews* 36.2 (mar. de 2012). Publisher: Oxford Academic, págs. 408-434 (vid. pág. 82).
- [151]Wei Hien Cheong, Yung Chie Tan, Soon Joo Yap y Kee Peng Ng. «ClicO FS: an interactive web-based service of Circos». En: *Bioinformatics* 31.22 (nov. de 2015). Publisher: Oxford Academic, págs. 3685-3687 (vid. pág. 84).
- [152]Sameeh M. Salama, Elizabeth Newnham, Nicholas Chang y Diane E. Taylor. «Genome map of Campylobacter fetus subsp. fetus ATCC 27374». en. En: *FEMS Microbiology Letters* 132.3 (oct. de 1995), págs. 239-245 (vid. pág. 91).
- [153]G. Iraola, L. Betancor, L. Calleros et al. «A rural worker infected with a bovineprevalent genotype of Campylobacter fetus subsp. fetus supports zoonotic transmission and inconsistency of MLST and whole-genome typing». En: *European Journal of Clinical Microbiology & Infectious Diseases* 34.8 (ago. de 2015), págs. 1593-1596 (vid. págs. 92, 145).

- [154]Gregorio Iraola, Samuel C. Forster, Nitin Kumar et al. «Distinct Campylobacter fetus lineages adapted as livestock pathogens and human pathobionts in the intestinal microbiota». En: *Nature Communications* 8.1 (dic. de 2017). Publisher: Nature Publishing Group, pág. 1367 (vid. págs. 92, 144, 147, 150, 151, 215-217, 227, 236, 245).
- [155]Chirag Jain, Luis M. Rodriguez-R, Adam M. Phillippy, Konstantinos T. Konstantinidis y Srinivas Aluru. «High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries». en. En: *Nature Communications* 9.1 (nov. de 2018), pág. 5114 (vid. págs. 97, 105).
- [156]H Tettelin, V. Masignani, M. J. Cieslewicz et al. «Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: Implications for the microbial "pan-genome"». En: *Proceedings of the National Academy of Sciences* 102.39 (sep. de 2005), págs. 13950-13955 (vid. pág. 98).
- [157] Jaap A. Wagenaar, Marcel A. P. Van Bergen, Diane G. Newell, Rose Grogono-Thomas y Birgitta Duim. «Comparative Study Using Amplified Fragment Length Polymorphism Fingerprinting, PCR Genotyping, and Phenotyping To Differentiate Campylobacter fetus Strains Isolated from Animals». En: *Journal* of Clinical Microbiology 39.6 (jun. de 2001), págs. 2283-2286 (vid. pág. 99).
- [158]Eric W. Sayers, Evan E. Bolton, J. Rodney Brister et al. «Database resources of the national center for biotechnology information». eng. En: *Nucleic Acids Research* 50.D1 (ene. de 2022), págs. D20-D26 (vid. pág. 105).
- [159]Alice R. Wattam, David Abraham, Oral Dalay et al. «PATRIC, the bacterial bioinformatics database and analysis resource». En: *Nucleic Acids Research* 42.D1 (2014). ISBN: 0305-1048 (vid. pág. 105).
- [160]Alice R. Wattam, James J. Davis, Rida Assaf et al. «Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center». En: *Nucleic Acids Research* 45.D1 (ene. de 2017), págs. D535-D542 (vid. pág. 105).
- [161]Samuel K. Sheppard, Keith A. Jolley y Martin C. J. Maiden. «A Gene-By-Gene Approach to Bacterial Population Genomics: Whole Genome MLST of Campylobacter». En: *Genes* 3.2 (abr. de 2012). Publisher: Molecular Diversity Preservation International, págs. 261-277 (vid. págs. 106, 230).
- [162]Ignacio Ferrés y Gregorio Iraola. «MLSTar: automatic multilocus sequence typing of bacterial genomes in R». En: *PeerJ* 6.6 (jun. de 2018). Publisher: PeerJ Inc., e5098 (vid. págs. 106, 231, 253).

- [163]Todd J. Treangen, Brian D. Ondov, Sergey Koren y Adam M. Phillippy. «The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes». En: *Genome Biology* 15.11 (2014). ISBN: 1465-6906 (vid. págs. 106, 231, 253).
- [164]Nicholas J. Croucher, Andrew J. Page, Thomas R. Connor et al. «Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins». En: *Nucleic Acids Research* 43.3 (feb. de 2015). Publisher: Oxford University Press, e15-e15 (vid. págs. 106, 231, 253).
- [165]Alexandros Stamatakis. «RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies». En: *Bioinformatics* 30.9 (mayo de 2014). arXiv: 10.1093/bioinformatics/btu033 ISBN: 1367-4811, págs. 1312-1313 (vid. págs. 106, 231, 253).
- [166]Torsten Seeman, Fabian Klötzl y Andrew J. Page. *snp-dist*. 2020 (vid. págs. 106, 253).
- [167]Guangchuang Yu, David K. Smith, Huachen Zhu, Yi Guan y Tommy Tsan-Yuk Lam. «ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data». En: *Methods in Ecology and Evolution* 8.1 (ene. de 2017). Ed. por Greg McInerny. Publisher: John Wiley & Sons, Ltd (10.1111), págs. 28-36 (vid. págs. 106, 231, 253).
- [168]Guangchuang Yu, Tommy Tsan-Yuk Lam, Huachen Zhu y Yi Guan. «Two Methods for Mapping and Visualizing Associated Data on Phylogeny Using *Ggtree*». En: *Molecular Biology and Evolution* 35.12 (dic. de 2018). Ed. por Fabia Ursula Battistuzzi. Publisher: Narnia, págs. 3041-3043 (vid. pág. 106).
- [169] James Hadfield, Nicholas J. Croucher, Richard J. Goater et al. «Phandango: An interactive viewer for bacterial population genomics». En: *Bioinformatics* 34.2 (ene. de 2018). Ed. por Janet Kelso. Publisher: Oxford Academic, págs. 292-293 (vid. págs. 107, 253).
- [170]Lu Cheng, Thomas R. Connor, J. Siren, David M. Aanensen y Jukka Corander.
 «Hierarchical and Spatially Explicit Clustering of DNA Sequences with BAPS Software». En: *Molecular Biology and Evolution* 30.5 (mayo de 2013). ISBN: 1537-1719 (Electronic)\r0737-4038 (Linking) Publisher: Oxford University Press, págs. 1224-1228 (vid. págs. 107, 121, 253).
- [171]Gerry Tonkin-Hill, John A. Lees, Stephen D. Bentley, Simon D. W. Frost y Jukka Corander. «RhierBAPS: An R implementation of the population clustering algorithm hierBAPS». En: Wellcome Open Research 3 (jul. de 2018). Publisher: The Wellcome Trust, pág. 93 (vid. págs. 107, 253).

- [172]Ignacio Ferrés y Gregorio Iraola. «Pagoo: An Encapsulated and Object-Oriented Framework for Evolutionary Analysis of Bacterial Pangenomes». En: SSRN Electronic Journal (2021) (vid. págs. 107, 253).
- [173]Carlos P. Cantalapiedra, Ana Hernández-Plaza, Ivica Letunic, Peer Bork y Jaime Huerta-Cepas. «eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale». En: *Molecular Biology and Evolution* 38.12 (dic. de 2021). Publisher: Oxford Academic, págs. 5825-5829 (vid. págs. 107, 133, 164, 253).
- [174] Jaime Huerta-Cepas, Damian Szklarczyk, Davide Heller et al. «eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses». En: *Nucleic Acids Research* 47.D1 (ene. de 2019), págs. D309-D314 (vid. págs. 107, 165).
- [175]Minoru Kanehisa, Yoko Sato y Masayuki Kawashima. «KEGG mapping tools for uncovering hidden features in biological data». En: *Protein Science* 31.1 (ene. de 2022). Publisher: John Wiley and Sons Inc, págs. 47-53 (vid. págs. 107, 253).
- [176] Minoru Kanehisa, Yoko Sato, Masayuki Kawashima, Miho Furumichi y Mao Tanabe. «KEGG as a reference resource for gene and protein annotation». En: *Nucleic acids research* 44.D1 (2016). Publisher: Nucleic Acids Res, págs. D457-D462 (vid. pág. 107).
- [177]M. Kanehisa y Susumu Goto. «KEGG: Kyoto Encyclopedia of Genes and Genomes». En: *Nucleic Acids Research* 28.1 (ene. de 2000). Publisher: Oxford University Press, págs. 27-30 (vid. pág. 107).
- [178]Sébastien Lê, Julie Josse y François Husson. «FactoMineR: An R Package for Multivariate Analysis». En: Journal of Statistical Software 25.1 (mar. de 2008). Publisher: American Statistical Association, págs. 1-18 (vid. págs. 107, 253).
- [179]Thibaut Jombart. «adegenet: a R package for the multivariate analysis of genetic markers». En: *Bioinformatics* 24.11 (jun. de 2008). Publisher: R Foundation for Statistical Computing, Vienna, Austria, págs. 1403-1405 (vid. págs. 107, 253).
- [180]Martin Christen Frølund Thomsen, Henrik Hasman, Henrik Westh, Hülya Hulya Kaya y Ole Lund. «RUCS: Rapid identification of PCR primers for unique core sequences». En: *Bioinformatics* 33.24 (dic. de 2017). Publisher: Oxford University Press, págs. 3917-3921 (vid. págs. 108, 138, 253).
- [181] Andreas Untergasser, Ioana Cutcutache, Triinu Koressaar et al. «Primer3—new capabilities and interfaces». En: *Nucleic Acids Research* 40.15 (ago. de 2012). Publisher: Oxford University Press, e115-e115 (vid. págs. 108, 138).

- [182]Ruslan Kalendar, Bekbolat Khassenov, Yerlan Ramankulov, Olga Samuilova y Konstantin I. Ivanov. «FastPCR: An in silico tool for fast primer and probe design and advanced sequence analysis». En: *Genomics* 109.3-4 (jul. de 2017). Publisher: Elsevier BV, págs. 312-319 (vid. págs. 108, 138, 253).
- [183]Konstantin Okonechnikov, Olga Golosova, Mikhail Fursov et al. «Unipro UGE-NE: a unified bioinformatics toolkit». En: *Bioinformatics* 28.8 (abr. de 2012). Publisher: Oxford Academic, págs. 1166-1167 (vid. págs. 108, 253).
- [184]Hervé Tettelin, David Riley, Ciro Cattuto y Duccio Medini. «Comparative genomics: the bacterial pan-genome». En: *Current Opinion in Microbiology* 11.5 (oct. de 2008), págs. 472-477 (vid. pág. 129).
- [185]Cláudia Balzan, Rosangela Estel Ziech, Letícia Trevisan Gressler y Agueda Palmira Castagna de Vargas. «Bovine genital campylobacteriosis: main features and perspectives for diagnosis and control». En: *Ciência Rural* 50.3 (2020). Publisher: Universidade Federal de Santa Maria (vid. pág. 143).
- [186]Mostafa Y. Abdel-Glil, Helmut Hotzel, Herbert Tomaso y Jörg Linde. «Phylogenomic Analysis of Campylobacter fetus Reveals a Clonal Structure of Insertion Element ISCfe1 Positive Genomes». En: *Frontiers in Microbiology* 11 (nov. de 2020). Publisher: Frontiers, pág. 585374 (vid. págs. 144, 148, 221, 245).
- [187]Elizabeth A Cummins, Rebecca J Hall, James O McInerney y Alan McNally. «Prokaryote pangenomes are dynamic entities». en. En: *Current Opinion in Microbiology* 66 (abr. de 2022), págs. 73-78 (vid. pág. 146).
- [188]Linda Van der Graaf-Van Bloois, William G. Miller, Emma Yee et al. «Campy-lobacter fetus Subspecies Contain Conserved Type IV Secretion Systems on Multiple Genomic Islands and Plasmids». En: *PLOS ONE* 11.4 (abr. de 2016). Ed. por Dipshikha Chakravortty. Publisher: Public Library of Science, e0152832 (vid. págs. 147, 151).
- [189]Linda Van der Graaf-Van Bloois, Birgitta Duim, William G. Miller et al. «Whole genome sequence analysis indicates recent diversification of mammalassociated Campylobacter fetus and implicates a genetic factor associated with H2S production». en. En: *BMC Genomics* 17.1 (dic. de 2016), pág. 713 (vid. págs. 147, 150).
- [190]a McGoldrick, J. Chanter, S. Gale et al. «Real Time PCR to detect and differentiate Campylobacter fetus subspecies fetus and Campylobacter fetus subspecies venerealis». En: *Journal of Microbiological Methods* 94.3 (2013). ISBN: 0167-7012 Publisher: Elsevier B.V., págs. 199-204 (vid. págs. 148, 149).

- [191]Linda Van der Graaf-Van Bloois, William G. Miller, Emma Yee et al. «Inconsistency of phenotypic and genomic characteristics of campylobacter fetus subspecies requires reevaluation of current diagnostics». En: Journal of Clinical Microbiology 52.12 (dic. de 2014). Publisher: American Society for Microbiology, págs. 4183-4188 (vid. pág. 150).
- [192]Masahiro Asakura, Worada Samosornsuk, Masumi Taguchi et al. «Comparative analysis of cytolethal distending toxin (cdt) genes among Campylobacter jejuni, C. coli and C. fetus strains». En: *Microbial Pathogenesis* 42.5 (mayo de 2007), págs. 174-183 (vid. pág. 152).
- [193]Orhan Sahin, Michael Yaeger, Zuowei Wu y Qijing Zhang. «Campylobacter-Associated Diseases in Animals». En: Annual Review of Animal Biosciences 5 (2017) (vid. pág. 152).
- [194]W. M. Johnson y H. Lior. «A new heat-labile cytolethal distending toxin (CLDT) produced by Campylobacter spp». eng. En: *Microbial Pathogenesis* 4.2 (feb. de 1988), págs. 115-126 (vid. pág. 152).
- [195]Maarten J. Gilbert, William G. Miller, Emma Yee et al. «Comparative Genomics of Campylobacter fetus from Reptiles and Mammals Reveals Divergent Evolution in Host-Associated Lineages». En: *Genome Biology and Evolution* 8.6 (jun. de 2016). Publisher: Oxford University Press, págs. 2006-2019 (vid. págs. 153, 215).
- [196]Erin R. Green y Joan Mecsas. «Bacterial Secretion Systems: An Overview». en. En: *Microbiology Spectrum* 4.1 (ene. de 2016). Ed. por Indira T. Kudva, pág. 4.1.13 (vid. págs. 153-156).
- [197]Stuart A. Thompson, Omer L. Shedd, Kevin C. Ray et al. «Campylobacter fetus Surface Layer Proteins Are Transported by a Type I Secretion System». en. En: Journal of Bacteriology 180.24 (dic. de 1998), págs. 6450-6458 (vid. pág. 153).
- [198]Rebecca S. Wiesner, David R. Hendrixson y Victor J. DiRita. «Natural Transformation of *Campylobacter jejuni* Requires Components of a Type II Secretion System». en. En: *Journal of Bacteriology* 185.18 (sep. de 2003), págs. 5408-5418 (vid. pág. 154).
- [199]Rémi Denise, Sophie S. Abby y Eduardo P. C. Rocha. «Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility». en. En: *PLOS Biology* 17.7 (jul. de 2019). Ed. por Morgan Beeby, e3000390 (vid. pág. 154).

- [200] Joel J. Maki, Mondraya Howard, Sara Connelly et al. «Species Delineation and Comparative Genomics within the *Campylobacter ureolyticus* Complex». en. En: *Journal of Clinical Microbiology* 61.5 (mayo de 2023). Ed. por Nathan A. Ledeboer, e00046-23 (vid. pág. 156).
- [201]Alistair B. Russell, S. Brook Peterson y Joseph D. Mougous. «Type VI secretion system effectors: poisons with a purpose». en. En: *Nature Reviews Microbiology* 12.2 (feb. de 2014), págs. 137-148 (vid. pág. 156).
- [202] Amber D. Gabbert, Jennifer L. Mydosh, Prabhat K. Talukdar et al. «The Missing Pieces: The Role of Secretion Systems in Campylobacter jejuni Virulence». en. En: *Biomolecules* 13.1 (ene. de 2023), pág. 135 (vid. pág. 156).
- [203]Hanna Sprenger, Sabine Kienesberger, Brigitte Pertschy et al. «Fic Proteins of Campylobacter fetus subsp. venerealis Form a Network of Functional Toxin–Antitoxin Systems». En: *Frontiers in Microbiology* 8 (oct. de 2017). Publisher: Frontiers, pág. 1965 (vid. págs. 156, 166, 219, 220).
- [204]Alexander Harms, Frédéric V. Stanger y Christoph Dehio. «Biological Diversity and Molecular Plasticity of FIC Domain Proteins». en. En: Annual Review of Microbiology 70.1 (sep. de 2016), págs. 341-360 (vid. pág. 157).
- [205]Julia Carolin Golz y Kerstin Stingl. «Natural Competence and Horizontal Gene Transfer in Campylobacter». en. En: *Fighting Campylobacter Infections*. Ed. por Steffen Backert. Vol. 431. Series Title: Current Topics in Microbiology and Immunology. Cham: Springer International Publishing, 2021, págs. 265-292 (vid. pág. 158).
- [206]Christopher M. Johnson y Alan D. Grossman. «Integrative and Conjugative Elements (ICEs): What They Do and How They Work». en. En: *Annual Review of Genetics* 49.1 (nov. de 2015), págs. 577-601 (vid. pág. 158).
- [207]Rodolphe Barrangou, Christophe Fremaux, Hélène Deveau et al. «CRISPRProvides Acquired Resistance Against Viruses in Prokaryotes». En: *Science* 315.2007 (mar. de 2007). ISBN: 1095-9203 (Electronic)\r0036-8075 (Linking) Publisher: American Association for the Advancement of Science, págs. 1709-1712 (vid. pág. 159).
- [208] Timothy R. Sampson y David S. Weiss. «Alternative Roles for CRISPR/Cas Systems in Bacterial Pathogenesis». En: *PLoS Pathogens* 9.10 (oct. de 2013). Ed. por Virginia Miller, e1003621 (vid. pág. 159).
- [209] Joe Parker, Andrew Rambaut y Oliver G. Pybus. «Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty». En: *Infection, Genetics and Evolution* 8.3 (mayo de 2008). Publisher: Elsevier, págs. 239-246 (vid. págs. 162, 173).

- [210]Andrew J. Page, Carla A. Cummins, Martin Hunt et al. «Roary: Rapid largescale prokaryote pan genome analysis». En: *Bioinformatics* 31.22 (jul. de 2015). Publisher: Oxford University Press, btv421 (vid. pág. 163).
- [211]Erik S. Wright. «DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment». en. En: *BMC Bioinformatics* 16.1 (dic. de 2015), pág. 322 (vid. pág. 163).
- [212] Trevor C Bruen, Hervé Philippe y David Bryant. «A Simple and Robust Statistical Test for Detecting the Presence of Recombination». en. En: *Genetics* 172.4 (abr. de 2006), págs. 2665-2681 (vid. pág. 163).
- [213]Remco Bouckaert, Timothy G. Vaughan, Joëlle Barido-Sottani et al. «BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis». En: *PLOS Computational Biology* 15.4 (abr. de 2019). Ed. por Mihaela Pertea. Publisher: Public Library of Science (PLoS), e1006650 (vid. págs. 163, 164, 253).
- [214]Patricio Maturana Russel, Brendon J Brewer, Steffen Klaere y Remco R Bouckaert. «Model Selection and Parameter Inference in Phylogenetics Using Nested Sampling». En: *Systematic Biology* 68.2 (mar. de 2019). Ed. por Tanja Stadler, págs. 219-233 (vid. pág. 163).
- [215]Joëlle Barido-Sottani, Veronika Bošková, Louis Du Plessis et al. «Taming the BEAST—A Community Teaching Material Resource for BEAST 2». En: *Systematic Biology* 67.1 (ene. de 2018), págs. 170-174 (vid. pág. 163).
- [216]Subha Kalyaanamoorthy, Bui Quang Minh, Thomas K F Wong, Arndt von Haeseler y Lars S Jermiin. «ModelFinder: fast model selection for accurate phylogenetic estimates». En: *Nature Methods* 14.6 (jun. de 2017), págs. 587-589 (vid. pág. 164).
- [217]Bui Quang Minh, Heiko A Schmidt, Olga Chernomor et al. «IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era». eng. En: *Molecular Biology and Evolution* 37.5 (mayo de 2020). Ed. por Emma Teeling, págs. 1530-1534 (vid. pág. 164).
- [218]Andrew Rambaut, Alexei J Drummond, Dong Xie, Guy Baele y Marc A Suchard. «Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7». eng. En: *Systematic Biology* 67.5 (sep. de 2018). Ed. por Edward Susko, págs. 901-904 (vid. pág. 164).
- [219]Billy Bourke, Voon Loong Chan y Philip Sherman. «Campylobacter upsaliensis : Waiting in the Wings». En: *Clinical Microbiology Reviews* 11.3 (jul. de 1998). Publisher: American Society for Microbiology Journals, págs. 440-449 (vid. pág. 164).

- [220]Torsten Seemann. ABRicate. 2015 (vid. pág. 165).
- [221]Lihong Chen, Jian Yang, Jun Yu et al. «VFDB: A reference database for bacterial virulence factors». En: *Nucleic Acids Research* 33.DATABASE ISS. (Dic. de 2005), págs. D325-D328 (vid. pág. 165).
- [222]Bo Liu, Dandan Zheng, Siyu Zhou, Lihong Chen y Jian Yang. «VFDB 2022: a general classification scheme for bacterial virulence factors». En: *Nucleic Acids Research* 50.D1 (ene. de 2022), págs. D912-D917 (vid. pág. 165).
- [223]Sean R. Eddy. «Accelerated Profile HMM Searches». En: *PLoS Computational Biology* 7.10 (oct. de 2011). Ed. por William R. Pearson. Publisher: Public Library of Science, e1002195 (vid. pág. 165).
- [224]Sophie S. Abby, Bertrand Néron, Hervé Ménager, Marie Touchon y Eduardo P. C. Rocha. «MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems». En: *PLoS ONE* 9.10 (oct. de 2014). Ed. por Néstor V. Torres, e110726 (vid. pág. 165).
- [225]Matías Giménez, Ignacio Ferrés y Gregorio Iraola. Improved detection and classification of plasmids from circularized and fragmented assemblies. en. preprint. Bioinformatics, ago. de 2022 (vid. pág. 166).
- [226]Georges P Schmartz, Anna Hartung, Pascal Hirsch et al. «PLSDB: advancing a comprehensive database of bacterial plasmids». en. En: *Nucleic Acids Research* 50.D1 (ene. de 2022), págs. D273-D278 (vid. pág. 167).
- [227]Ben Langmead y Steven L Salzberg. «Fast gapped-read alignment with Bowtie 2». En: *Nature Methods* 9.4 (abr. de 2012). Publisher: Nature Publishing Group, págs. 357-359 (vid. págs. 167, 253).
- [228]H. Li, B. Handsaker, A. Wysoker et al. «The Sequence Alignment/Map format and SAMtools». En: *Bioinformatics* 25.16 (ago. de 2009). Publisher: Oxford University Press, págs. 2078-2079 (vid. págs. 167, 253).
- [229]Martin Morgan, Hervè Pagès, Valerie Obenchain y Nathaniel Hayden. *Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import.* 2019 (vid. págs. 167, 253).
- [230]Claire Bertelli, Matthew R Laird, Kelly P Williams et al. «IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets». En: *Nucleic Acids Research* 45.W1 (jul. de 2017). Publisher: Oxford University Press, W30-W35 (vid. pág. 168).

- [231]David Couvin, Aude Bernheim, Claire Toffano-Nioche et al. «CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins». En: *Nucleic Acids Research* 46.W1 (jul. de 2018). Publisher: Oxford University Press, W246-W251 (vid. pág. 169).
- [232]Anne Chao. «Estimating the Population Size for Capture-Recapture Data with Unequal Catchability». En: *Biometrics* 43.4 (dic. de 1987). Publisher: JSTOR, pág. 783 (vid. pág. 179).
- [233]Meredith H. Prysak, Christopher J. Mozdzierz, Angela M. Cook et al. «Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and framedependent mRNA cleavage». en. En: *Molecular Microbiology* 71.5 (mar. de 2009), págs. 1071-1087 (vid. págs. 208, 223).
- [234]Christine Pourcel, Marie Touchon, Nicolas Villeriot et al. «CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers». En: *Nucleic Acids Research* 48.D1 (ene. de 2020). Publisher: Oxford University Press, págs. D535-D544 (vid. pág. 213).
- [235]Sabine Kienesberger, Hanna Sprenger, Stella Wolfgruber et al. «Comparative Genome Analysis of Campylobacter fetus Subspecies Revealed Horizontally Acquired Genetic Elements Important for Virulence and Niche Specificity». En: *PLoS ONE* 9.1 (ene. de 2014). Ed. por Thomas Alter. ISBN: 1932-6203 Publisher: Public Library of Science, e85491 (vid. págs. 215, 220).
- [236]Pablo Daniel Farace, José Matías Irazoqui, Claudia Graciela Morsella et al. «Phylogenomic analysis for Campylobacter fetus ocurring in Argentina Phylogenomic analysis for Campylobacter fetus ocurring in Argentina». En: May (2021) (vid. págs. 215, 216).
- [237]Nerea Pena-Fernández, Medelin Ocejo, Linda Van Der Graaf-van Bloois et al. «Comparative pangenomic analysis of Campylobacter fetus isolated from Spanish bulls and other mammalian species». en. En: *Scientific Reports* 14.1 (feb. de 2024), pág. 4347 (vid. págs. 215, 225).
- [238]Petra Müllner, Julie M. Collins-Emerson, Anne C. Midwinter et al. «Molecular Epidemiology of *Campylobacter jejuni* in a Geographically Isolated Country with a Uniquely Structured Poultry Industry». En: *Applied and Environmental Microbiology* 76.7 (abr. de 2010), págs. 2145-2154 (vid. pág. 216).
- [239]L. Calleros, F. Paolicchi, S. Silveyra et al. «Assessing the intra-species genetic variability in the clonal pathogen Campylobacter fetus: CRISPRs are highly polymorphic DNA markers». En: *Journal of Microbiological Methods* 132 (2017). Publisher: Elsevier (vid. págs. 217, 224).

- [240]Samuel K. Sheppard, Frances M. Colles, Noel D. McCarthy et al. «Niche segregation and genetic structure of Campylobacter jejuni populations from wild and agricultural host species». En: *Molecular Ecology* 20.16 (ago. de 2011), págs. 3484-3490 (vid. pág. 217).
- [241]Evangelos Mourkas, Koji Yahara, Sion C Bayliss et al. «Host ecology regulates interspecies recombination in bacteria of the genus Campylobacter». en. En: *eLife* 11 (feb. de 2022), e73552 (vid. pág. 217).
- [242]Ben Pascoe, Guillaume Méric, Koji Yahara et al. «Local genes for local bacteria: Evidence of allopatry in the genomes of transatlantic *Campylobacter* populations». en. En: *Molecular Ecology* 26.17 (sep. de 2017), págs. 4497-4508 (vid. págs. 217, 245).
- [243]Luis Fratti. La calidad de la carne bovina de Uruguay: el resultado de un largo camino. INAC. Goiania, Brasil (vid. pág. 218).
- [244]Estancias Ferguson. *Principales Razas Bovinas productoras de carne de Argentina* (vid. pág. 218).
- [245]Hemant Naikare, Kiran Palyada, Roger Panciera, Denver Marlow y Alain Stintzi. «Major Role for FeoB in *Campylobacter jejuni* Ferrous Iron Acquisition, Gut Colonization, and Intracellular Survival». en. En: *Infection and Immunity* 74.10 (oct. de 2006), págs. 5433-5444 (vid. pág. 219).
- [246]Christina Bronowski, Chloe E. James y Craig Winstanley. «Role of environmental survival in transmission of *Campylobacter jejuni*». en. En: *FEMS Microbiology Letters* 356.1 (jul. de 2014), págs. 8-19 (vid. pág. 219).
- [247]Si-Ping Zhang, Qian Wang, Shuo-Wei Quan et al. «Type II toxin–antitoxin system in bacteria: activation, function, and mode of action». en. En: *Biophysics Reports* 6.2-3 (jun. de 2020), págs. 68-79 (vid. págs. 219, 222).
- [248]Stuart A. Thompson. «Campylobacter Surface-Layers (S-Layers) and Immune Evasion». en. En: Annals of Periodontology 7.1 (dic. de 2002), págs. 43-53 (vid. pág. 220).
- [249]Samuel K. Sheppard, Xavier Didelot, Guillaume Meric et al. «Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter.» En: Proceedings of the National Academy of Sciences of the United States of America 110.29 (jul. de 2013). Publisher: National Academy of Sciences, págs. 11923-7 (vid. pág. 220).
- [250]Roye Rozov, Aya Brown Kav, David Bogumil et al. «Recycler: An algorithm for detecting plasmids from de novo assembly graphs». En: *Bioinformatics* 33.4 (2017) (vid. pág. 222).

- [251]Alessandra Carattoli y Henrik Hasman. «PlasmidFinder and In Silico pMLST: Identification and Typing of Plasmid Replicons in Whole-Genome Sequencing (WGS)». en. En: *Horizontal Gene Transfer*. Ed. por Fernando De La Cruz. Vol. 2075. Series Title: Methods in Molecular Biology. New York, NY: Springer US, 2020, págs. 285-294 (vid. pág. 222).
- [252]M Couturier, F Bex, P L Bergquist y W K Maas. «Identification and classification of bacterial plasmids». en. En: *Microbiological Reviews* 52.3 (sep. de 1988), págs. 375-395 (vid. pág. 222).
- [253]Ramón Díaz-Orejas, Manuel Espinosa y Chew Chieng Yeo. «The Importance of the Expendable: Toxin–Antitoxin Genes in Plasmids and Chromosomes». En: *Frontiers in Microbiology* 8.AUG (ago. de 2017). Publisher: Frontiers Media S.A., pág. 1479 (vid. pág. 223).
- [254]Quan Zhang y Yuzhen Ye. «Not all predicted CRISPR–Cas systems are equal: isolated cas genes and classes of CRISPR like elements». en. En: *BMC Bioinformatics* 18.1 (dic. de 2017), pág. 92 (vid. pág. 224).
- [255]R P Rennie, D Strong, D E Taylor et al. «Campylobacter fetus diarrhea in a Hutterite colony: epidemiological observations and typing of the causative organism». en. En: *Journal of Clinical Microbiology* 32.3 (mar. de 1994), págs. 721-724 (vid. pág. 227).
- [256] Tatsuya Morooka, Akiko Umeda, Masaki Fujita et al. «Epidemiologic Application of Pulsed-Field Gel Electrophoresis to an Outbreak of Campylobacter fetus Meningitis in a Neonatal Intensive Care Unit». en. En: Scandinavian Journal of Infectious Diseases 28.3 (ene. de 1996), págs. 269-270 (vid. pág. 227).
- [257]Martin C. J. Maiden, Melissa J. Jansen van Rensburg, James E. Bray et al. «MLST revisited: the gene-by-gene approach to bacterial genomics». En: *Nature Reviews Microbiology* 11.10 (oct. de 2013). arXiv: 1011.1669v3 ISBN: 1740-1534, págs. 728-736 (vid. pág. 230).
- [258]Daniela Costa, Laura Betancor, Pilar Gadea et al. «Polyclonal Campylobacter fetus Infections Among Unrelated Patients, Montevideo, Uruguay, 2013–2018».
 En: *Clinical Infectious Diseases* 70.September (jul. de 2019). Publisher: Oxford Academic, págs. 2-6 (vid. págs. 233, 238).
- [259]Martin J. Blaser, Diane G. Newell, Stuart A. Thompson y Ellen L. Zechner. «Pathogenesis of Campylobacter fetus». En: *Campylobacter*. Washington, DC, USA: ASM Press, abr. de 2014, págs. 401-428 (vid. pág. 247).
- [260]Alexei J. Drummond y Andrew Rambaut. «BEAST: Bayesian evolutionary analysis by sampling trees». En: *BMC Evolutionary Biology* 7.1 (2007) (vid. pág. 253).

- [261]Ignacio Ferres, Gregorio Iraola, Pablo Fresia y Daniela Costa. *Pewit: Pangenome Estimation Walks Inside Taxonomy*. Montevideo, 2017 (vid. pág. 253).
- [262]I. Milne, G. Stephen, M. Bayer et al. «Using Tablet for visual exploration of second-generation sequencing data». En: *Briefings in Bioinformatics* 14.2 (mar. de 2013), págs. 193-202 (vid. pág. 253).

Índice de figuras

2.1	Árbol filogenético de especies de <i>Campylobacter</i> 9
2.2	Árbol filogenético de especies de Campylobacter con las subespe-
	cies de <i>Campylobacter fetus</i>
3.1	Procesamiento de datos de secuenciación
3.2	Workflow del procesamiento de datos Illumina 73
3.3	Determinación de los sitios de origen y final de replicación con
	oriloc
3.4	Mapa del genoma de la cepa 17144
3.5	Alineamiento de los contigs de U10 al genoma de referencia
	cfv-84112
3.6	Mapa genómico de la cepa U10
3.7	Alineamiento U10_p1 - pCfviMP1
4.1	Estrategia de trabajo
4.2	Distribución geográfica de los genomas
4.3	Hospederos, fuentes y origen geográfico de los aislamientos 112
4.4	Distribución de ANI
4.5	MLST en genomas de aislamientos humanos y bovinos 117
4.6	Árbol filogenético de <i>C. fetus sin raíz</i> de mamíferos 119
4.7	Árbol filogenético C. fetus de mamíferos y metadatos asociados . 120
4.8	Distancias pareadas
4.9	Distancias pareadas por hospedero, subespecie y linaje 124
4.10	Proporción de bases recombinantes por genoma

4.11	Distribución de los bloques recombinantes
4.12	DAPC con linaje, subespecie y hospedero como variables de
	agrupamiento
4.13	Pangenoma
4.14	PCA pangenoma
4.15	MCA de las categorías funcionales del genoma accesorio 134
4.16	Perfiles de presencia/ausencia de genes linaje específicos 136
4.17	Gráficos de barras donde se muestra el número de grupos de
	ortólogos discriminantes de linaje para cada categoría funcional. 137
4.18	Amplicones predichos para la amplificación específica de L1141
4.19	Mapa del entorno de un <i>locus</i> ISCfe1 en el genoma cfv84-112 . 149
5.1	Sistemas de secreción bacterianos
5.2	Estructura básica de un <i>locus</i> CRISPR-Cas
5.3	Distribución geográfica de los genomas del <i>dataset</i>
5.4	Árbol filogenético obtenido a partir del alineamiento genómico
	de genomas del clon bovino ST4
5.5	Distribución de los genomas por sublinaje y su correspondencia
	con el origen geográfico de los aislamientos
5.6	MCA obtenido a partir de las posiciones polimórficas del marco
	clonal
5.7	Diversidad nucleotídica pareada
5.8	Filogenia datada
5.9	Distribuciones de probabilidades posteriores de tMRCA 178
5.10	Curva de rarefacción del pangenoma
5.11	Número de genes accesorios por genoma
5.12	PCA & DAPC del genoma accesorio
5.13	Número de genes accesorios por categoría COG
5.14	Categorias COG de los genes más contribuyentes
5.15	MCA del genoma accesorio

5.16	Frecuencias de los grupos de ortólogos discriminantes por sublinaje185
5.17	Pangenoma por sublinaje
5.18	Localización y estructura de los bloques recombinantes en el
	genoma <i>core</i>
5.19	Número de bases en bloques recombinantes
5.20	Heatmap <i>cdtB</i>
5.21	Contexto génico de <i>cdtB</i>
5.22	Distribución de las proteínas Fic codificadas por genoma 194
5.23	Boxplots mapeo de Fics
5.24	<i>Heatmap</i> Fics
5.25	Sistemas de secreción identificados
5.26	Distribución de los sistemas de secreción por tipo por genoma . 198
5.27	Contexto génico de <i>fhaC</i>
5.28	Distribución de T4SS por genoma
5.29	Sistemas de conjugación por genoma
5.30	Distribución y número de sistemas de conjugación por genoma . 205
5.31	Genes en sistemas T4SS y de conjugación
5.32	Presencia de plásmidos por mapeo
5.33	Boxplots de profundidades de mapeo de plásmidos entre SA y EU210
5.34	<i>Heatmap</i> de resultados de BLASTn de ISCfe1
5.35	Profundidades de mapeo de la secuencia de inserción ISCfe1 213
5.36	Número de CRISPRs por genoma
5.37	CRISPR-Cas
6.1	Árbol filogenético y recombinación de genomas de aislamientos
	clínicos
8.1	Distancias inter-sublinajes
8.2	NZ_CP05442.1

Índice de cuadros

2.1	Especies validadas en LSPD 7
3.1	Comparación de las principales plataformas de secuenciación
	utilizadas
3.2	Información asociada a los genomas de C. fetus obtenidos en
	este trabajo
3.3	Análisis de calidad de <i>reads</i> cortos pre y post-filtrado 77
3.4	Calidad de <i>reads</i> largos
3.5	Calidad de genomas ensamblados
3.6	Medidas resumen de las métricas de anotación de los genomas. 81
3.7	Genomas de <i>C. fetus</i> uruguayos depositados en el NCBI 92
4.1	Calidad de los genomas ensamblados
4.2	Medidas resumen ANI por subespecie
4.3	Número de genomas por ST por hospedero
4.4	Cebadores diseñados
4.5	Resultados de validación de cebadores de L1
5.1	Proteínas Fic de referencia utilizadas en la búsqueda mediante
	BLASTp
5.2	Distancias intra linaje
5.3	Medidas resumen tMRCA
5.4	Composición del pangenoma
5.5	Número de genes accesorios por genoma
5.6	Grupos de genes ortólogos discriminantes

5.7	Plásmidos utilizados como referencia
6.1	Resultados de la tipificación de los aislamientos mediante MLST 232
6.2	Distancias genómicas entre los aislamientos
8.1	Categorías funcionales COG
8.2	Metadatos de aislamientos bovinos ST4 analizados en el Capítulo 5 255
8.3	Metadatos de genomas utilizados para reconstruir la filogenia
	datada en el Capítulo 5
8.4	Datos de secuenciación masiva utilizados para los mapeos 262