



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

FACULTAD DE CIENCIAS BIOLÓGICAS

CENTRO DE INVESTIGACIÓN EN BIOTECNOLOGÍA



CEIB
CENTRO DE INVESTIGACIÓN
EN BIOTECNOLOGÍA UAEM

**GENÓMICA COMPARATIVA DE NUEVE CEPAS
DEL PATÓGENO *Anaplasma marginale*
E IDENTIFICACIÓN DE POSIBLES CANDIDATOS
VACUNALES CONTRA LA ANAPLASMOSIS
BOVINA EN MÉXICO**

T E S I S

QUE PARA OBTENER EL GRADO DE
DOCTOR EN CIENCIAS NATURALES
(LÍNEA TERMINAL DE BIOTECNOLOGÍA)

P R E S E N T A :

M. en B. Fernando Martínez Ocampo

DIRECTOR DE TESIS:

Dr. Edgar Dantán González

CO-DIRECTOR DE TESIS:

Dra. Rosa Estela Quiroz Castañeda

COMITÉ REVISOR

Dr. Edgar Dantán González

Laboratorio de Estudios Ecogenómicos
Centro de Investigación en Biotecnología
Universidad Autónoma del Estado de Morelos

Dra. Rosa Estela Quiroz Castañeda

Unidad de Anaplasmosis
Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad
Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias

Dr. Ramón Suárez Rodríguez

Laboratorio de Fisiología Molecular de Plantas
Centro de Investigación en Biotecnología
Universidad Autónoma del Estado de Morelos

Dr. José Augusto Ramírez Trujillo

Laboratorio de Fisiología Molecular de Plantas
Centro de Investigación en Biotecnología
Universidad Autónoma del Estado de Morelos

Dr. Armando Hernández Mendoza

Laboratorio de Bioinformática Evolutiva
Centro de Investigación en Dinámica Celular
Universidad Autónoma del Estado de Morelos

Dra. María del Rayo Sánchez Carbente

Laboratorio de Biología Molecular de Hongos
Centro de Investigación en Biotecnología
Universidad Autónoma del Estado de Morelos

Dra. Itzel Amaro Estrada

Unidad de Anaplasmosis
Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad
Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias

El presente trabajo se realizó en el Laboratorio de Estudios Ecogenómicos (LEE) del Centro de Investigación en Biotecnología (CEIB), UAEM; la Unidad de Anaplasmosis del Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad (CENID-SAI), INIFAP; el Laboratorio de Bioinformática Evolutiva del Centro de Investigación en Dinámica Celular (CIDC), UAEM; y el Programa de Genómica Evolutiva del Centro de Ciencias Genómicas (CCG), UNAM.



Laboratorio de
Estudios
Ecogenómicos



Instituto Nacional de Investigaciones
Forestales, Agrícolas y Pecuarias



Centro de
Investigación en
Dinámica Celular



Centro de Ciencias Genómicas

El presente trabajo de investigación fue posible gracias al apoyo del Consejo Nacional de Ciencia y Tecnología (CONACYT) a través de la beca nacional número 293552 correspondiente al periodo: Septiembre de 2016 – Agosto de 2020.

A mi hija Julieta Martínez Hernández

A mi madre Ernestina Ocampo González

A mi padre Fernando Martínez Martínez

AGRADECIMIENTOS TÉCNICOS

Al Dr. **Armando Hernández Mendoza** del CIDC, UAEM, por el apoyo con el servidor Lotr, en el cual se realizaron los análisis de los genomas de las nueve cepas mexicanas de *Anaplasma marginale* mediante herramientas bioinformáticas.

Al Dr. **Luis Fernando Lozano Aguirre Beltrán** del CCG, UNAM, por el apoyo con el ensamblaje, la anotación y el análisis de los nueve genomas de cepas mexicanas de *Anaplasma marginale*.

AGRADECIMIENTOS

A mi hija **Julieta Martínez Hernández** por alegrar mi vida con esa hermosa sonrisa. Soy la persona más feliz del universo por tenerte a mi lado ¡Te amo con todo mi ser!

A mis padres, **Ernestina Ocampo González** y **Fernando Martínez Martínez**, por todo el apoyo y amor incondicional que me han brindado durante toda mi vida. Los admiro y respeto. ¡Son los mejores padres y abuelos del universo!

A mi esposa **Karem Hernández Lechuga** por ser mi compañera de vida y apoyarme en todo momento. ¡Eres la persona más increíble e interesante que he conocido en mi vida! ¡Te amo infinito!

A mi hermana **Elisa Martínez Ocampo** y su esposo **Ricardo Marcos Ávila Olac** por ser mi ejemplo de vida en todos los aspectos. ¡Ahora serán la guía de mi hija Julieta!

A la Dra. **Itzel Amaro Estrada** por los comentarios, sugerencias y observaciones que ayudaron a enriquecer la calidad del presente trabajo.

A la Dra. **María del Rayo Sánchez Carbente** por la valiosa contribución con la revisión de los escritos y presentaciones de la presente tesis.

Al Dr. **José Augusto Ramírez Trujillo** por las invaluable aportaciones realizadas a la presente tesis del doctorado.

Al Dr. **Ramón Suárez Rodríguez** por contribuir, de manera significativa, a mi formación académica, profesional y personal durante la maestría y el doctorado.

Al Dr. **Armando Hernández Mendoza** por ser un extraordinario mentor y pieza importante en mi formación como bioinformático. ¡Gracias por creer en mí!

A mi tutora la Dra. **Rosa Estela Quiroz Castañeda** por abrirme las puertas de su laboratorio y brindarme un apoyo total en todo momento. Gracias por transmitir la pasión por la ciencia y la investigación. ¡Estaré eternamente agradecido con usted!

A mi tutor el Dr. **Edgar Dantán González** por el cariño y la confianza que me ha brindado durante casi diez años. Ha sido un mentor en todos los aspectos, pero sobre todo, ha sido un excelente amigo y una extraordinaria persona. ¡Fue un verdadero placer trabajar con usted!

¡GRACIAS INFINITAS!

ÍNDICE GENERAL

ÍNDICE GENERAL	i
ÍNDICE DE FIGURAS	iii
ÍNDICE DE TABLAS	v
RESUMEN.....	1
ABSTRACT.....	2
1. INTRODUCCIÓN	3
1.1 Género <i>Anaplasma</i>	3
1.1.1 Ciclo de vida y transmisión.....	5
1.1.2 Anaplasmosis	6
1.2 <i>Anaplasma marginale</i>	7
1.2.1 Genomas de <i>Anaplasma marginale</i>	8
1.2.2 Superinfección de <i>Anaplasma marginale</i>	10
1.3 Anaplasmosis bovina	11
1.3.1 Vacunología tradicional.....	13
1.3.2 Vacunología reversa	15
1.4 Anaplasmosis bovina en México	16
1.5 Cepas mexicanas de <i>Anaplasma marginale</i>	17
2. HIPÓTESIS	20
3. OBJETIVOS	20
3.1 Objetivo general.....	20
3.2 Objetivos particulares.....	20
4. ESTRATEGIA METODOLÓGICA.....	21
5. MATERIALES Y MÉTODOS	22
5.1 Selección de cepas mexicanas de <i>Anaplasma marginale</i>	22
5.2 Secuenciación masiva del DNA genómico	23
5.3 Ensamblaje de los genomas.....	24
5.4 Análisis de MSP1a	25
5.5 Anotación de los genomas	27
5.6 Análisis filogenético.....	29
5.7 Análisis pangénómico	30
5.8 Genómica comparativa	31
5.9 Identificación de los posibles candidatos vacunales	32
6. RESULTADOS Y DISCUSIÓN.....	34
6.1 Selección de cepas mexicanas de <i>Anaplasma marginale</i>	34
6.2 Secuenciación masiva del DNA genómico	38
6.3 Ensamblaje de los genomas.....	50
6.4 Análisis de MSP1a	54
6.5 Anotación de los genomas	70
6.6 Análisis filogenético.....	77

6.7 Análisis pangénómico	80
6.8 Genómica comparativa	83
6.9 Identificación de los posibles candidatos vacunales	85
7. CONCLUSIONES.....	90
8. PERSPECTIVAS	91
9. BIBLIOGRAFÍA.....	92
10. APÉNDICES	100
Apéndice 1 Genoma reportado de <i>A. marginale</i> MEX-01-001-01	100
Apéndice 2 Genoma reportado de <i>A. marginale</i> MEX-14-010-01	102
Apéndice 3 Genoma reportado de <i>A. marginale</i> MEX-15-099-01	104
Apéndice 4 Genoma reportado de <i>A. marginale</i> MEX-17-017-01	106
Apéndice 5 Genoma reportado de <i>A. marginale</i> MEX-30-184-02	108
Apéndice 6 Genoma reportado de <i>A. marginale</i> MEX-30-193-01	110
Apéndice 7 Genoma reportado de <i>A. marginale</i> MEX-31-096-01	112
Apéndice 8 Genoma reportado de 'Ca. <i>Mycoplasma haemobos</i> ' INIFAP01	114
Apéndice 9 Genoma reportado de <i>Mycoplasma wenyonii</i> INIFAP02	116
Apéndice 10 Artículo publicado.....	118
Apéndice 11 Artículo publicado.....	120
Apéndice 12 Artículo publicado.....	122
Apéndice 13 Artículo publicado.....	124
Apéndice 14 Artículo publicado.....	126
Apéndice 15 Artículo publicado.....	128
Apéndice 16 Artículo publicado.....	130
Apéndice 17 Artículo publicado.....	132
Apéndice 18 Artículo publicado.....	134
Apéndice 19 Capítulo de libro publicado	136
Apéndice 20 Artículo de divulgación científica publicado.....	138

ÍNDICE DE FIGURAS

Figura 1	Distribución de la prevalencia de la anaplasmosis bovina en México.....	34
Figura 2	Distribución geográfica de nueve cepas mexicanas de <i>A. marginale</i>	35
Figura 3	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-01-001-01 (Aguascalientes, Aguascalientes).....	41
Figura 4	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-07-068-01 (Pichucalco, Chiapas).....	42
Figura 5	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-14-010-01 (Atitalaquia, Hidalgo).....	43
Figura 6	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-15-099-01 (Texcoco, Estado de México).....	44
Figura 7	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-17-017-01 (Puente de Ixtla, Morelos).....	45
Figura 8	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-28-037-02 (Soto la Marina, Tamaulipas).....	46
Figura 9	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-30-184-02 (Tlapacoyan, Veracruz).....	47
Figura 10	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-30-193-01 (Veracruz, Veracruz).....	48
Figura 11	Gráficas de calidad Phred por base y distribución del contenido G+C por lectura de <i>A. marginale</i> MEX-31-096-01 (Tizimín, Yucatán).....	49
Figura 12	Diagrama representativo del alineamiento entre MSP1a de <i>A. marginale</i> St. Maries contra el contig 1 de los genomas finales de siete cepas mexicanas de <i>A. marginale</i> ..	55
Figura 13	Diagrama representativo del alineamiento entre MSP1a de <i>A. marginale</i> St. Maries contra los contigs del genoma final de <i>A. marginale</i> MEX-07-068-01 (Pichucalco, Chiapas).....	56
Figura 14	Diagrama representativo del alineamiento entre MSP1a de <i>A. marginale</i> St. Maries contra los contigs del genoma final de <i>A. marginale</i> MEX-28-037-02 (Soto la Marina, Tamaulipas).....	57

Figura 15	Alineamiento entre las lecturas crudas y el gen <i>msp1a</i> de las cepas MEX-15-099-01, MEX-17-017-01, MEX-30-184-03 y MEX-31-096-01 de <i>A. marginale</i>	60
Figura 16	Distribución de las cuatro ecorregiones a nivel mundial.....	65
Figura 17	Relaciones filogenéticas de las cepas mexicanas, australianas, brasileñas, americanas y puertorriqueña de <i>A. marginale</i>	78
Figura 18	Número de secuencias codificantes (CDS) que comparten las siete cepas mexicanas de <i>A. marginale</i>	81
Figura 19	Número de secuencias codificantes (CDS) que comparten 24 genomas de la especie <i>A. marginale</i>	82
Figura 20	Estructura del genoma (sintenia) de las siete cepas mexicanas, una cepa americana, dos cepas australianas y dos cepas brasileñas de <i>A. marginale</i>	84
Figura 21	Mapa circular de la genómica comparativa de las siete cepas mexicanas, dos cepas americanas, dos cepas australianas y dos cepas brasileñas de <i>A. marginale</i>	85

ÍNDICE DE TABLAS

Tabla 1	Características de las seis especies del género <i>Anaplasma</i>	4
Tabla 2	Características de los 24 genomas de <i>A. marginale</i> que están depositados en la base de datos GenBank.....	9
Tabla 3	Genotipos de <i>msp1a</i> en la superinfección de cepas mexicanas de <i>A. marginale</i>	11
Tabla 4	Estructura de las secuencias de repetidas del gen <i>msp1a</i> de 28 diferentes cepas mexicanas de <i>A. marginale</i>	18
Tabla 5	Características de nueve cepas mexicanas de <i>A. marginale</i>	22
Tabla 6	Características de las lecturas crudas, editadas y recortadas de las nueve cepas mexicanas de <i>A. marginale</i>	39
Tabla 7	Características de los genomas originales y depurados de las nueve cepas mexicanas de <i>A. marginale</i>	51
Tabla 8	Características de los genomas finales de las nueve cepas mexicanas de <i>A. marginale</i>	53
Tabla 9	Estructura de las repetidas de MSP1a de las nueve cepas mexicanas de <i>A. marginale</i>	58
Tabla 10	Características general de las cuatro ecorregiones.....	65
Tabla 11	Estructura de los microsatélites de MSP1a de <i>A. marginale</i> y distribución en los cuatro grupos de ecorregiones.....	66
Tabla 12	Primeras (RI) y últimas (RL) repetidas de la estructura MSP1a de <i>A. marginale</i> que están asociadas a las cuatro ecorregiones.....	66
Tabla 13	Características de los microsatélites de MSP1a de las siete cepas mexicanas de <i>A. marginale</i>	68
Tabla 14	Cobertura final de ensamblaje de los genomas finales de las siete cepas mexicanas de <i>A. marginale</i>	70
Tabla 15	Características de la anotación automática de los genomas finales de las siete cepas mexicanas de <i>A. marginale</i>	71

Tabla 16	Categorías de familias (subsistemas) de las proteínas anotadas de las siete cepas mexicanas de <i>A. marginale</i> obtenidas con el servidor RAST.....	73
Tabla 17	Conjuntos de proteínas ortólogas de las siete cepas mexicanas de <i>A. marginale</i> obtenidos con la base de datos COG.....	75
Tabla 18	Secuencias codificantes (CDS) únicas en los genomas finales de las siete cepas mexicanas de <i>A. marginale</i>	81
Tabla 19	Secuencias codificantes (CDS) que fueron anotadas en los genomas finales de las siete cepas mexicanas de <i>A. marginale</i> , al utilizar cuatro diferentes algoritmos.....	86
Tabla 20	Proteínas de membrana externa y extracelulares identificadas en las siete cepas mexicanas de <i>A. marginale</i>	87
Tabla 21	Antígenos protectores identificados en las siete cepas mexicanas de <i>A. marginale</i>	87
Tabla 22	12 antígenos protectores conservados en las siete cepas mexicanas de <i>A. marginale</i>	88
Tabla 23	47 epítomos tipo B que están conservados en las siete cepas mexicanas de <i>A. marginale</i>	89

RESUMEN

Anaplasma marginale es una bacteria intraeritrocítica y es el principal agente etiológico de la anaplasmosis bovina. Los signos clínicos de la enfermedad son: anemia hemolítica, fiebre, pérdida de peso, pérdida en la producción de leche, abortos, ictericia, e incluso la muerte súbita de los bovinos si los medicamentos no se aplican de manera oportuna. Este patógeno se distribuye ampliamente en todo el mundo, principalmente en las regiones tropicales y subtropicales del planeta. Actualmente, no existe una vacuna contra la anaplasmosis bovina a nivel mundial. Una alternativa para prevenir esta enfermedad es mediante el diseño de vacunas que están basadas en la información genómica. Sin embargo, solamente se han reportado genomas de cepas americanas (Estados Unidos), australianas, brasileñas y puertorriqueña de *A. marginale* en la base de datos GenBank. Por lo tanto, el objetivo de este trabajo es identificar los péptidos o proteínas con potencial inmunogénico para el control de la anaplasmosis bovina en México, mediante el análisis de genómica comparativa entre nueve cepas mexicanas de *A. marginale*. En este trabajo se reportó que el DNA genómico de las cepas MEX-15-099-01 (Texcoco, Estado de México) y MEX-31-096-01 (Tizimín, Yucatán) contenía información genómica adicional de otros patógenos, los cuales se reportaron como *Mycoplasma wenyonii* INIFAP02 y 'Candidatus Mycoplasma haemobos' INIFAP01, respectivamente. Además, los genomas de las cepas MEX-07-068-01 (Pichucalco, Chiapas) y MEX-28-037-02 (Soto la Marina, Tamaulipas) mostraron una superinfección de diferentes cepas de *A. marginale* y fueron excluidos de los análisis genómicos posteriores. Por otra parte, los borradores de genomas de siete cepas mexicanas fueron ensamblados con una elevada calidad, entre 32 y 46 contigs con una longitud total de ~1.16 Mb, un contenido G+C de ~49.7% y una cobertura de ensamblaje mayor de 19X. Los borradores de genomas de las siete cepas mexicanas fueron depositados en la base de datos GenBank. La estructura de repetidas de MSP1a de cinco cepas mexicanas no coincidieron con las estructuras reportadas en la literatura. Estos errores pueden disminuir considerablemente mediante la secuenciación masiva de los genomas y no solamente del gen *msh1a*. La anotación automática de los borradores de genomas de las siete cepas mexicanas obtuvo entre 1,138 y 1,178 CDS. Además, los siete borradores de genomas contienen una copia de los tres genes rRNAs y 37 genes tRNAs. El análisis filogenético del genoma core de 24 genomas de *A. marginale*, mostró que: las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) están altamente relacionadas con cepas de la región de Norteamérica; las cepas MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México) y MEX-30-193-01 (Veracruz, Veracruz) están altamente relacionadas con cepas de la región de Sudamérica; y la cepa MEX-17-017-01 (Puente de Ixtla, Morelos) está altamente relacionada con cepas de la región de Asia. El análisis pangénomico entre las siete cepas mexicanas reveló que el genoma core contiene 883 genes, mientras que este valor disminuye a 534 genes al utilizar los 24 genomas de *A. marginale*. La genómica comparativa mostró que los genomas de *A. marginale* están altamente conservados con valores de identidad mayores del 98%. Por último, 12 posibles candidatos vacunales fueron identificados. Estos candidatos vacunales están presentes en las siete cepas mexicanas, no muestran homología con las proteínas de eucariotas, tienen una localización subcelular en la membrana externa y de manera extracelular, son antígenos protectores y contienen una amplia variedad de epítomos de tipo B con longitudes entre 10 y 25 aminoácidos. Además, se identificaron 47 epítomos de tipo B de los 12 posibles candidatos vacunales que pueden ser utilizados para el diseño de MAPs (*Multiple Antigenic Peptides*) como una alternativa para el diseño de vacunas contra la anaplasmosis bovina en México.

ABSTRACT

Anaplasma marginale is an intraerythrocytic bacterium and is the main etiological agent of bovine anaplasmosis. The clinical signs of the disease are: hemolytic anemia, fever, weight loss, loss of milk production, abortions, jaundice, and even sudden death in cattle if medications are not applied in a timely manner. This pathogen is widely distributed throughout the world, mainly in the tropical and subtropical regions of the planet. Currently, there is no vaccine against bovine anaplasmosis worldwide. An alternative to prevent this disease is through the design of vaccines based on genomic information. However, only genomes from American (United States), Australian, Brazilian, and Puerto Rican strains of *A. marginale* have been reported in the GenBank database. Therefore, the objective of this work is to identify peptides or proteins with immunogenic potential for the control of bovine anaplasmosis in Mexico, through comparative genomic analysis of nine Mexican strains of *A. marginale*. In this work, it was reported that the genomic DNA of the MEX-15-099-01 (Texcoco, Estado de México) and MEX-31-096-01 (Tizimín, Yucatán) strains contained additional genomic information of other pathogens, which were reported as *Mycoplasma wenyonii* INIFAP02 and '*Candidatus* *Mycoplasma haemobos*' INIFAP01, respectively. In addition, the genomes of MEX-07-068-01 (Pichucalco, Chiapas) and MEX-28-037-02 (Soto la Marina, Tamaulipas) strains showed a superinfection of different strains of *A. marginale* and were excluded from the next genomics analyses. On the other hand, the draft genomes of seven Mexican strains were assembled with a high quality, between 32 and 46 contigs with a total length of ~1.16 Mb, a G+C content of ~49.7% and an assembly coverage greater than 19X. The draft genomes of the seven Mexican strains were deposited in the GenBank database. The repeats structure of MSP1a of five Mexican strains did not coincide with the structures reported in the literature. These errors can be considerably reduced by massive sequencing of genomes and not only of the *msp1a* genes. Automatic annotation of the draft genomes of the seven Mexican strains obtained between 1,138 and 1,178 CDS. Furthermore, the seven draft genomes contain one copy of the three rRNAs and 37 tRNAs. Phylogenetic analysis of the core genome of 24 *A. marginale* genomes showed that: MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-30-184-02 (Tlapacoyan, Veracruz) and MEX-31-096-01 (Tizimín, Yucatán) strains are highly related to strains from the North America region; MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México) and MEX-30-193-01 (Veracruz, Veracruz) strains are highly related to strains from South America region; and MEX-17-017-01 (Puente de Ixtla, Morelos) strain is highly related to strains from Asian region. The pangenomic analysis among the seven Mexican strains revealed that the core genome contains 883 genes, while this value decreases to 534 genes when using the 24 genomes of *A. marginale*. Comparative genomics showed that *A. marginale* genomes are highly conserved with identity values greater than 98%. Finally, 12 potential vaccine candidates were identified. These vaccine candidates are present in the seven Mexican strains, do not show homology with eukaryotic proteins, have a subcellular location in the outer membrane and extracellularly, are protective antigens, and contain a wide variety of B-type epitopes with lengths between 10 and 25 amino acids. In addition, 47 type B epitopes of the 12 possible vaccine candidates were identified that can be used for the design of MAPs (Multiple Antigenic Peptides) as an alternative for the design of vaccines against bovine anaplasmosis in Mexico.

1. INTRODUCCIÓN

1.1 Género *Anaplasma*

Del griego “*an*” y “*plasma*”, que significan “sin” y “cualquier cosa formada o moldeada”, respectivamente, los organismos que pertenecen al género *Anaplasma* son α -proteobacterias, Gram-negativas, del orden de los Rickettsiales y de la familia Anaplasmataceae. Estas diminutas bacterias no móviles, de formas cocoides a elipsoidales y a menudo pleomórficas, son parásitos intracelulares obligados de células eucariotas que se replican por fisión binaria [1].

El género *Anaplasma*, descubierto en 1910 por Sir Arnold Theiler y reorganizado en el año 2001 por Dumler y colaboradores [2], comprende seis especies de manera oficial: *A. marginale*, *A. centrale*, *A. bovis*, *A. ovis*, *A. phagocytophilum* y *A. platys* [1]. Estas seis especies difieren en los animales hospederos, las células blanco, los vectores biológicos y la distribución geográfica (Tabla 1).

Estudios posteriores basados en el análisis filogenético de la secuencia del gen 16S RNA ribosomal (rRNA) y *groESL*, mostraron que *Aegyptinaella pullorum* está fuertemente relacionado con las especies de *Anaplasma*. Sin embargo, aún no ha sido clasificado de manera oficial como miembro del género *Anaplasma* [3].

Tabla 1. Características de las seis especies del género *Anaplasma* (Editada de Battilani M., et al. 2017) [1].

Especie	Enfermedad(es)	Animales hospederos	Células blanco	Vectores biológicos	Distribución geográfica
<i>A. marginale</i>	Anaplasmosis bovina	Ganado bovino y rumiantes silvestres	Eritrocitos	Especies de los géneros: <i>Ixodes</i> , <i>Dermacentor</i> y <i>Rhipicephalus</i>	Regiones tropicales y subtropicales en todo el mundo
<i>A. centrale</i>	Anaplasmosis bovina asintomática	Ganado bovino	Eritrocitos	<i>Rhipicephalus simus</i>	Regiones tropicales y subtropicales en todo el mundo
<i>A. bovis</i>	Anaplasmosis bovina	Ganado bovino y búfalos	Monocitos	Especies de los géneros: <i>Amblyomma</i> , <i>Rhipicephalus</i> , <i>Hyalomma</i> y <i>Haemaphysalis</i>	África, Asia, Europa, Estados Unidos y Sudamérica
<i>A. ovis</i>	Anaplasmosis ovina	Ovejas, cabras y rumiantes silvestres	Eritrocitos	<i>Melophagus ovinus</i> y especies de los géneros: <i>Dermacentor</i> y <i>Rhipicephalus</i>	África, Asia, Europa y Estados Unidos
<i>A. phagocytophilum</i>	Anaplasmosis granulocítica humana; Anaplasmosis equina; Fiebre transmitida por garrapatas de los rumiantes; Anaplasmosis de perros y gatos	Humanos, caballos, rumiantes, roedores, carnívoros e insectívoros	Granulocitos y células endoteliales	Especies del género <i>Ixodes</i>	Todo el mundo
<i>A. platys</i>	Trombocitopenia cíclica en perros	Perros	Plaquetas	<i>Rhipicephalus sanguineus</i>	Todo el mundo

Recientemente, se han identificado otras posibles especies del género *Anaplasma*. En el año 2012, se reportó una posible especie que está fuertemente relacionada con *A. phagocytophilum* y fue detectada en el ciervo sika (*Cervus nippon yezoensis*) en Japón [4]. Posteriormente, en el año 2013 se reportó a *Anaplasma odocoilei* sp. nov., la cual fue detectada en el ciervo de cola blanca (*Odocoileus virginianus*) en Estados Unidos [5]. Por último, en el año 2015 se reportó una nueva especie de *Anaplasma*, nombrada de manera provisional como

Anaplasma capra, la cual fue detectada en cabras asintomáticas en China y es potencialmente patógena de humanos [6].

1.1.1 Ciclo de vida y transmisión

El ciclo de vida de las especies del género *Anaplasma*, involucra la replicación de las bacterias en los animales vertebrados y las garrapatas de la familia Ixodidae [7]. Los animales vertebrados funcionan como hospederos y se consideran reservorios de estas bacterias porque pueden desarrollar infecciones de manera persistente. Además, los animales hospederos pueden actuar como una fuente de infección para las garrapatas, las cuales funcionan como vectores biológicos [8–10]. El ciclo de vida inicia cuando las garrapatas se alimentan de la sangre de los animales hospederos infectados con especies del género *Anaplasma*. Después, estas bacterias ingresan al epitelio del intestino medio de las garrapatas, en donde ocurre el primer ciclo de replicación. Posteriormente, las bacterias migran al epitelio de las glándulas salivales de las garrapatas, en donde ocurre el segundo ciclo de replicación. Por último, las bacterias ingresan a la saliva de las garrapatas, cuando estas se alimentan de la sangre de otros animales vertebrados [11]. Los vectores biológicos transmiten a estas bacterias de manera transestadial (es decir, de larvas a ninfas y de ninfas a adultos), pero no de manera transovárica (es decir, a las generaciones sucesivas) [12,13].

La transmisión mecánica de las especies del género *Anaplasma* ocurre mediante la picadura de insectos (por ejemplo, las moscas hematófagas) o de

fómites (por ejemplo, las agujas), cuando están contaminados con la sangre de los animales hospederos de estas bacterias. La transmisión mecánica ocurre cuando los vectores biológicos están presentes o ausentes [14].

1.1.2 Anaplasmosis

Las especies del género *Anaplasma* son los agentes etiológicos de la enfermedad llamada Anaplasmosis. Esta enfermedad afecta a diversos organismos, incluidos los rumiantes, pero la enfermedad clínica es más notable en el ganado. Sin embargo, los rumiantes silvestres pueden estar infectados de manera persistente. La Anaplasmosis es una enfermedad infecciosa pero no contagiosa [15].

La anaplasmosis bovina es causada por la infección de *A. bovis*, *A. phagocytophilum*, *A. centrale* o *A. marginale* al ganado bovino. El patógeno *A. bovis* infecta a los monocitos de los bovinos que induce una enfermedad clínica leve, como la presencia de fiebre, anemia, somnolencia, convulsiones, pérdida de peso y agrandamiento de los ganglios linfáticos [16]. *A. phagocytophilum* es un parásito intragranulocítico que infecta a los neutrófilos y eosinófilos de los animales hospederos, los cuales pueden ser bovinos y humanos [17]. Los animales infectados con *A. phagocytophilum* pueden desarrollar leucopenia, neutropenia y la reducción de la función de los neutrófilos, lo cual puede ocasionar una inmunosupresión y la aparición de infecciones oportunistas [18]. Por otra parte, *A. centrale* es un parásito intraeritrocítico que causa una enfermedad clínica leve o

asintomática en el ganado bovino, es decir, es una especie naturalmente atenuada [13]. *A. centrale* se considera una especie menos virulenta pero fuertemente relacionada con *A. marginale*. Por lo tanto, *A. centrale* puede proporcionar un tipo de inmunidad protectora contra la infección de algunas cepas de *A. marginale*. Por este motivo, *A. centrale* se ha utilizado como una opción de vacuna viva para prevenir la infección de *A. marginale* en países como Israel, Sudáfrica, Sudamérica y Australia [14,19]. Por último, *A. marginale* también es un parásito intraeritrocítico, el cual es considerado el principal agente etiológico de la anaplasmosis bovina a nivel mundial porque puede causar una enfermedad potencialmente mortal [1,14].

1.2 *Anaplasma marginale*

El patógeno *A. marginale* fue descrito por primera vez en 1910 por Sir Arnold Theiler, quien observó pequeñas inclusiones cerca del borde (área marginal) de los eritrocitos del ganado bovino enfermo [20]. *A. marginale* se distribuye ampliamente a nivel mundial, especialmente en las regiones tropicales y subtropicales del planeta. Sin embargo, la diseminación de este patógeno está en constante aumento. Una propuesta es debido al cambio climático, el cual puede influir en el movimiento del ganado bovino por los drásticos cambios de temperatura, y por lo tanto, ocasiona el movimiento de las poblaciones de garrapatas como vectores biológicos [21]. Los vectores biológicos de *A. marginale* incluyen al menos a 20 especies de garrapatas, principalmente de los géneros *Dermacentor* y *Rhipicephalus* [14]. La transmisión transplacentaria de *A. marginale* en el ganado bovino, resulta en el nacimiento de terneros sanos pero infectados de manera

persistente [22]. *A. marginale* también puede infectar una amplia diversidad de animales rumiantes silvestres, que incluyen al búfalo de agua (*Bubalus bubalis*), el bisonte americano (*Bison bison*), el ciervo de cola blanca (*Odocoileus virginianus*), el ciervo mulo (*Odocoileus hemionus hemionus*), el ciervo de cola negra (*Odocoileus hemionus columbianus*), el antílope americano (*Antilocapra americana*), el borrego cimarrón (*Ovis canadensis*) y la jirafa (*Giraffa camelopardalis*), por mencionar algunos [15].

1.2.1 Genomas de *Anaplasma marginale*

Hasta el día 10 de marzo del año 2022, se han reportado 24 genomas de *A. marginale* que están depositados en la base de datos GenBank de NCBI (siglas del inglés *National Center for Biotechnology Information*) (<https://bit.ly/3vJiL11>), de los cuales dos genomas están completos, cuatro genomas están cerrados en un cromosoma y 18 son borradores de genomas (Tabla 2). En Estados Unidos se reportaron tres genomas que pertenecen al mismo microorganismo: el genoma completo de *A. marginale* Florida y los dos borradores de genomas de *A. marginale* Florida y Florida Relapse (DNA genómico obtenido 129 días después de la infección). También, en Estados Unidos se reportaron dos genomas que pertenecen al mismo microorganismo: el genoma completo y el borrador de genoma de *A. marginale* St. Maries.

Tabla 2. Características de los 24 genomas de *A. marginale* que están depositados en la base de datos GenBank (datos actualizados hasta el 10 de marzo de 2022).

Cepa	Longitud total (pb)	Cromosomas	Contigs	Contenido G+C (%)*	Genes**	CDS**	País
<i>A. marginale</i> Dawn	1,196,760	1	-----	49.73	1,145	1,107	Australia
<i>A. marginale</i> Gypsy Plains	1,198,622	1	-----	49.72	1,189	1,149	Australia
<i>A. marginale</i> Jaboticabal	1,195,321	1	-----	49.77	1,238	1,198	Brasil
<i>A. marginale</i> Palmeira	1,195,200	1	-----	49.75	1,219	1,179	Brasil
<i>A. marginale</i> Florida	1,202,435	1	-----	49.77	1,227	1,187	Estados Unidos
<i>A. marginale</i> Florida	1,136,981	-----	204	49.84	1,186	1,147	Estados Unidos
<i>A. marginale</i> Florida Relapse	1,154,411	-----	61	49.81	1,189	1,149	Estados Unidos
<i>A. marginale</i> Mississippi	1,141,520	-----	82	49.79	1,208	1,168	Estados Unidos
<i>A. marginale</i> Okeechobee	1,390,987	-----	403	47.49	1,267	1,225	Estados Unidos
<i>A. marginale</i> Oklahoma	1,156,921	-----	57	49.82	1,167	1,127	Estados Unidos
<i>A. marginale</i> Oklahoma-2	1,160,766	-----	49	49.79	1,188	1,148	Estados Unidos
<i>A. marginale</i> South Idaho	1,409,432	-----	358	46.73	1,316	1,262	Estados Unidos
<i>A. marginale</i> St. Maries	1,197,687	1	-----	49.76	1,250	1,210	Estados Unidos
<i>A. marginale</i> St. Maries	1,155,236	-----	60	49.79	1,194	1,154	Estados Unidos
<i>A. marginale</i> Virginia	1,153,875	-----	70	49.79	1,241	1,201	Estados Unidos
<i>A. marginale</i> Washington Okanogan	1,383,255	-----	332	46.94	1,331	1,274	Estados Unidos
<i>A. marginale</i> MEX-01-001-01	1,179,425	-----	34	49.79	1,218	1,178	México
<i>A. marginale</i> MEX-14-010-01	1,172,327	-----	46	49.79	1,190	1,150	México
<i>A. marginale</i> MEX-15-099-01	1,169,440	-----	32	49.79	1,185	1,145	México
<i>A. marginale</i> MEX-17-017-01	1,172,716	-----	41	49.79	1,203	1,163	México
<i>A. marginale</i> MEX-30-184-02	1,176,681	-----	40	49.79	1,205	1,165	México
<i>A. marginale</i> MEX-30-193-01	1,167,111	-----	41	49.80	1,178	1,138	México
<i>A. marginale</i> MEX-31-096-01	1,176,579	-----	43	49.79	1,204	1,164	México
<i>A. marginale</i> Puerto Rico	1,158,530	-----	59	49.80	1,220	1,180	Puerto Rico

* Datos obtenidos con el programa QUAST.

** Datos obtenidos con el servidor RAST.

En la base de datos GenBank están reportados los borradores de genomas de siete cepas mexicanas de *A. marginale*: MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México), MEX-17-017-01 (Puente de Ixtla, Morelos), MEX-30-184-02 (Tlapacoyan, Veracruz), MEX-30-193-01 (Veracruz, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán); los cuales son el resultado del trabajo de la presente tesis de doctorado. La información genómica de *A. marginale* que está disponible en las bases de datos, ha sido reportada de cinco países (Australia, Brasil, Estados Unidos, México y Puerto Rico) y resulta poco representativa en comparación con la amplia distribución geográfica que tiene este patógeno a nivel mundial.

1.2.2 Superinfección de *Anaplasma marginale*

Previamente, se creía que la infección de los eritrocitos del ganado bovino por un genotipo de *A. marginale*, excluía la infección de otros genotipos diferentes de *A. marginale* en el mismo animal. A este fenómeno se le nombró como exclusión de la infección [23]. En el año 2014, Palmer y colaboradores estudiaron un rebaño de bovinos en Estados Unidos, donde se detectaron 70 animales que estaban infectados con un genotipo de *A. marginale*. Sin embargo, se detectaron cinco animales que estaban infectados con dos genotipos diferentes de *A. marginale* en el mismo bovino. Es decir, cada animal estaba superinfectado con dos cepas diferentes de *A. marginale* [24]. Posteriormente, en el año 2015, Castañeda-Ortiz y colaboradores estudiaron dos rebaños de bovinos en la república mexicana: en la Joya, Jalisco y el Verdineño, Nayarit. Con base en la amplificación, la clonación y la

secuenciación del gen *msp1a*, en el rebaño de la Joya se detectaron siete animales infectados con una cepa de *A. marginale*. Y además, se detectaron siete animales superinfectados hasta con cinco cepas diferentes de *A. marginale*. Por otra parte, en el rebaño del Verdineño se detectaron seis animales infectados con una cepa de *A. marginale*, mientras que se detectaron 37 animales superinfectados hasta con nueve cepas diferentes de *A. marginale* (Tabla 3) [25].

Tabla 3. Genotipos de *msp1a* en la superinfección de cepas mexicanas de *A. marginale* (Editada de Castañeda-Ortiz E. J., *et al.* 2015) [25].

Número de genotipos diferentes de <i>A. marginale</i> identificados en el mismo bovino	Número de animales infectados	
	Rebaño de la Joya, Jalisco	Rebaño del Verdineño, Nayarit
1	7	6
2	2	11
3	2	10
4	2	5
5	1	3
6	0	4
7	0	2
8	0	0
9	0	2

1.3 Anaplasmosis bovina

La anaplasmosis bovina es una enfermedad que se ha reportado en las regiones tropicales y subtropicales del planeta, que incluyen Sudamérica, América Central, Estados Unidos, el sur de Europa, África, Asia y Australia. Además, se ha reportado que es una enfermedad endémica del ganado bovino en México, América Central, el Caribe y Sudamérica [14]. En el año 1969, Canadá legisló el control para

la importación del ganado bovino, y en la actualidad, existe una política estricta para la erradicación de la anaplasmosis bovina.

Los signos clínicos de la anaplasmosis bovina son: anemia hemolítica progresiva, fiebre, pérdida excesiva de peso, pérdida en la producción de leche, abortos en el último trimestre del embarazo, ictericia en la etapa final de la enfermedad, e incluso la muerte súbita de los bovinos si los medicamentos no se aplican de manera oportuna [15]. Sin embargo, la enfermedad clínica sólo se puede confirmar mediante la identificación del patógeno *A. marginale*. Esto se logra mediante la observación en el microscopio de frotis de sangre que están teñidos con Giemsa u otra tinción de Romanowsky [26]; o mediante la amplificación de genes o fragmentos de genes de *A. marginale* por PCR (siglas del inglés *Polymerase Chain Reaction*) [27]. La anaplasmosis bovina es responsable de importantes pérdidas económicas en la industria de la ganadería a nivel mundial [14,20].

Actualmente, el tratamiento de los casos clínicos agudos contra la anaplasmosis bovina, consiste en la administración de tetraciclinas, las cuales son medicamentos bacteriostáticos que funcionan al inhibir la síntesis de proteínas de *A. marginale* [28]. La clorotetraciclina y oxitetraciclina son los únicos compuestos derivados de la tetraciclina que están aprobados en Estados Unidos como tratamiento contra la anaplasmosis bovina. También, el Imidocarb® es un medicamento que se utiliza para el tratamiento de esta enfermedad. Sin embargo,

el tratamiento con estos químicos resulta altamente costoso y restringe la venta de la carne y leche de los bovinos en un periodo de aproximadamente 30 días [29].

Por otra parte, una medida de control contra la anaplasmosis bovina es el uso de acaricidas que permiten disminuir considerablemente las poblaciones de garrapatas. Sin embargo, el uso de estas sustancias permite la selección de las poblaciones de garrapatas resistentes. Además, la presencia de residuos de acaricidas en la carne y leche de los bovinos, es un tema de preocupación para la salud pública [30]. Una opción para prevenir la anaplasmosis bovina es mediante la aplicación de vacunas de cepas vivas atenuadas o inactivadas de *A. marginale*. No obstante, estas vacunas han tenido resultados variables, principalmente por el deficiente y costoso proceso de purificación de las cepas de *A. marginale* [29]. Por lo tanto, el diseño y desarrollo de vacunas que permitan prevenir la anaplasmosis bovina tiene un especial interés económico y científico a nivel mundial.

1.3.1 Vacunología tradicional

A partir del siglo pasado, se han propuesto vacunas de cepas vivas atenuadas o inactivadas de *A. marginale* o *A. centrale* contra la anaplasmosis bovina [31–35]. Sin embargo, estas vacunas permiten diseminar a otros organismos patógenos de transmisión sanguínea, pueden causar una reversión de la patogenicidad, e incluso pueden generar una contaminación con los antígenos de los eritrocitos del bovino hospedero [36–38]. También, se han desarrollado vacunas recombinantes a partir de las proteínas de membrana externa, como las proteínas Vir del Sistema de

Secreción Tipo IV [39]. No obstante, estas proteínas recombinantes han generado una inmunoprotección deficiente debido a la amplia diversidad y variabilidad genética de las cepas de *A. marginale* en diferentes regiones geográficas, en el mismo rebaño e incluso en el mismo animal [24,25,40,41]. Las vacunas recombinantes tienen la limitante de invertir tiempo, recursos y esfuerzo para identificar y analizar sólo una proteína como candidato vacunal. En la actualidad, no existen vacunas recombinantes comerciales contra la anaplasmosis bovina a nivel mundial.

1.3.1.1 Principales proteínas de superficie

Actualmente, las cinco principales proteínas de superficie (MSPs del inglés *Major Surface Proteins*) mejor caracterizadas en *A. marginale* son las proteínas de membrana externa, las cuales están codificadas en los genes de copia única *msp1a*, *msp4* y *msp5* que están conservados a nivel de especie y son ampliamente utilizados como marcadores moleculares para determinar la diversidad genética entre las cepas de *A. marginale* [15]. La diversidad del gen *msp1a* se utiliza para identificar las diferentes cepas de *A. marginale* a nivel mundial. Además, la secuencia del gen *msp4* es utilizada para estudios filogenéticos [40]. Por último, la presencia del gen *msp5* se utiliza para el diagnóstico molecular de la infección de *A. marginale* en los eritrocitos del ganado bovino. Además, la proteína Msp5 es altamente inmunogénica y se utiliza para el diagnóstico serológico de la anaplasmosis bovina [27].

La estructura de la proteína Msp1a está formada por un dominio conservado intracelular en el extremo C-terminal y por un dominio variable extracelular en el extremo N-terminal. El dominio variable contiene una o más secuencias similares que son llamadas repetidas, las cuales tienen entre 23 y 31 aminoácidos de longitud cada una [42,43]. En el año 2007, de la Fuente y colaboradores identificaron 79 diferentes secuencias de repetidas en 131 cepas de *A. marginale* de América del Norte, Sudamérica, Europa, África, Asia y Australia. También, reportaron que el 78% de las secuencias de repetidas están presentes en las cepas de *A. marginale* de cualquier región geográfica. Además, existen secuencias de repetidas únicas para algunas cepas presentes en regiones geográficas como Italia, España, China, Argentina y Sudamérica. Por último, se reportaron que cuatro cepas de *A. marginale* de Australia tienen un genotipo de *msp1a* con la secuencia de la repetida única llamada “8” [34,44].

1.3.2 Vacunología reversa

En la actualidad, el progreso de las tecnologías de secuenciación masiva del DNA está generando una inmensa cantidad de información genómica que está disponible en las bases de datos a nivel mundial. El diseño de vacunas que están basadas en la información genómica de los organismos patógenos, cultivables y no cultivables *in vitro*, es una alternativa segura y eficaz para identificar los péptidos o proteínas que tienen el potencial antigénico e inmunogénico predicho para ser considerados como posibles candidatos vacunales [34,45]. Hasta la fecha, sólo se han reportado los genomas de siete cepas mexicanas de *A. marginale* en las bases

de datos a nivel mundial, como resultado del trabajo de la presente tesis de doctorado. Por lo tanto, previamente no había sido posible diseñar vacunas contra la anaplasmosis bovina para la región geográfica de México mediante la vacunología reversa.

Por otra parte, los péptidos antigénicos múltiples, también llamados MAPs (del inglés *Multiple Antigenic Peptides*), están formados por uno o más péptidos sintéticos, los cuales están unidos de manera covalente a una macromolécula de matriz sólida con ramificaciones de residuos de lisina [46]. El diseño de MAPs que contengan los péptidos sintéticos obtenidos a partir de los genomas de las cepas mexicanas de *A. marginale*, podrían ser utilizados como una alternativa para el diseño de vacunas contra la anaplasmosis bovina en México.

1.4 Anaplasmosis bovina en México

La industria ganadera es una de las actividades agroindustriales más redituables a nivel mundial y México ocupa el octavo puesto en inventario del ganado bovino [47]. En el año 2020 se registraron 35,653,619 cabezas de ganado, de las cuales 2,606,311 y 33,047,308 son utilizadas para la producción de leche y carne, respectivamente (datos obtenidos del Servicio de Información Agroalimentaria y Pesquera; <https://bit.ly/35luVNq>).

En México, la anaplasmosis bovina provoca importantes pérdidas económicas a la industria ganadera. Estas pérdidas económicas son difíciles de calcular porque

existen varios factores a considerar, como la pérdida de peso, la pérdida en la producción de leche, el aborto y la mortalidad, por mencionar algunos ejemplos. No existen datos exactos, pero se estima que esta enfermedad es responsable de aproximadamente el 25% de la mortalidad del ganado bovino en los programas de mejoramiento genético [48].

Por otra parte, no existen vacunas comerciales disponibles contra la anaplasmosis bovina a nivel mundial. Sin embargo, en México se ha utilizado como vacuna a la cepa viva de *A. marginale* Yucatán, la cual induce protección sólida y sin efectos secundarios [32]. Además, no se ha utilizado a *A. centrale* como vacuna viva porque se considera un organismo exótico y no permitido en México.

1.5 Cepas mexicanas de *Anaplasma marginale*

En la actualidad, son escasos los reportes sobre las cepas mexicanas de *A. marginale*. En el año 2012, Jiménez-Ocampo y colaboradores reportaron el análisis estructural de la proteína Msp1a e identificaron 28 diferentes cepas mexicanas de *A. marginale* de diferentes regiones geográficas de la república mexicana (Tabla 4) [49]. Posteriormente, en el año 2015, Castañeda-Ortiz y colaboradores reportaron diferentes cepas mexicanas de *A. marginale* en dos rebaños de bovinos: en la Joya, Jalisco y el Verdineño, Nayarit [25]. La tasa de infección de las cepas mexicanas de *A. marginale* fue del 36% (30 animales infectados de 84 posibles) y del 100% (82 animales) en los rebaños de bovinos de la Joya y el Verdineño, respectivamente. Además, se reportaron dos (LJ1 y LJ2) y 12 (EV1 hasta EV12) nuevas secuencias

de repetidas del gen *msp1a* a nivel mundial, las cuales provienen de los rebaños de bovinos de la Joya y el Verdineño, respectivamente. También, se identificaron 20 y 60 genotipos diferentes de *msp1a* de las cepas mexicanas de *A. marginale* en los bovinos infectados de la Joya y el Verdineño, respectivamente [25]. Por lo tanto, la amplia diversidad genética de las cepas mexicanas de *A. marginale*, ha dificultado la identificación de péptidos o proteínas como posibles candidatos vacunales contra la anaplasmosis bovina en México.

Tabla 4. Estructura de las secuencias de repetidas del gen *msp1a* de 28 diferentes cepas mexicanas de *A. marginale* (Editada de Jiménez-Ocampo R., *et al.* 2012 [49]).

Cepa mexicana	Estructura de repetidas de <i>msp1a</i>							Número de repetidas
Mex-31-089-01 - Ticul, Yucatán	F	M	M					3
Mex-30-184-03 - Tlapacoyan, Veracruz	T	C	B	B	C	C		6
Mex-30-130-01- Playa Vicente, Veracruz	T	C	B	B	C	B	C	7
Mex-31-096-01 - Tizimin, Yucatán	T	C	B	B	C	B	π	7
Mex-15-099-01 - Texcoco, Edo. de México	α	β	β	Γ				4
Mex-17-030-01 - Yauatepec, Morelos	α	β	β	Γ				4
Mex-30-193-01 - Veracruz, Veracruz	α	β	β	Γ				4
Mex-18-017-01 - Tepic, Nayarit	α	β	β	Γ				4
Mex-07-068-02 - Sta. Martha, Chiapas	α	β	β	Γ	β			5
Mex-07-068-01 - Pichucalco, Chiapas	α	β	β	Γ	β	Γ		6
Tamaulipas 17 y 18 (G10)	α	β	β	β	Γ			5
Mex-07-065-01 - Palenque, Chiapas	β	β	β	β	Γ			5
Mex-30-184-02 - Tlapacoyan, Veracruz	73	β	β	β	Γ			5
Tamaulipas 15 (G9)	τ	57	13	18				4
Mex-14-010-01 - Atitalaquia, Hidalgo	τ	57	13	18				4
Mex-28-037-01 - Soto la Marina, Tamaulipas	τ	57	13	18				4
Mex-01-001-01 - Aguascalientes, Aguascalientes	4	9	10	11	9			5
Tamaulipas 13 (G2)	4	9	10	10	9			5
Mex-17-017-01 - Puente de Ixtla, Morelos	12	13	14					3
Mex-28-037-02 - Soto la Marina, Tamaulipas	28	29	74	29	M	F		6
Tamaulipas 1 (G1)	56	57	58	59				4
Tamaulipas 4 (G3)	60	61	61	62	61			5
Tamaulipas 7, 9, 10 y 12 (G4)	4	63	63	27	12			5
Tamaulipas 11 (G5)	67	68	63	27	12			5
Tamaulipas 14 (G6)	69	61	70	71	61			5
Tamaulipas 5 (G7)	64	65	D	65	66			5
Tamaulipas 6 y 8 (G8)	D	65	D	65	66			5
Mex-30-184-01 - Tlapacoyan, Veracruz	72	C	F					3

Por lo antes expuesto, el presente trabajo tiene el propósito de identificar a las secuencias codificantes que tengan el potencial para ser utilizadas como posibles candidatos vacunales contra la anaplasmosis bovina en México. Esto se logrará mediante el análisis de genómica comparativa entre diferentes cepas mexicanas de *A. marginale*. Además, se identificarán las similitudes y diferencias genómicas que existen entre las cepas mexicanas, americanas (Estados Unidos), australianas, brasileñas y puertorriqueña de *A. marginale*.

2. HIPÓTESIS

El análisis de genómica comparativa entre nueve cepas mexicanas de *Anaplasma marginale* permite identificar los posibles inmunógenos contra la anaplasmosis bovina en México.

3. OBJETIVOS

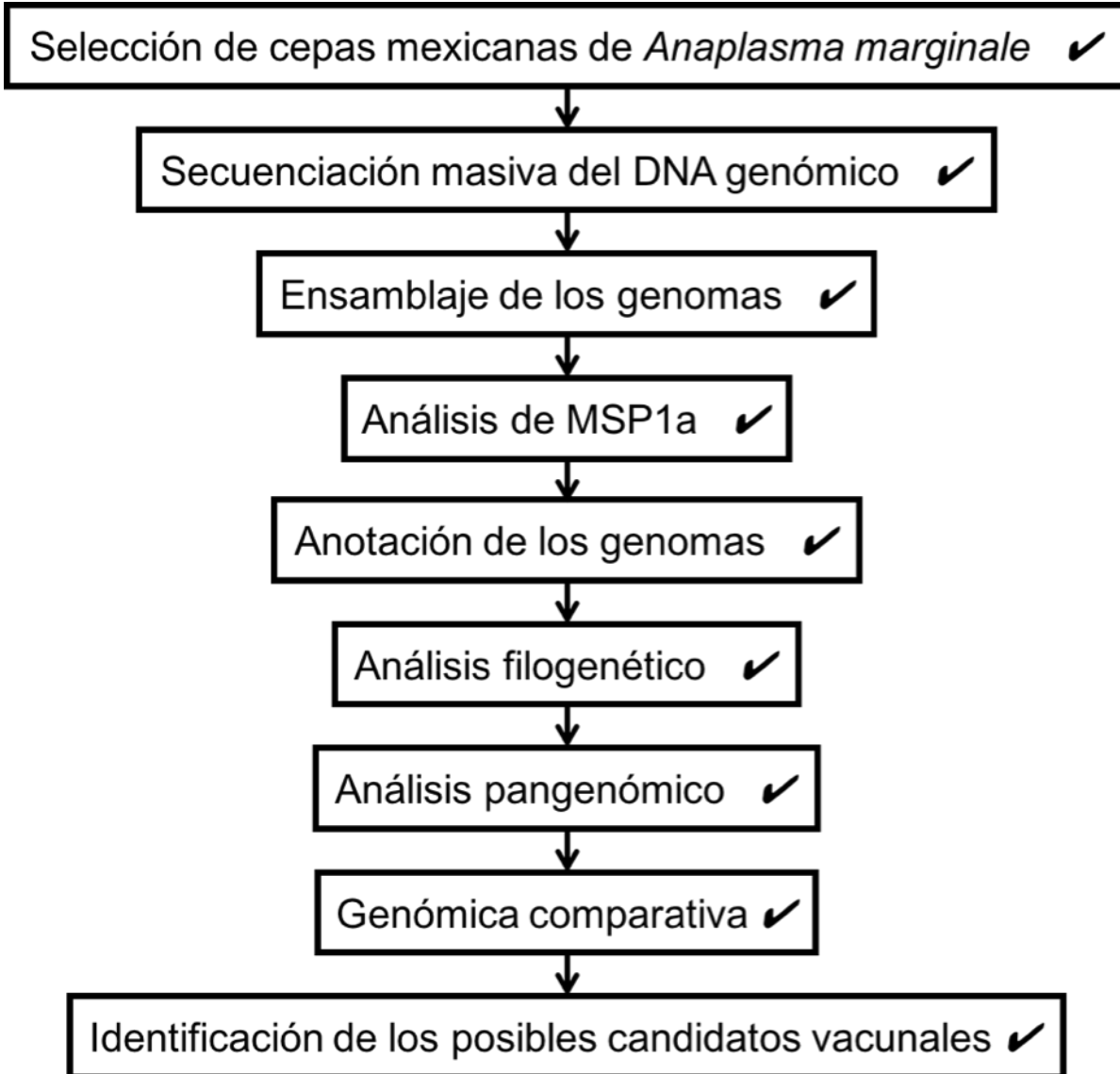
3.1 Objetivo general

Identificar péptidos o proteínas con potencial inmunogénico mediante el análisis de genómica comparativa para el control de la anaplasmosis bovina en México.

3.2 Objetivos particulares

- I. Secuenciar, ensamblar y anotar los genomas de nueve cepas mexicanas de *Anaplasma marginale*.
- II. Identificar los péptidos o proteínas con potencial inmunogénico mediante análisis bioinformáticos.

4. ESTRATEGIA METODOLÓGICA



5. MATERIALES Y MÉTODOS

5.1 Selección de cepas mexicanas de *Anaplasma marginale*

La Unidad de Anaplasmosis del CENID-SAI que pertenece al INIFAP, posee una colección exclusiva de diferentes cepas mexicanas de *A. marginale* [49]. Nueve cepas mexicanas (Tabla 5) fueron seleccionadas bajo los siguientes criterios: (i) cepas que provienen de diferentes regiones geográficas de la república mexicana; (ii) cepas que tienen diferente estructura de repetidas de MSP1a; y (iii) dos cepas que tienen reportado un diferente grado de virulencia: la cepa de alto grado de virulencia de Texcoco, Estado de México [50] y la cepa de bajo grado de virulencia de Tizimín, Yucatán [32]. La nomenclatura de las nueve cepas mexicanas utiliza los códigos del INEGI que describen los números del estado, municipio y aislado, respectivamente, los cuales están separados por guiones, como fue reportado previamente [49]. Las nueve muestras de sangre que fueron obtenidas de bovinos infectados con las cepas mexicanas de *A. marginale*, se almacenaron en nitrógeno líquido hasta su uso.

Tabla 5. Características de nueve cepas mexicanas de *Anaplasma marginale*.

Cepa	Región geográfica	Estructura de repetidas de MSP1a
MEX-01-001-01	Aguascalientes, Aguascalientes	4 9 10 11 9
MEX-07-068-01	Pichucalco, Chiapas	α β β Γ β Γ
MEX-14-010-01	Atitalaquia, Hidalgo	τ 57 13 18
MEX-15-099-01	Texcoco, Estado de México	α β β Γ
MEX-17-017-01	Puente de Ixtla, Morelos	12 13 14
MEX-28-037-02	Soto la Marina, Tamaulipas	28 29 74 29 M F
MEX-30-184-02	Tlapacoyan, Veracruz	73 β β β Γ
MEX-30-193-01	Veracruz, Veracruz	α β β Γ
MEX-31-096-01	Tizimín, Yucatán	T C B B C B π

5.2 Secuenciación masiva del DNA genómico

La extracción del DNA genómico (gDNA) se realizó a partir de 200 μ l de sangre de bovino infectada con cada una de las nueve cepas mexicanas de *A. marginale*, al utilizar el kit UltraClean® BloodSpin® DNA Isolation (Laboratorios MO BIO) siguiendo las instrucciones del fabricante. Las nueve muestras del gDNA se almacenaron a -20°C hasta su uso.

La secuenciación masiva del gDNA de las nueve cepas mexicanas de *A. marginale* se realizó a partir de 2 μ g de cada muestra del gDNA mediante una corrida de tipo Paired-End, al utilizar la plataforma MiSeq (tecnología Illumina) en la Universidad de Arizona en Estados Unidos. Como resultado, la información genómica de las lecturas crudas (cada lectura cruda tiene 300 bases de longitud) fue almacenada en 18 archivos con formato FASTQ en el servidor Lotr, el cual está ubicado en las instalaciones del Centro de Investigación en Dinámica Celular (CIDC), UAEM.

Las lecturas editadas se obtuvieron al eliminar las secuencias completas o parciales de los adaptadores de Illumina (los adaptadores fueron incorporados al gDNA durante el proceso de la secuenciación masiva) al utilizar el algoritmo ILLUMINACLIP del programa Trimmomatic (versión 0.36) [51].

Las lecturas recortadas se obtuvieron al eliminar los nucleótidos que tuvieron una calidad Phred menor de 13 ($Q < 13$) en el proceso de la secuenciación masiva

del gDNA, al utilizar el algoritmo dynamictrim de la paquetería de software SolexaQA++ (versión 3.1.7.1) [52].

Las lecturas crudas, editadas y recortadas fueron analizadas con el programa FastQC (versión 0.11.7) (<https://bit.ly/3Cq1qMa>) para obtener la siguiente información: (i) número total de lecturas; (ii) longitud total de las lecturas; (iii) gráficas de calidad Phred por base; y (iv) gráficas de la distribución del contenido G+C por lectura.

5.3 Ensamblaje de los genomas

Los genomas originales de las nueve cepas mexicanas de *A. marginale* se ensamblaron *de novo* con el programa SPAdes (versión 3.11.1) [53], al utilizar las siguientes opciones: (i) solamente se ejecutó el módulo de ensamblaje (--only-assembler); (ii) se redujo el número de desajustes, también llamados mismatches (-careful); y (iii) se utilizaron los valores (números impares) de k-mer entre 21 hasta 127.

Los genomas depurados de las nueve cepas mexicanas de *A. marginale* se obtuvieron al eliminar los contigs que tuvieron una cobertura de ensamblaje menor de 2.0 y/o una longitud menor de 500 pb, al utilizar un script de Perl diseñado por el Dr. Luis Fernando Lozano Aguirre Beltrán (CCG, UNAM).

Los genomas finales de las nueve cepas mexicanas de *A. marginale* se obtuvieron al eliminar los contigs que pertenecen a organismos de otros géneros. Los contigs de los genomas finales cumplieron con tres criterios de selección: (i) los contigs de los genomas depurados se alinearon contra la base de datos de las secuencias de nucleótidos (nr/nt) del género *Anaplasma*, al utilizar el programa Blastn (<https://bit.ly/3sSUMVa>) [54]. Los contigs de los genomas finales tuvieron una cobertura de alineamiento igual o mayor de 50%, una identidad igual o mayor de 70% y un valor de expectancia (E-value) igual o menor de 1e-100; (ii) los contigs de los genomas finales tuvieron un contenido G+C entre 49 y 50%, al utilizar un script de Python diseñado por el M. en B. Fernando Martínez Ocampo (CEIB, UAEM) (Nota: el script de Python está disponible en: <https://bit.ly/3MCKgzq>); y (iii) los contigs de los genomas finales tuvieron una cobertura de ensamblaje igual o mayor de 5.0, al utilizar el programa de ensamblaje SPAdes (versión 3.11.1) [53].

Los genomas originales, depurados y finales de las nueve cepas mexicanas de *A. marginale*, fueron analizados con el programa QUAST (versión 4.6.2) [55] para obtener la siguiente información: (i) número total de contigs; (ii) longitud total del genoma; (iii) longitud N50, y (iv) contenido G+C.

5.4 Análisis de MSP1a

La posible superinfección de diferentes cepas mexicanas de *A. marginale* presentes en el gDNA, fue verificada mediante la detección de dos o más secuencias completas o parciales del gen *mSP1a* en cada uno de los nueve

genomas finales. La secuencia de MSP1a de *A. marginale* St. Maries (número de acceso GenBank AAV86554.1) se alineó contra los contigs que pertenecen a los nueve genomas finales, al utilizar el programa tBlastn (<https://bit.ly/3l0mOZs>) [54]. Las secuencias completas o parciales del gen *mSP1a* que fueron detectadas en los nueve genomas finales, tuvieron una identidad igual o mayor del 50% y un valor de expectancia (E-value) igual o menor de 1e-50.

Las estructuras de repetidas de la región variable de MSP1a de las nueve cepas mexicanas de *A. marginale*, fueron obtenidas al utilizar el programa RepeatAnalyzer (versión 2.8) [56]. La calidad de ensamblaje de las estructuras de repetidas de MSP1a se verificó de manera visual mediante el alineamiento entre las lecturas crudas contra los genes *mSP1a*, al utilizar los programas Bowtie2 (versión 2.3.4.3) [57], SAMtools (versión 1.9) [58] y BamView (versión 17.0.1) [59]. Además, las estructuras de repetidas de MSP1a se compararon contra la base de datos de *A. marginale* (actualizada hasta el día 13 de diciembre de 2017 con 412 cepas y 274 secuencias de repetidas) del programa RepeatAnalyzer (versión 2.8) [56] para obtener la posible distribución geográfica a nivel mundial.

Como resultado, los genomas finales de las cepas MEX-07-068-01 (Pichucalco, Chiapas) y MEX-28-037-02 (Soto la Marina, Tamaulipas) muestran una superinfección, es decir, cada genoma final contiene la información genómica de dos o más cepas mexicanas diferentes de *A. marginale*. Por lo tanto, los dos genomas finales fueron excluidos de los análisis genómicos posteriores.

Las secuencias de los microsatélites que se ubican en la región 5'-UTR (del inglés *Untranslated Region*) de los genes *msp1a*, fueron localizadas entre la secuencia Shine-Dalgarno (SD): GTAGG, y el codón de inicio de la traducción: ATG [60]. Las estructuras de los microsatélites (letras en negrita) se determinaron mediante la secuencia: GTAGG (**G/A TTT**)**m** (**GT**)**n** T ATG [61]. El análisis de los genotipos se realizó de acuerdo con la nomenclatura propuesta por de la Fuente y colaboradores en el año 2007 [44]. Las distancias (en nucleótidos) SD-ATG se calcularon al utilizar la siguiente fórmula: $(4 \times m) + (2 \times n) + 1$ [61].

La cobertura final de ensamblaje de los genomas finales de siete cepas mexicanas de *A. marginale*, fue obtenida mediante el alineamiento entre las lecturas crudas contra los contigs que pertenecen a los genomas finales, al utilizar el programa Bowtie2 (versión 2.3.4.3) [57]. El valor de la cobertura final de ensamblaje de siete genomas finales se calculó al utilizar la siguiente ecuación:

$$\frac{[\text{Número total de lecturas crudas alineadas contra el genoma final}] \times [300 (\text{longitud de las lecturas crudas})]}{[\text{Longitud total del genoma final}]}$$

5.5 Anotación de los genomas

La anotación automática de los genomas finales de las siete cepas mexicanas de *A. marginale* se realizó mediante la predicción de los marcos abiertos de lectura (ORFs del inglés *Open Reading Frames*) y las secuencias codificantes (CDS del inglés *Coding Sequences*), al utilizar el algoritmo Classic RAST del servidor RAST

(del inglés *Rapid Annotation using Subsystem Tecnology*) (versión 2.0) (<https://bit.ly/3tl5KIT>) [62].

La anotación automática de los genes RNA ribosomales (rRNAs) y de transferencia (tRNAs) en los genomas finales de las siete cepas mexicanas de *A. marginale*, se realizó al utilizar los servidores RNAmmer (versión 1.2) (<https://bit.ly/35B7CVL>) [63] y ARAGORN (versión 1.2.38) (<https://bit.ly/3sV0BI7>) [64], respectivamente.

Las proteínas que fueron anotadas en los genomas finales de las siete cepas mexicanas de *A. marginale*, se agruparon en diferentes conjuntos predeterminados de proteínas ortólogas de la base de datos COG (del inglés *Clusters of Orthologous Groups of proteins*) (versión 2003-2014) [65].

La búsqueda de las proteínas que están involucradas con la resistencia a antimicrobianos en los genomas finales de las siete cepas mexicanas de *A. marginale*, se realizó al utilizar los servidores ResFinder (versión 3.0) (<https://bit.ly/3sVeaat>) [66], antiSMASH (versión 3.0) (<https://bit.ly/3hPFH6T>) [67] y CARD 2017 (del inglés *Comprehensive Antibiotic Resistance Database*) (<https://bit.ly/3vPzrnR>) [68]. El servidor CARD 2017 tiene la opción de realizar la búsqueda de las proteínas con los programas BLAST (<https://bit.ly/365TNhy>) y RGI (del inglés *Resistance Gene Identifier*) (versión 4.0.3) (<https://bit.ly/3HX0Bvx>) [68].

La búsqueda de las proteínas que están involucradas con la virulencia y/o la patogenicidad, se realizó al utilizar los servidores VirulenceFinder (versión 1.5) (<https://bit.ly/3sVlgup>) [69], PathogenFinder (versión 1.1) (<https://bit.ly/3vZoeRs>) [70], VFDB (del inglés *Virulence Factor DataBase*) (<https://bit.ly/3KvLwm4>) [71] y PAIDB (del inglés *Pathogenicity Island DataBase*) (versión 2.0) (<https://bit.ly/3KxTVp1>) [72].

5.6 Análisis filogenético

Las secuencias concatenadas de los genes únicos del genoma core (también llamado genoma núcleo o básico) de las siete cepas mexicanas y las cepas americanas (Estados Unidos), australianas, brasileñas y puertorriqueña de *A. marginale* (Tabla 2), y *A. centrale* Israel (número de acceso GenBank CP001759.1) como grupo externo (outgroup), se obtuvieron con los programas Prokka (versión 1.12) [73] y Roary (versión 3.8.0) [74]. Brevemente, los genomas fueron anotados de manera automática al utilizar el programa Prokka (del inglés *rapid prokaryotic genome annotation*). Los archivos con formato GFF (del inglés *General Feature Format*) fueron utilizados como archivos de entrada por el programa Roary (the *pan-genome pipeline*), el cual realizó el alineamiento múltiple entre los genes únicos concatenados de los genomas core al utilizar la opción: create a multiFASTA alignment of core genes (-e). El programa JModelTest (versión 2.1.10) [75] seleccionó el mejor modelo de sustitución de nucleótidos al utilizar el criterio de información Akaike (AIC del inglés *Akaike information criterion*). Dos árboles filogenéticos fueron estimados con el método de máxima verosimilitud (ML del

inglés *Maximum Likelihood*) al utilizar el programa PhyML (versión 3.1) [76] con las opciones: (i) probabilidades bayesianas posteriores (PP del inglés *Bayesian posterior probabilities*); y (ii) 1,000 repeticiones bootstrap. Los dos árboles filogenéticos se visualizaron y editaron al utilizar el programa FigTree (versión 1.4.3) (<https://bit.ly/3KM1SHF>).

5.7 Análisis pangenómico

Dos análisis pangenómicos se realizaron a partir de las proteínas anotadas en los genomas de *A. marginale* al utilizar la paquetería de software GET_HOMOLOGUES (v11042019) [77] con las opciones: (i) entre los genomas finales de las siete cepas mexicanas; y (ii) entre los genomas de las cepas mexicanas, americanas, australianas, brasileñas y puertorriqueña de *A. marginale* (Tabla 2). Brevemente, los archivos con formato FAA (del inglés *Fasta Amino Acid*) obtenidos con el algoritmo Classic RAST del servidor RAST (ver sección 5.5), fueron utilizados como archivos de entrada por la paquetería de software GET_HOMOLOGUES. Los scripts de Perl `get_homologues.pl` y `compare_clusters.pl` se utilizaron para calcular el pangenoma consenso, el cual resultó del alineamiento tipo todos-contra-todos (*all-against-all*) de las proteínas con el programa Blastp, y además, de la agrupación de los resultados al utilizar los algoritmos COGtriangles y OMCL. Como resultado, la matriz binaria (presencia o ausencia de las proteínas) permitió obtener los grupos de proteínas que pertenecen a los diferentes tipos de genomas: unique (proteínas presentes solamente en un genoma), cloud (proteínas presentes entre el 0 y menos del 15% de los genomas), shell (proteínas presentes

entre el 15 y menos del 95% de los genomas), soft core (proteínas presentes entre el 95 y menos del 100% de los genomas) y core (proteínas presentes en el 100% de los genomas).

5.8 Genómica comparativa

Los valores de la identidad promedio de nucleótidos (ANI del inglés *Average Nucleotide Identity*) se calcularon mediante la comparación de los genomas de las cepas mexicanas, americanas, australianas, brasileñas y puertorriqueña de *A. marginale* (Tabla 2), al utilizar los algoritmos ANIb (basado en Blast) y ANIm (basado en MUMmer) del script de Python `calculate_ani.py` (<https://bit.ly/3hS8LuA>).

El orden y la orientación de los contigs que pertenecen a los genomas de las siete cepas mexicanas de *A. marginale*, se obtuvieron mediante el alineamiento entre los contigs de los genomas finales contra el genoma de referencia de *A. marginale* St. Maries (número de acceso GenBank CP000030.1), al utilizar el algoritmo Move Contigs del programa Mauve (versión 2.4.0) [78].

La presencia de eventos evolutivos a gran escala, como el reordenamiento y la inversión de segmentos genómicos, se detectó mediante el alineamiento entre los contigs ordenados de los genomas finales de las siete cepas mexicanas (ver párrafo anterior) y los genomas de las cepas Dawn, Gypsy Plains, Jaboticabal, Palmeira y Florida (números de acceso GenBank CP006847.1, CP006846.1, CP023731.1, CP023730.1 y CP001079.1, respectivamente), contra el genoma de

referencia de *A. marginale* St. Maries (número de acceso GenBank CP000030.1), al utilizar el algoritmo “Align with progressiveMauve...” del programa Mauve (versión 2.4.0) [78].

El análisis de genómica comparativa a nivel de sintenia se realizó mediante el alineamiento entre los genomas de las siete cepas mexicanas y las cepas Dawn, Gypsy Plains, Jaboticabal, Palmeira y Florida (números de acceso GenBank CP006847.1, CP006846.1, CP023731.1, CP023730.1 y CP001079.1, respectivamente), contra el genoma de referencia de *A. marginale* St. Maries (número de acceso GenBank CP000030.1), al utilizar el programa BRIG (del inglés *BLAST Ring Image Generator*) (v0.95) [79].

5.9 Identificación de los posibles candidatos vacunales

La anotación automática de las CDS que pertenecen a los genomas finales de las siete cepas mexicanas de *A. marginale*, se realizó al utilizar los algoritmos Classic RAST (ver sección 5.5) y RASTtk del servidor RAST (versión 2.0) (<https://bit.ly/3IZirz1>) [62], y los programas Prokka (versión 1.12) [73] (ver sección 5.6) y GeneMark (versión 4.33) [80]. Las proteínas que fueron anotadas con el algoritmo Classic RAST se utilizaron para los análisis posteriores.

La predicción de la localización subcelular de las proteínas de membrana externa y extracelulares que pertenecen a las siete cepas mexicanas de *A. marginale*, se determinó al utilizar los servidores PSORTb (versión 3.0.2)

(<https://bit.ly/3MB63HF>) [81], CELLO (versión 2.5) (<https://bit.ly/3hUabo4>) [82] y SOSUIGramN (<https://bit.ly/3MDBx04>) [83]. Las proteínas de membrana externa y extracelulares que fueron predichas por al menos dos servidores, se utilizaron para los análisis posteriores.

La predicción de los antígenos protectores se realizó al utilizar el servidor VaxiJen (versión 2.0) (<https://bit.ly/3KxArkK>) [84] con un valor de corte igual a 0.7, los cuales deben tener las siguientes características: (i) estar conservados en las siete cepas mexicanas (cobertura igual o mayor del 90% e identidad igual o mayor a 98%); (ii) no mostrar homología con las proteínas de eucariotas (cobertura menor del 60% y similitud menor del 60%); y (iii) también fueron anotados por el algoritmo RASTtk del servidor RAST y los programas Prokka y GenMark (cobertura igual o mayor del 90% e identidad igual o mayor a 98%) (ver el primer párrafo de la sección 5.9).

La predicción de los epítomos de tipo B de los antígenos protectores se realizó al utilizar los servidores: (i) Bcepred (<https://bit.ly/3hYvMf3>) [85]; (ii) ABCpred (<https://bit.ly/3vS0wXB>) [86]; y (iii) IEDB-AR (del inglés *Immune Epitope DataBase Analysis Resource*) (versión 2.19) (<https://bit.ly/3tKyt9A>) [87]. Los epítomos de tipo B fueron predichos por al menos un servidor y tienen una longitud entre 10 y 25 aa.

6. RESULTADOS Y DISCUSIÓN

6.1 Selección de cepas mexicanas de *Anaplasma marginale*

En México, estudios serológicos y moleculares han reportado la presencia de anaplasmosis bovina en las regiones tropicales y subtropicales del país [48,88–90]. La Figura 1 muestra que la anaplasmosis bovina tiene una prevalencia mayor del 50% en seis estados que comprenden el Golfo de México y la península de Yucatán (Campeche, Quintana Roo, Tabasco, Tamaulipas, Veracruz y Yucatán) y tres estados que colindan con el océano Pacífico (Chiapas, Nayarit y Sinaloa). Además, la anaplasmosis bovina tiene una prevalencia del 10 al 50% en ocho estados del centro del país (Aguascalientes, Estado de México, Guanajuato, Hidalgo, Morelos, Puebla, Querétaro y Tlaxcala) y cinco estados que colindan con el océano Pacífico (Colima, Guerrero, Jalisco, Michoacán y Oaxaca) [48].



Figura 1. Distribución de la prevalencia de la anaplasmosis bovina en México. En el año 2020, se reportaron 35.6 millones de cabezas de ganado en el país (ver la sección 1.4). La prevalencia de la anaplasmosis bovina en cada estado de la república mexicana se muestra con el siguiente código de colores: el blanco tiene una prevalencia del 1 al 10% y 7.1 millones de cabezas de ganado (20%); el verde tiene una prevalencia del 10 al 50% y 15.7 millones de cabezas de ganado (44%); y el azul tiene una prevalencia mayor del 50% y 12.8 millones de cabezas de ganado (36%) (Figura editada de Rodríguez Camarillo SD., *et al.* 2009) [48].

Las nueve cepas mexicanas de *A. marginale* que fueron seleccionadas para el presente trabajo, provienen de ocho diferentes estados de la república mexicana y tienen ocho diferentes estructuras de repetidas de MSP1a (Tabla 5 y Figura 2).



Figura 2. Distribución geográfica de nueve cepas mexicanas de *A. marginale*. El nombre del municipio y estado donde fue obtenida cada cepa de *A. marginale* se muestra entre paréntesis. La cepa de alto grado de virulencia se muestra en letras de color rojo y la cepa de bajo grado de virulencia se muestra en letras de color azul.

Cinco cepas mexicanas provienen de cuatro estados que tienen una prevalencia de la anaplasmosis bovina mayor del 50%: las cepas MEX-07-068-01

(Pichucalco, Chiapas), MEX-28-037-02 (Soto la Marina, Tamaulipas), MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-30-193-01 (Veracruz, Veracruz) fueron obtenidas a partir de bovinos infectados en campo que presentaron signos clínicos típicos de la enfermedad [49,91,92]; y la cepa MEX-31-096-01 (Tizimín, Yucatán) fue obtenida a partir de un bovino portador asintomático [50]. Además, tres cepas aisladas en México: MEX-07-068-01 (Pichucalco, Chiapas), MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-30-193-01 (Veracruz, Veracruz) comparten las repetidas α , β y Γ en diferentes estructuras de MSP1a, mientras que las dos cepas mexicanas restantes: MEX-28-037-02 (Soto la Marina, Tamaulipas) y MEX-31-096-01 (Tizimín, Yucatán) no comparten secuencias de repetidas de MSP1a (Tabla 5).

Por otra parte, las cuatro cepas mexicanas que provienen de cuatro estados del centro del país tienen una prevalencia de la anaplasmosis bovina del 10 al 50%: las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 (Atitalaquia, Hidalgo) y MEX-17-017-01 (Puente de Ixtla, Morelos); fueron obtenidas a partir de bovinos infectados en campo que presentaron signos clínicos típicos de la enfermedad [49,91]; y la cepa MEX-15-099-01 (Texcoco, Estado de México) fue obtenida a partir de un brote agudo en un lote de acopio [50]. Además, estas cuatro cepas mexicanas no comparten secuencias de repetidas de MSP1a (Tabla 5).

Es importante mencionar que las cepas MEX-15-099-01 (Texcoco, Estado de México) y MEX-30-193-01 (Veracruz, Veracruz) comparten la misma estructura de repetidas de MSP1a (α β β Γ), pero provienen de diferentes regiones geográficas

con diferente porcentaje de prevalencia de la anaplasmosis bovina (Tabla 5 y Figuras 1 y 2). No existe información concluyente para determinar si las dos cepas mexicanas corresponden al mismo microorganismo, el cual probablemente consiguió trasladarse de una región geográfica a otra mediante el movimiento del ganado o vectores biológicos (garrapatas).

Dos cepas mexicanas que tienen reportado un diferente grado de virulencia, fueron incluidas para el presente trabajo. En el año 2000, García Ortiz y colaboradores evaluaron el grado de virulencia de las cepas MEX-15-099-01 (Texcoco, Estado de México) y MEX-31-096-01 (Tizimín, Yucatán) [50]. Brevemente, dos grupos de cuatro bovinos fueron inoculados con una dosis de 1×10^6 eritrocitos infectados por bovino. Como resultado, todos los animales infectados con la cepa MEX-15-099-01 (Texcoco, Estado de México) desarrollaron anaplasmosis clínica, mientras que solo un animal infectado con la cepa MEX-31-096-01 (Tizimín, Yucatán) desarrolló la enfermedad. Los autores del trabajo concluyeron que la cepa MEX-15-099-01 (Texcoco, Estado de México) muestra un mayor grado de virulencia [50]. Por otra parte, en el año 2008, Rodríguez Camarillo y colaboradores evaluaron el grado de virulencia de la cepa MEX-31-096-01 (Tizimín, Yucatán) [32]. Brevemente, 113 bovinos fueron inoculados con el patógeno a dosis crecientes. Como resultado, sólo un animal requirió tratamiento para la enfermedad clínica. Los autores del trabajo concluyeron que la cepa MEX-31-096-01 (Tizimín, Yucatán) tiene un bajo grado de virulencia [32]. Por lo tanto, en el presente trabajo se considera que las cepas MEX-15-099-01 (Texcoco, Estado

de México) y MEX-31-096-01 (Tizimín, Yucatán) tienen un alto y bajo grado de virulencia, respectivamente (Figura 2).

6.2 Secuenciación masiva del DNA genómico

A. marginale es una bacteria intracelular obligada que es difícil de cultivar *in vitro*. La línea celular IDE8 fue obtenida a partir de embriones de la garrapata *Ixodes scapularis* [93] y ha sido utilizada para establecer algunas cepas de *A. marginale* a nivel mundial [94–96]. En México, algunas cepas de *A. marginale* se han intentado establecer y mantener en condiciones de laboratorio sin obtener resultados concluyentes [97]. Actualmente, en el Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad (CENID-SAI) del INIFAP, se realiza el trabajo de obtener y desarrollar líneas celulares en condiciones de laboratorio, con el objetivo de establecer un modelo que permita el cultivo *in vitro* de diferentes cepas mexicanas de *A. marginale* [98,99]. Por lo tanto, el gDNA de cepas mexicanas de *A. marginale* solamente se puede extraer a partir de la sangre de bovinos infectados, lo cual dificulta obtener el gDNA puro de *A. marginale* y favorece la presencia del gDNA de otros organismos, por ejemplo, de bovinos y otros patógenos.

En el año 1993, Alleman y colaboradores determinaron que el genoma de *A. marginale* está constituido por un cromosoma circular con una longitud total entre 1.20 y 1.26 Mb, y un contenido G+C de 56% en moles, lo cual se determinó mediante una electroforesis en gel de campo pulsado (PFGE por las siglas en

inglés) y un análisis espectral, respectivamente [100]. Hasta la fecha, 17 genomas de cepas australianas, brasileñas, americanas y puertorriqueña de *A. marginale*, han sido depositados en la base de datos GenBank con una longitud total entre 1.13 y 1.40 Mb, y un contenido G+C entre 46.73 y 49.84% (Tabla 2), lo cual es inusualmente elevado para un organismo intracelular obligado. Ya que los genomas de otras especies que pertenecen al orden de los Rickettsiales tienen un contenido G+C promedio de 31% [101].

La secuenciación masiva del gDNA de las nueve cepas mexicanas de *A. marginale*, generó lecturas recortadas de alrededor de 83 y 281 Mb de información genómica con una elevada calidad Phred (Q>13, es decir, con una precisión de la base mayor del 90%), que además, no contienen secuencias de los adaptadores de Illumina (Tabla 6 y lado izquierdo de las Figuras 3-11).

Tabla 6. Características de las lecturas crudas, editadas y recortadas de las nueve cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Lecturas crudas		Lecturas editadas		Lecturas recortadas	
		Número total de lecturas	Longitud total (nucleótidos)	Número total de lecturas	Longitud total (nucleótidos)	Número total de lecturas	Longitud total (nucleótidos)
MEX-01-001-01	Aguascalientes, Aguascalientes	1,100,300	330,090,000	1,100,300	324,358,591	1,100,208	160,181,089
MEX-07-068-01	Pichucalco, Chiapas	660,182	198,054,600	660,182	193,924,310	660,010	101,797,462
MEX-14-010-01	Atitalaquia, Hidalgo	1,418,888	425,666,400	1,418,888	416,290,003	1,418,550	227,063,992
MEX-15-099-01	Texcoco, Edo. de México	605,370	181,611,000	605,370	178,268,652	605,318	91,290,317
MEX-17-017-01	Puente de Ixtla, Morelos	567,482	170,244,600	567,482	166,200,132	567,434	83,320,463
MEX-28-037-02	Soto la Marina, Tamaulipas	1,408,630	422,589,000	1,408,630	412,156,618	1,408,550	213,544,001
MEX-30-184-02	Tlapacoyan, Veracruz	671,050	201,315,000	671,050	193,209,796	669,396	102,199,946
MEX-30-193-01	Veracruz, Veracruz	940,598	282,179,400	940,598	270,978,018	940,392	173,803,802
MEX-31-096-01	Tizimin, Yucatán	1,671,658	501,497,400	1,671,658	493,256,198	1,671,606	281,138,662

Las gráficas de distribución del contenido G+C por lectura de las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-07-068-01 (Pichucalco, Chiapas), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-17-017-01 (Puente de Ixtla, Morelos), MEX-28-037-02 (Soto la Marina, Tamaulipas), MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-30-193-01 (Veracruz, Veracruz), muestran la presencia de una cresta de ~49% (lado derecho de las Figuras 3-5 y 7-10, respectivamente). El contenido G+C es similar a los valores reportados para los 17 genomas de cepas australianas, brasileñas, americanas y puertorriqueña de *A. marginale* (Tabla 2). Por lo tanto, la mayor cantidad de información genómica pertenece a las siete cepas mexicanas de *A. marginale*.

Las gráficas de distribución del contenido G+C por lectura de la cepa MEX-15-099-01 (Texcoco, Estado de México), muestran la presencia de dos crestas, principal y secundaria de ~49 y ~34%, respectivamente (lado derecho de la Figura 6). Por lo tanto, la mayor cantidad de información genómica pertenece a la cepa mexicana de *A. marginale*, y la menor cantidad de información genómica pertenece a otro patógeno que fue reportado como *Mycoplasma wenyonni* INIFAP02 (Apéndices 12 y 13). Por otra parte, la cepa MEX-31-096-01 (Tizimín, Yucatán), muestra la presencia de dos crestas, principal y secundaria de ~31 y ~49%, respectivamente (lado derecho de la Figura 11). Por lo tanto, la mayor cantidad de información genómica pertenece a otro patógeno que fue reportado como 'Candidatus *Mycoplasma haemobos*' INIFAP01 (Apéndices 10 y 11), y la menor cantidad de información genómica pertenece a la cepa mexicana de *A. marginale*.

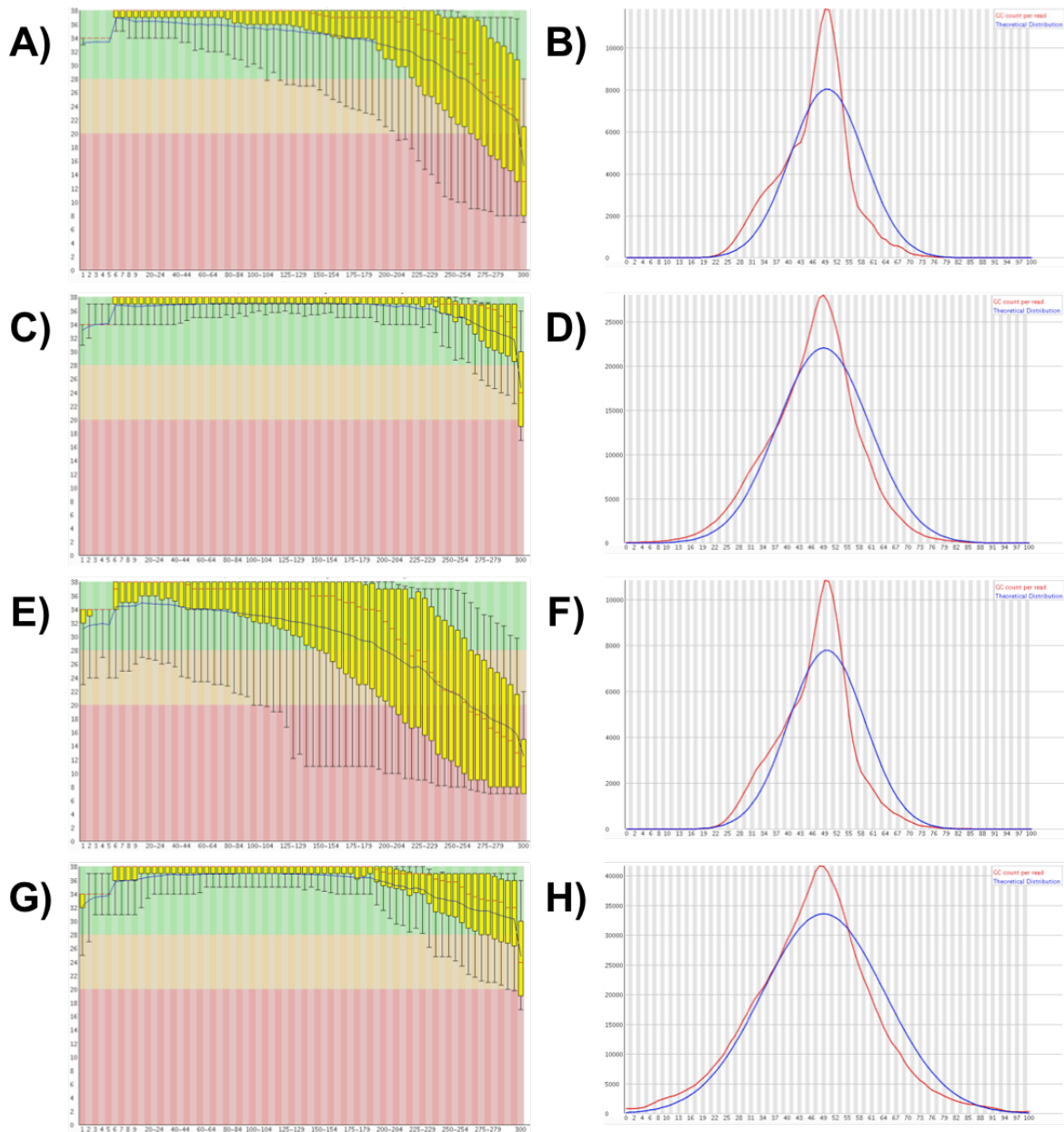


Figura 3. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H) de *A. marginale* MEX-01-001-01 (Aguascalientes, Aguascalientes). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 10,000 (B y F), 25,000 (D) o 40,000 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

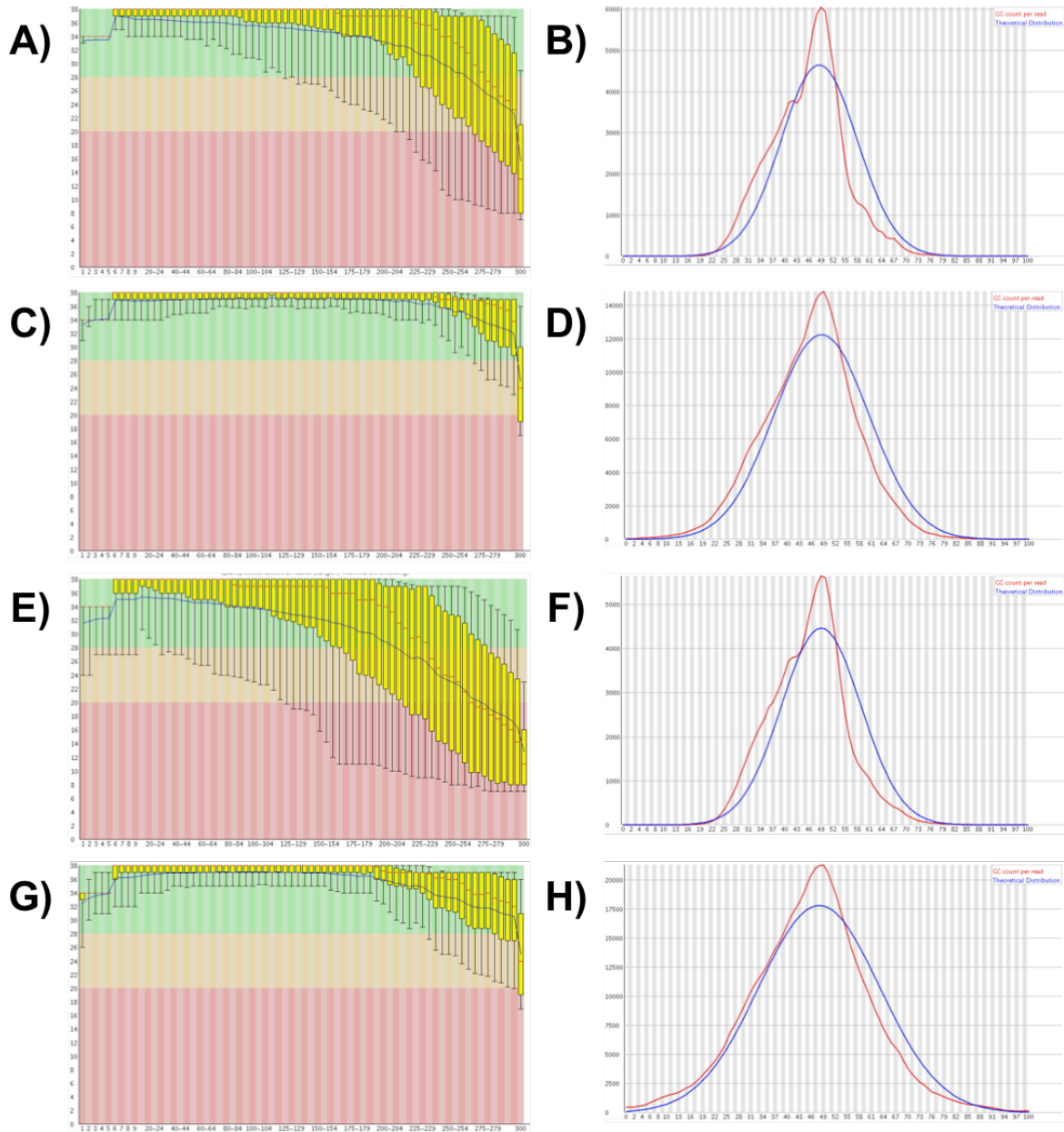


Figura 4. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-07-068-01 (Pichucalco, Chiapas). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 6,000 (B), 14,000 (D), 5,000 (F) o 20,000 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

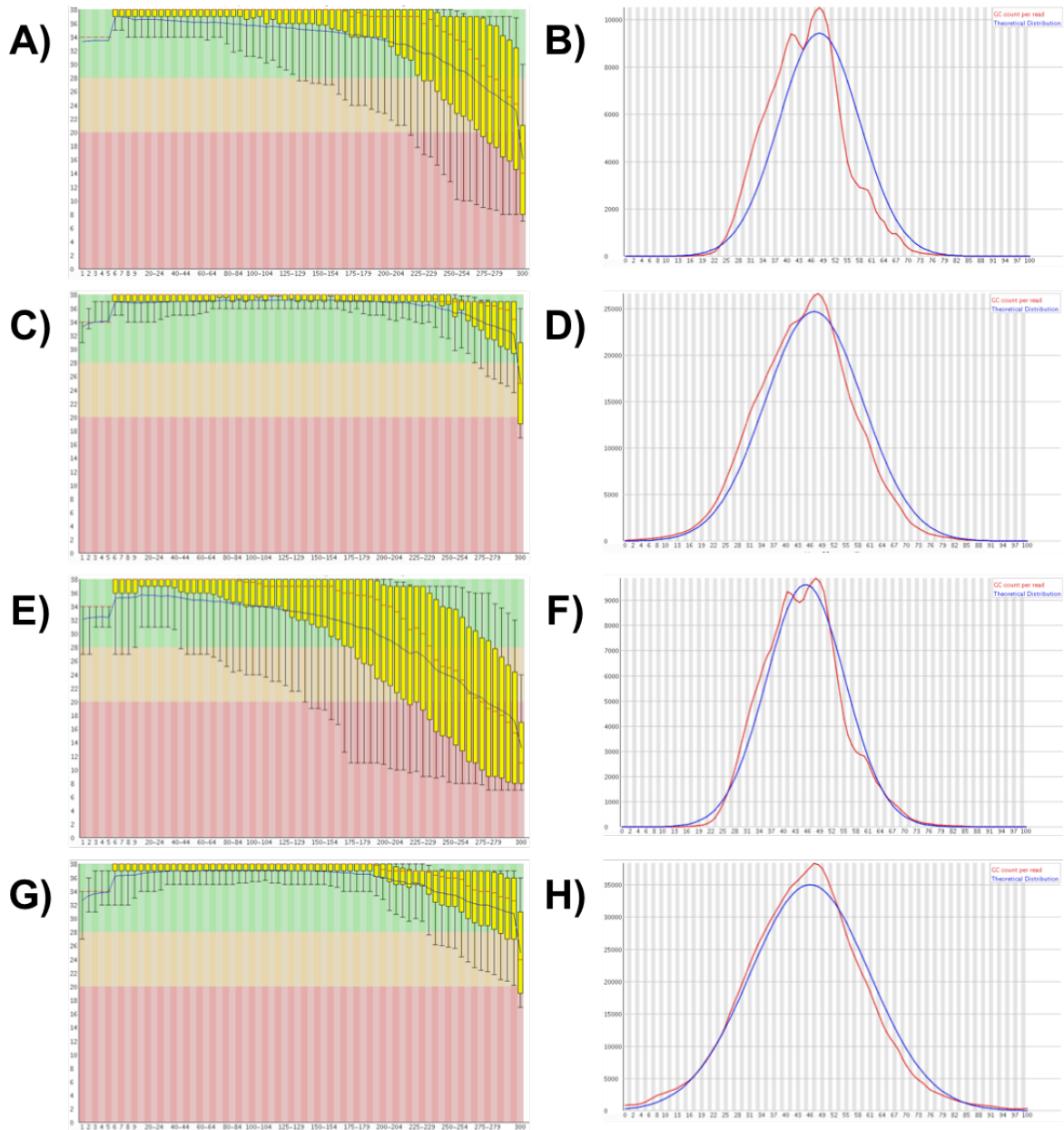


Figura 5. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-14-010-01 (Atitalaquia, Hidalgo). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 10,000 (B), 25,000 (D), 9,000 (F) o 35,000 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

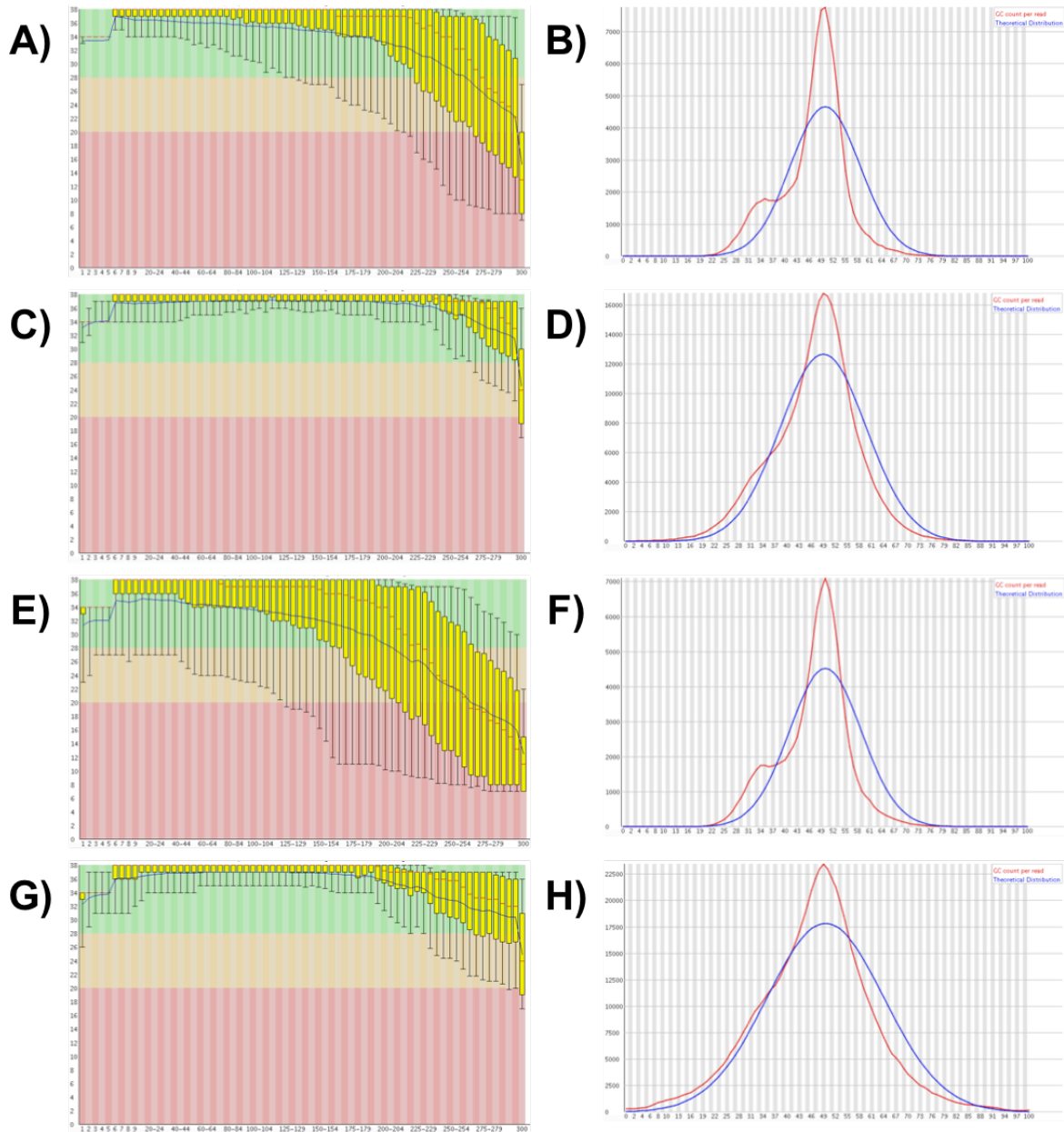


Figura 6. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-15-099-01 (Texcoco, Estado de México). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 7,000 (B y F), 16,000 (D) o 22,500 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

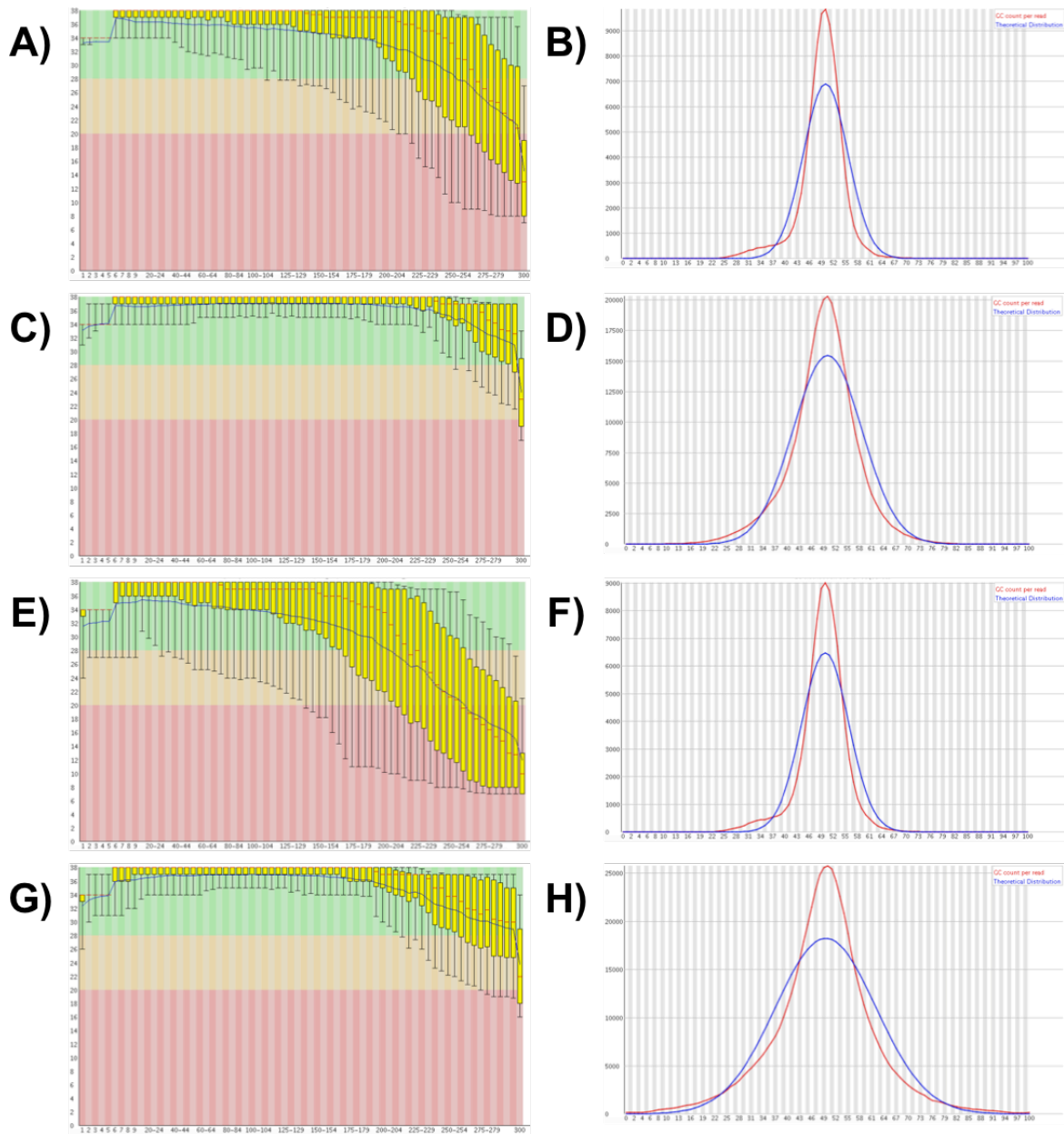


Figura 7. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-17-017-01 (Puente de Ixtla, Morelos). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 9,000 (B y F), 20,000 (D) o 25,000 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

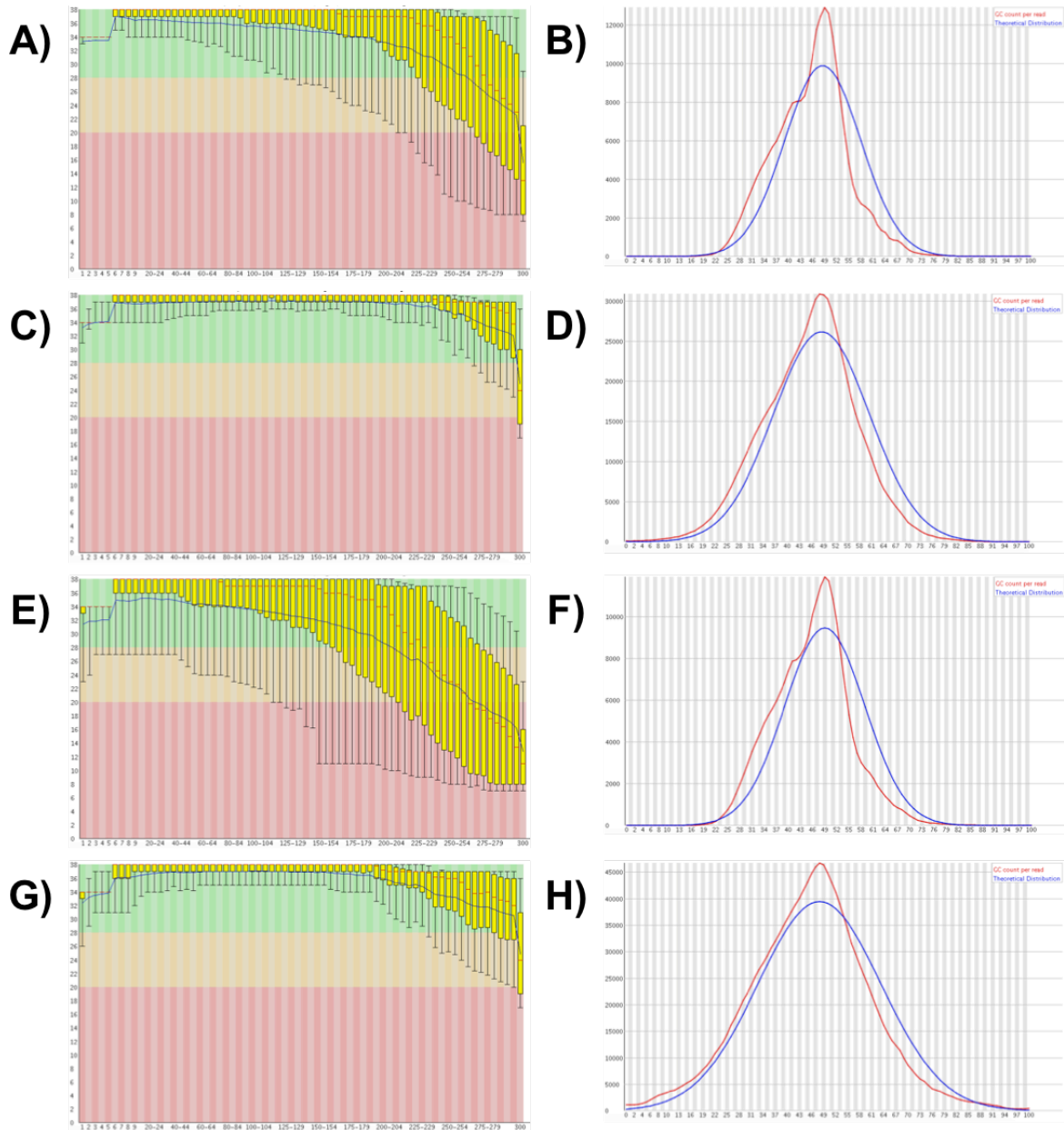


Figura 8. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-28-037-02 (Soto la Marina, Tamaulipas). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 12,000 (B), 30,000 (D), 10,000 (F) o 45,000 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

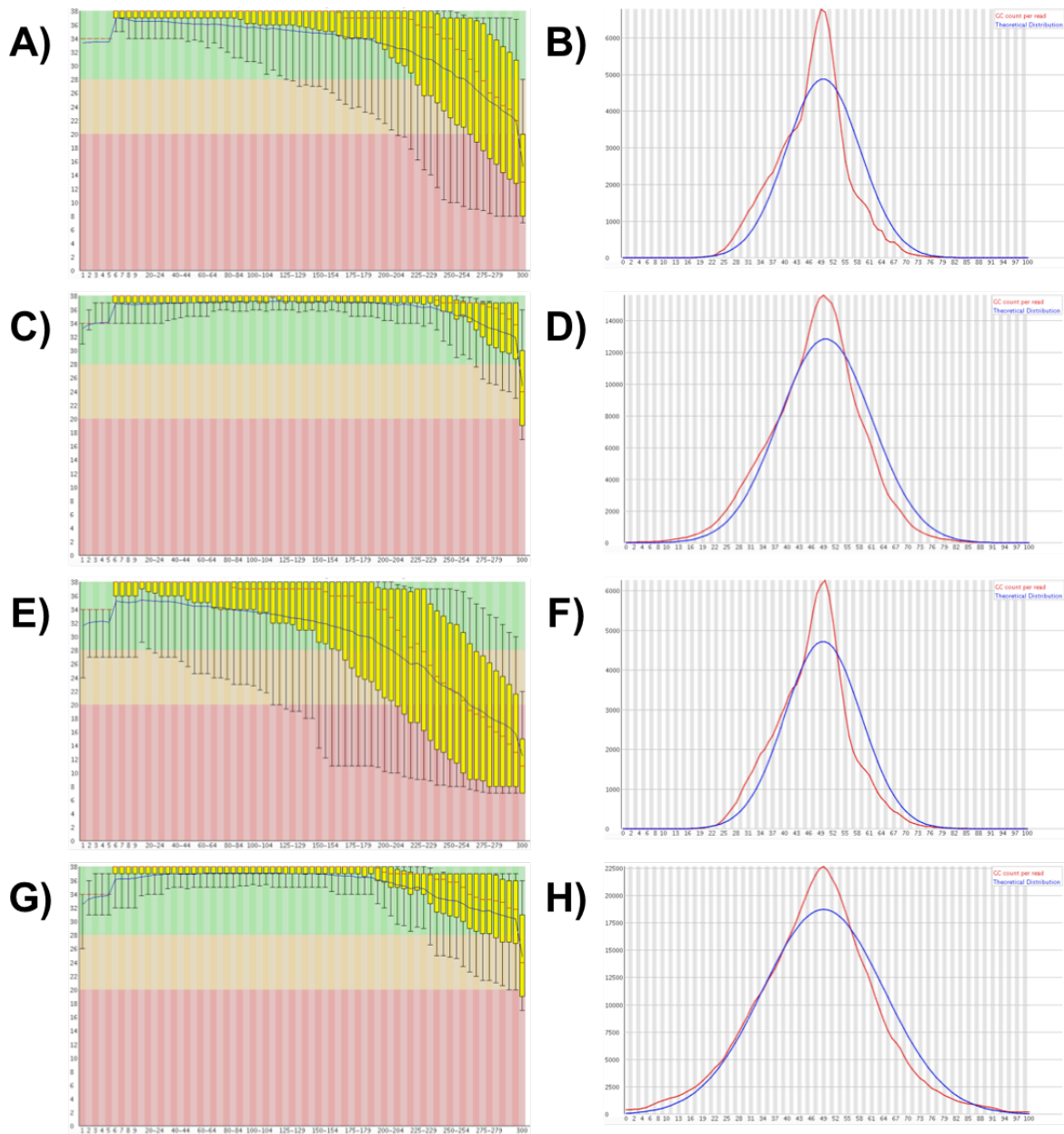


Figura 9. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-30-184-02 (Tlapacoyan, Veracruz). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 6,000 (B y F), 14,000 (D) o 22,500 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

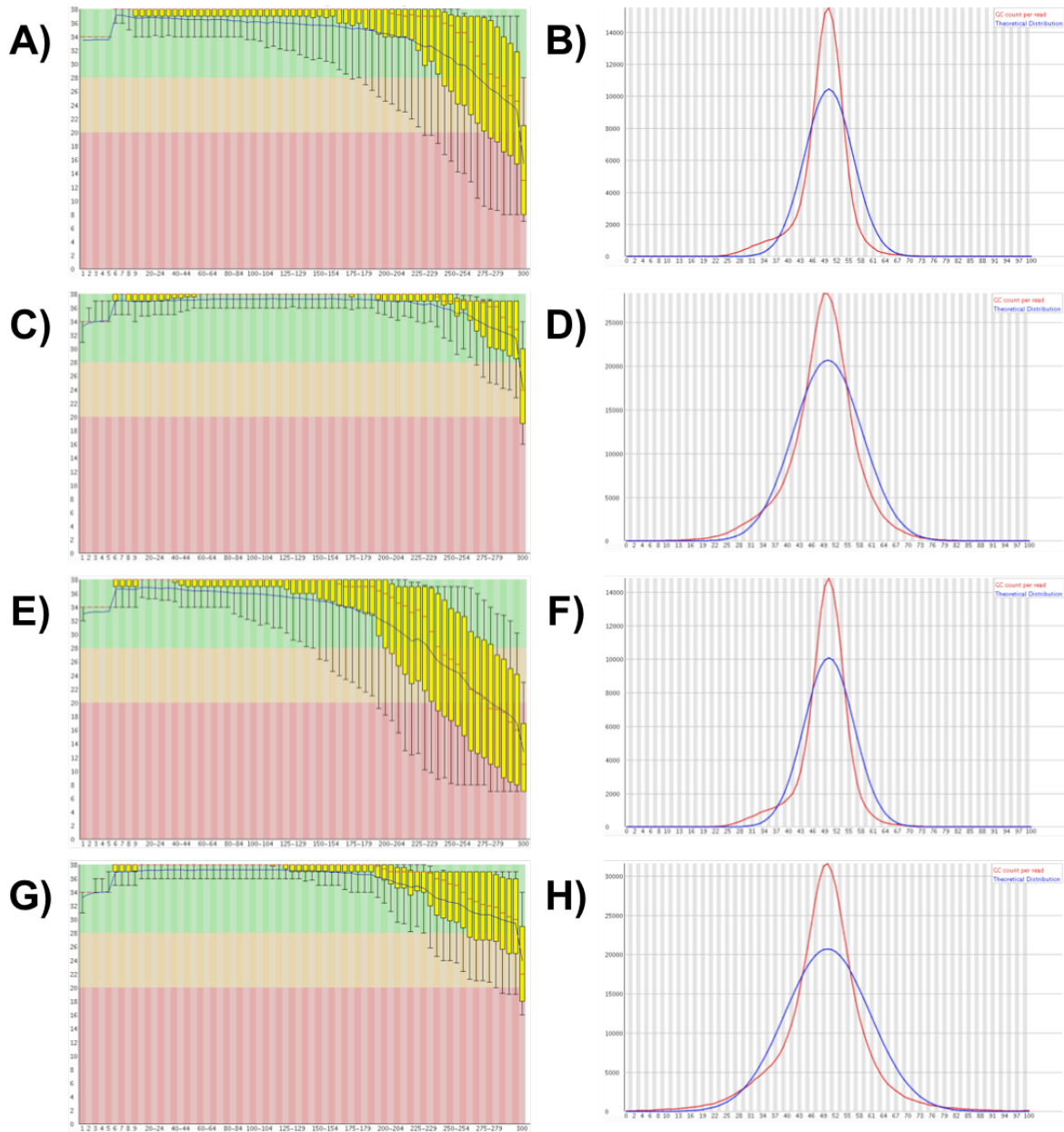


Figura 10. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-30-193-01 (Veracruz, Veracruz). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 14,000 (B y F), 25,000 (D) o 30,000 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

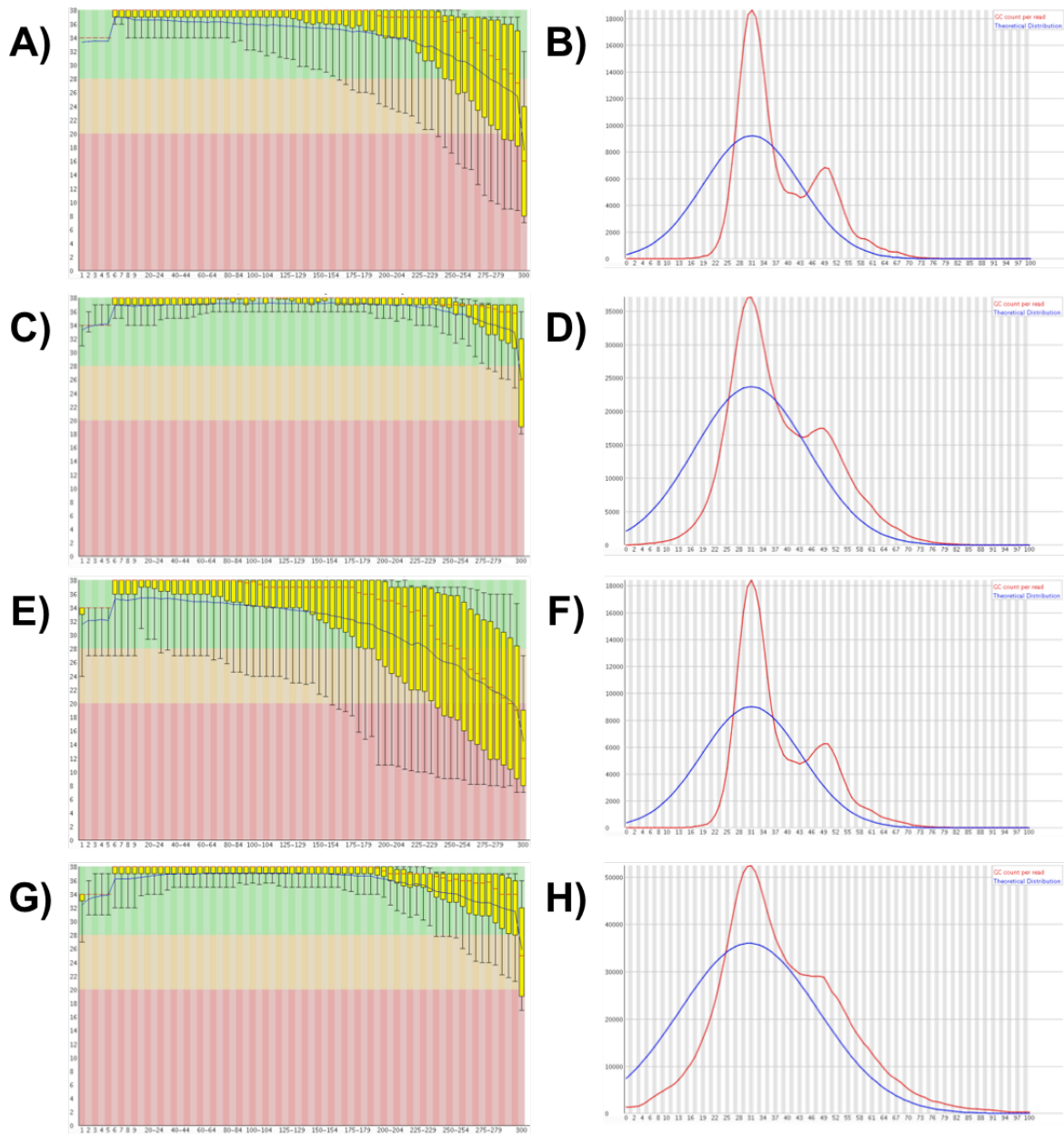


Figura 11. Gráficas de calidad Phred por base (izquierda) y distribución del contenido G+C por lectura (derecha) de las lecturas R1 (A-D) y R2 (E-H), para las lecturas crudas (A-B y E-F) y recortadas (C-D y G-H), de *A. marginale* MEX-31-096-01 (Tizimín, Yucatán). Las gráficas de calidad Phred por base muestran en el eje horizontal (abscisas) la posición de las 300 bases de las lecturas; y en el eje vertical (ordenadas) muestran los valores de calidad Phred entre 0 y 40: valores entre 0 y 20 tienen baja calidad (región roja), valores entre 20 y 28 tienen moderada calidad (región naranja) y valores entre 28 y 40 tienen elevada calidad (región verde). Las líneas rojas representan el valor de la mediana, las líneas azules representan el promedio de la calidad, las cajas amarillas representan el rango inter-cuartil (25 al 75%) y los bigotes inferiores y superiores representan los puntos del 10 y 90%, respectivamente. Por otra parte, las gráficas de distribución del contenido G+C por lectura muestran en el eje horizontal (abscisas) el porcentaje del contenido G+C entre 0 y 100; y en el eje vertical (ordenadas) muestran el número de lecturas entre 0 y 18,000 (B y F), 35,000 (D) o 50,000 (H). Las líneas rojas representan la distribución real y las líneas azules representan la distribución teórica del contenido G+C por lectura.

6.3 Ensamblaje de los genomas

En el año 2004, el genoma completo de *A. marginale* St. Maries fue depositado en la base de datos GenBank (número de acceso GenBank CP000030.1). Posteriormente, 16 genomas de cepas americanas, australianas, brasileñas y puertorriqueña de *A. marginale* han sido reportados a nivel mundial. Estos 17 genomas de *A. marginale* tienen una longitud total entre 1.13 y 1.40 Mb con un contenido G+C entre 46.73 y 49.84% (Tabla 2). Sin embargo, tres borradores de genomas de las cepas americanas: Okeechobee, South Idaho y Washington Okanogan (números de acceso GenBank AFMV00000000.1, AFMY00000000.1 y AFMW00000000.1, respectivamente), están ensamblados en más de 300 contigs con una longitud total mayor de 1.3 Mb y un contenido G+C menor de 47.49% (Tabla 2). En el presente trabajo, estos tres borradores de genomas están considerados como ensamblajes de baja calidad. Por otra parte, seis genomas de las cepas Dawn, Florida, Gypsy Plains, Jaboticabal, Palmeira y St. Maries (números de acceso GenBank CP006847.1, CP001079.1, CP006846.1, CP023731.1, CP023730.1 y CP000030.1, respectivamente), están ensamblados en un cromosoma con una longitud total entre 1.19 y 1.20 Mb, y un contenido G+C entre 49.72 y 49.77% (Tabla 2). En el presente trabajo, estos seis genomas cerrados en un cromosoma están considerados como ensamblajes de alta calidad. Los datos anteriores sugieren que los genomas de las nueve cepas mexicanas de *A. marginale* tienen una longitud total de ~1.19 Mb con un contenido G+C de ~49.7%.

Los genomas depurados de las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-07-068-01 (Pichucalco, Chiapas), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-17-017-01 (Puente de Ixtla, Morelos), MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-30-193-01 (Veracruz, Veracruz) se ensamblaron entre 58 y 121 contigs con una longitud total entre 1.19 y 1.24 Mb, y un contenido G+C entre 48.99 y 50% (Tabla 7). Los genomas depurados de estas seis cepas mexicanas muestran una longitud total y un contenido G+C similar a los valores reportados para los 17 genomas de cepas americanas, australianas, brasileñas y puertorriqueña de *A. marginale* (Tabla 2). Los seis genomas depurados se ensamblaron en menos de 122 contigs, lo cual sugiere que cada genoma depurado contiene la información genómica de una cepa mexicana de *A. marginale*.

Tabla 7. Características de los genomas originales y depurados de las nueve cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Genomas originales			Genomas depurados		
		Número de contigs	Longitud total (pb)	Contenido G+C (%)	Número de contigs	Longitud total (pb)	Contenido G+C (%)
MEX-01-001-01	Aguascalientes, Aguascalientes	2,502	2,263,366	48.83	70	1,224,488	49.78
MEX-07-068-01	Pichucalco, Chiapas	3,766	2,944,874	48.73	71	1,214,436	49.79
MEX-14-010-01	Atitalaquia, Hidalgo	9,135	5,444,003	44.76	97	1,249,350	50.00
MEX-15-099-01	Texcoco, Edo. de México	912	2,162,604	44.27	82	1,798,848	44.35
MEX-17-017-01	Puente de Ixtla, Morelos	541	1,466,308	47.71	121	1,247,542	48.99
MEX-28-037-02	Soto la Marina, Tamaulipas	3,660	2,111,924	49.84	1,078	1,436,355	49.84
MEX-30-184-02	Tlapacoyan, Veracruz	1,835	1,938,229	49.48	66	1,214,902	49.82
MEX-30-193-01	Veracruz, Veracruz	148	1,217,727	49.73	58	1,193,748	49.73
MEX-31-096-01	Tizimín, Yucatán	4,074	3,709,717	41.44	135	2,151,655	41.36

El genoma depurado de la cepa MEX-28-037-02 (Soto la Marina, Tamaulipas) se ensambló en 1,078 contigs con una longitud total de 1.43 Mb, y un contenido G+C de 49.84% (Tabla 7). El genoma depurado de la cepa mexicana muestra una longitud total y un contenido G+C similar a los valores reportados para los 17 genomas de cepas americanas, australianas, brasileñas y puertorriqueña de *A. marginale* (Tabla 2). Sin embargo, el genoma depurado de la cepa mexicana se ensambló en más de 1,000 contigs, lo cual sugiere que la información genómica de *A. marginale* está duplicada, es decir, existe la posibilidad que el genoma depurado contenga la información genómica de dos o más cepas mexicanas diferentes de *A. marginale*.

Los genomas depurados de las cepas MEX-15-099-01 (Texcoco, Estado de México) y MEX-31-096-01 (Tizimín, Yucatán) tienen una longitud total de 1.79 y 2.1 Mb, y un contenido G+C de 44.35 y 41.36%, respectivamente (Tabla 7). Los genomas depurados de las dos cepas mexicanas muestran una longitud total mayor y un contenido G+C considerablemente menor a los valores reportados para los 17 genomas de cepas americanas, australianas, brasileñas y puertorriqueña de *A. marginale* (Tabla 2). Las características de los genomas depurados de las dos cepas mexicanas sugieren que la información genómica contiene una combinación del gDNA de *A. marginale* con otros patógenos, lo cual coincide con los datos obtenidos de las gráficas de la distribución del contenido G+C por lectura en la sección 6.2.

Por otra parte, los genomas finales de las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-07-068-01 (Pichucalco, Chiapas), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México), MEX-17-017-01 (Puente de Ixtla, Morelos), MEX-30-184-02 (Tlapacoyan, Veracruz), MEX-30-193-01 (Veracruz, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) se ensamblaron entre 32 y 46 contigs con una longitud total entre 1.16 y 1.17 Mb, y un contenido G+C entre 49.78 y 49.80% (Tabla 8). Los genomas finales de las ocho cepas mexicanas muestran una longitud total y un contenido G+C similar a los valores reportados para los seis genomas que están ensamblados en un cromosoma (Tabla 2). Sin embargo, los genomas finales de las ocho cepas mexicanas se ensamblaron en un número considerablemente menor a los valores reportados para los 11 borradores de genomas de cepas americanas y puertorriqueña de *A. marginale* (Tabla 2), lo cual sugiere que los ocho genomas finales tienen una elevada calidad de ensamblaje.

Tabla 8. Características de los genomas finales de las nueve cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Número de contigs	Longitud total (pb)	Longitud N50 (pb)	Contenido G+C (%)
MEX-01-001-01	Aguascalientes, Aguascalientes	34	1,179,425	65,428	49.79
MEX-07-068-01	Pichucalco, Chiapas	46	1,174,034	58,899	49.78
MEX-14-010-01	Atitalaquia, Hidalgo	46	1,172,327	85,147	49.79
MEX-15-099-01	Texcoco, Estado de México	32	1,169,440	85,955	49.79
MEX-17-017-01	Puente de Ixtla, Morelos	41	1,172,716	65,203	49.79
MEX-28-037-02	Soto la Marina, Tamaulipas	982	1,345,681	1,706	49.79
MEX-30-184-02	Tlapacoyan, Veracruz	40	1,176,681	61,328	49.79
MEX-30-193-01	Veracruz, Veracruz	41	1,167,111	65,442	49.80
MEX-31-096-01	Tizimín, Yucatán	43	1,176,579	62,699	49.79

El genoma final de la cepa MEX-28-037-02 (Soto la Marina, Tamaulipas) se ensambló en 982 contigs con una longitud total de 1.34 Mb, y un contenido G+C de 49.79% (Tabla 8). El genoma final de la cepa mexicana muestra una longitud total y un contenido G+C similar a los valores reportados para los 17 genomas de cepas americanas, australianas, brasileñas y puertorriqueña de *A. marginale* (Tabla 2). Sin embargo, el genoma final de la cepa mexicana se ensambló en más de 900 contigs con una longitud N50 de 1.7 Kb (Tabla 8), lo cual sugiere que el genoma final tiene una baja calidad de ensamblaje porque contiene la información genómica de dos o más cepas mexicanas diferentes de *A. marginale*.

6.4 Análisis de MSP1a

En la actualidad, diversos trabajos han reportado que las regiones tropicales que tienen un elevado porcentaje de prevalencia de la anaplasmosis bovina, muestran tasas de superinfección de *A. marginale* que son considerablemente mayores en comparación con las regiones templadas o tropicales que tienen un menor porcentaje de prevalencia de la enfermedad [25,102,103]. En el año 2015, Castañeda-Ortiz y colaboradores utilizaron los genes *msp1a* (un alelo por genoma) y *msp2* (múltiples alelos por genoma) para determinar que el grado de superinfección de *A. marginale* es considerablemente mayor en el rebaño del Verdineño, Nayarit (región tropical que tiene una prevalencia de la anaplasmosis bovina mayor del 50%) en comparación con el rebaño de la Joya, Jalisco (región tropical que tiene una prevalencia de la anaplasmosis bovina del 10 al 50%). Cabe mencionar que las dos regiones geográficas comparten a la garrapata

Rhipicephalus (Boophilus) microplus como vector biológico [25]. En el presente trabajo, el gen *msp1a* fue utilizado para detectar la posible superinfección de *A. marginale* en los genomas finales de las nueve cepas mexicanas de *A. marginale*. Por otra parte, los nueve genomas finales solamente contienen secuencias parciales de las múltiples copias del gen *msp2*. Por lo tanto, el gen *msp2* no fue utilizado en el presente trabajo para detectar la posible superinfección de *A. marginale*.

Los genomas finales de las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México), MEX-17-017-01 (Puente de Ixtla, Morelos), MEX-30-184-02 (Tlapacoyan, Veracruz), MEX-30-193-01 (Veracruz, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán), contienen la secuencia completa de una copia del gen *msp1a* en el contig 1 (Figura 12). Por lo tanto, los genomas finales de las siete cepas mexicanas de *A. marginale* no muestran una superinfección, es decir, cada genoma final contiene la información genómica de una cepa mexicana de *A. marginale*.

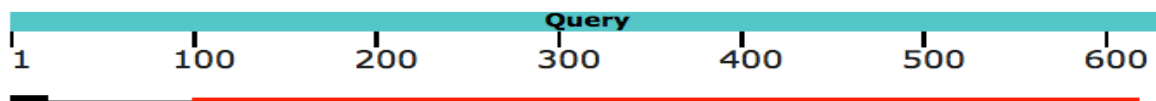


Figura 12. Diagrama representativo que muestra el resultado del alineamiento entre la proteína MSP1a (línea azul) de *A. marginale* St. Maries (número de acceso GenBank AAV86554.1) contra el contig 1 (líneas negra y roja) de los genomas finales de las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México), MEX-17-017-01 (Puente de Ixtla, Morelos), MEX-30-184-02 (Tlapacoyan, Veracruz), MEX-30-193-01 (Veracruz, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) de *A. marginale*, al utilizar el programa tBlastn. Los números que se ubican en la línea azul muestran la longitud de la proteína MSP1a en aminoácidos. Las líneas de color negro y rojo muestran la región variable y conservada de MSP1a, respectivamente. La línea delgada de color negro muestra la región donde se ubica la estructura de repeticiones de MSP1a.

Por otra parte, los genomas finales de las cepas MEX-07-068-01 (Pichucalco, Chiapas) y MEX-28-037-02 (Soto la Marina, Tamaulipas) contienen secuencias parciales de múltiples copias del gen *mSP1a* en diferentes contigs (Figuras 13 y 14, respectivamente). Por lo tanto, los genomas finales de las dos cepas mexicanas de *A. marginale* muestran una superinfección, es decir, cada genoma final contiene la información genómica de dos o más cepas mexicanas diferentes de *A. marginale*. Cabe mencionar que las dos cepas mexicanas de *A. marginale* fueron obtenidas de dos estados que se ubican en regiones tropicales con una prevalencia de la anaplasmosis bovina mayor del 50% (Figuras 1 y 2). Estos resultados apoyan el trabajo reportado por Castañeda-Ortiz y colaboradores en el año 2015 [25], en el cual muestran que la incidencia de la superinfección de *A. marginale* es considerablemente mayor en las regiones tropicales que tienen un elevado porcentaje de prevalencia de la anaplasmosis bovina, en comparación con las regiones templadas o tropicales que tienen un bajo porcentaje de prevalencia de la enfermedad. Esto se atribuye a que existe una mayor presión selectiva por las cepas antigénicamente distintas, y en consecuencia, cepas con diferentes genotipos de MSP1a.



Figura 13. Diagrama representativo que muestra el resultado del alineamiento entre la proteína MSP1a (línea azul) de *A. marginale* St. Maries (número de acceso GenBank AAV86554.1) contra los contigs (líneas rojas y negra) del genoma final de *A. marginale* cepa MEX-07-068-01 (Pichucalco, Chiapas), al utilizar el programa tBlastn. Los números que se ubican en la línea azul muestran la longitud de la proteína MSP1a en aminoácidos. Las líneas de color rojo muestran fragmentos de la región conservada de MSP1a.

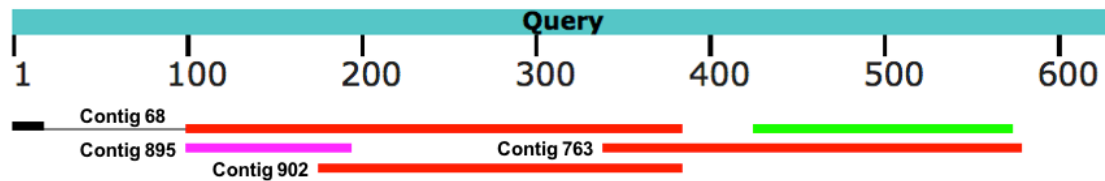


Figura 14. Diagrama representativo que muestra el resultado del alineamiento entre la proteína MSP1a (línea azul) de *A. marginale* St. Maries (número de acceso GenBank AAV86554.1) contra los contigs (líneas negras, rojas, rosa y verde) del genoma final de *A. marginale* cepa MEX-28-037-02 (Soto la Marina, Tamaulipas), al utilizar el programa tBlastn. Los números que se ubican en la línea azul muestran la longitud de la proteína MSP1a en aminoácidos. Las líneas de color negro muestran la región variable de MSP1a. La línea delgada de color negro muestra la región donde se ubica la estructura de repetidas de MSP1a. Las líneas de color rojo, rosa y verde muestran fragmentos de la región conservada de MSP1a.

La proteína MSP1a de *A. marginale* contiene una estructura de repetidas en la región variable en el extremo amino terminal. La estructura de repetidas contiene entre una o hasta 11 secuencias similares que son llamadas repetidas, las cuales tienen una longitud entre 23 y 31 aminoácidos cada una [29]. Hasta la fecha, más de 235 secuencias de repetidas han sido reportadas a nivel mundial [56]. La estructura de repetidas de MSP1a tiene la dificultad de ser reportada de manera correcta porque es difícil obtener la secuencia exacta de cada repetida. Y además, el nombre de cada repetida es asignado de manera manual con una tasa de error, por ejemplo, el mismo nombre es asignado a repetidas que tienen diferentes secuencias, o diferentes nombres son asignados a repetidas que tienen la misma secuencia [56]. En el presente trabajo, las estructuras de repetidas de MSP1a de las nueve cepas mexicanas de *A. marginale*, fueron identificadas, verificadas y analizadas para obtener la posible distribución geográfica a nivel mundial.

La proteína MSP1a de la cepa MEX-01-001-01 (Aguascalientes, Aguascalientes) tiene una estructura de cinco repetidas: 4 9 10 11 9; la cual es

idéntica a la estructura reportada en la literatura (Tabla 9). Además, la repetida 9 solamente se ha reportado en 15 estructuras de repetidas de cepas mexicanas. Por lo tanto, la cepa MEX-01-001-01 (Aguascalientes, Aguascalientes) está altamente relacionada con las cepas mexicanas de *A. marginale* en la región de Norteamérica.

Tabla 9. Estructura de repetidas de MSP1a de las nueve cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Contig	Estructura de repetidas reportada en la literatura*	Estructura de repetidas obtenida con RepeatAnalyzer
MEX-01-001-01	Aguascalientes, Aguascalientes	1	4 9 10 11 9	4 9 10 11 9
MEX-07-068-01	Pichucalco, Chiapas	97	α β β Γ β Γ	E
MEX-07-068-01	Pichucalco, Chiapas	268	α β β Γ β Γ	NA**
MEX-14-010-01	Atitalaquia, Hidalgo	1	τ 57 13 18	τ 22-2 13 18
MEX-15-099-01	Texcoco, Edo de México	1	α β β Γ	α β Γ
MEX-17-017-01	Puente de Ixtla, Morelos	1	12 13 14	12 14
MEX-28-037-02	Soto la Marina, Tamaulipas	68	28 29 74 29 M F	F M
MEX-28-037-02	Soto la Marina, Tamaulipas	895	28 29 74 29 M F	10 11 9
MEX-30-184-02	Tlapacoyan, Veracruz	1	73 β β β Γ	T C
MEX-30-193-01	Veracruz, Veracruz	1	α β β Γ	α β β Γ
MEX-31-096-01	Tizimín, Yucatán	1	T C B B C B π	T C

* Datos obtenidos de Jiménez-Ocampo R., *et al.* 2012 (Tabla 4) [49].

** NA: No disponible (siglas del inglés *Not Available*)

El genoma final de la cepa MEX-07-068-01 (Pichucalco, Chiapas) muestra una superinfección de diferentes cepas mexicanas de *A. marginale* (Figura 13). Los contigs 97 y 268 contienen las secuencias parciales de la región variable del gen *msp1a* (Figura 13). El contig 97 contiene la estructura parcial de una repetida: E; la cual no coincide con la estructura reportada en la literatura (Tabla 9). Por otra parte, el contig 268 no contiene una estructura de repetidas (Tabla 9). Por lo tanto, estos resultados confirmaron que el genoma final de la cepa MEX-07-068-01 (Pichucalco,

Chiapas) muestra una superinfección de diferentes cepas mexicanas de *A. marginale*. En consecuencia, el genoma final de la cepa MEX-07-068-01 (Pichucalco, Chiapas) fue excluido de los análisis genómicos posteriores.

La proteína MSP1a de la cepa MEX-14-010-01 (Atitalaquia, Hidalgo) tiene una estructura de cuatro repetidas: τ 22-2 13 18; la cual es similar a la estructura reportada en la literatura (Tabla 9). La segunda posición de la estructura de repetidas muestra que la repetida 57 cambió por la repetida 22-2, probablemente debido a las correcciones en la nomenclatura de las repetidas que se implementaron en el programa RepeatAnalyzer [56]. Además, la estructura de repetidas (τ 22-2 13 18) solamente se ha reportado en cuatro cepas argentinas. Por lo tanto, la cepa MEX-14-010-01 (Atitalaquia, Hidalgo) está altamente relacionada con las cepas argentinas de *A. marginale* en la región de Sudamérica.

La proteína MSP1a de la cepa MEX-15-099-01 (Texcoco, Estado de México) tiene una estructura de tres repetidas: α β Γ ; la cual es similar a la estructura reportada en la literatura (Tabla 9). La estructura de repetidas muestra que una repetida β fue eliminada (Tabla 9). La Figura 15A muestra que el ensamblaje de la secuencia completa del gen *msp1a* tiene una elevada calidad y la estructura de repetidas es correcta. Además, Argentina, Brasil, China, México y Venezuela reportaron 41 estructuras de repetidas que inician con la repetida α y contienen a las repetidas β y/o Γ . Por lo tanto, la cepa MEX-15-099-01 (Texcoco, Estado de

México) está altamente relacionada con las cepas argentinas, brasileñas y venezolanas de *A. marginale* en la región de Sudamérica.

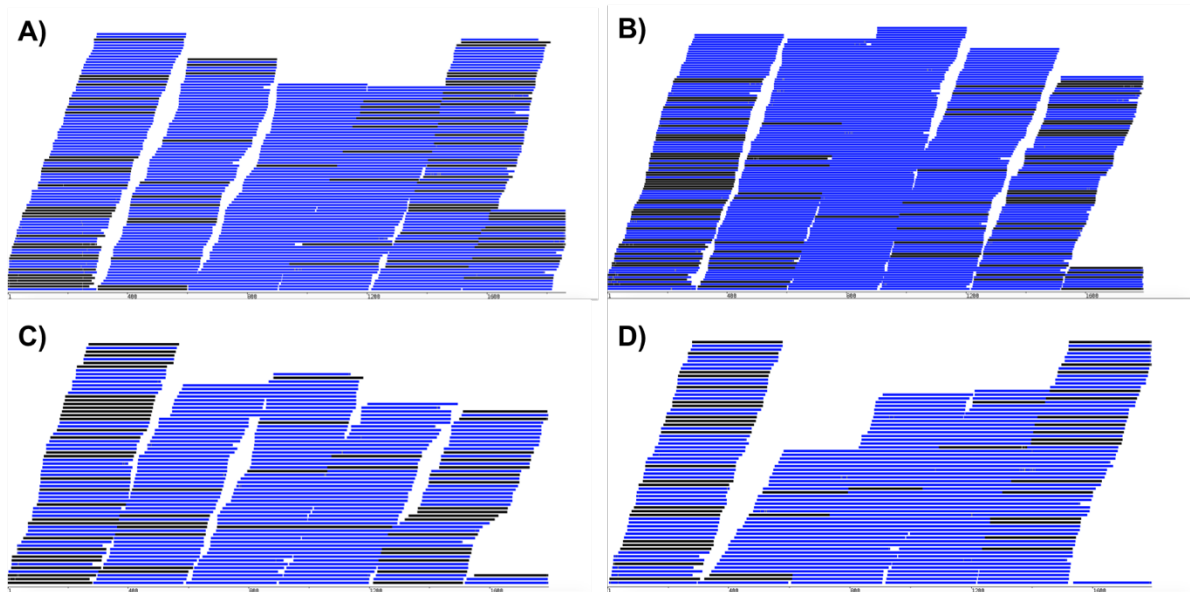


Figura 15. Diagramas representativos que muestran el resultado del alineamiento entre las lecturas crudas de 300 bases de longitud (líneas azules y negras) contra el gen *msp1a* de ~1.8 Kb de longitud (líneas blancas en la parte inferior) de las cepas MEX-15-099-01 de Texcoco, Estado de México (A); MEX-17-017-01 de Puente de Ixtla, Morelos (B); MEX-30-184-02 de Tlapacoyan, Veracruz (C); y MEX-31-096-01 de Tizimín, Yucatán (D) de *A. marginale*, al utilizar el programa BamView. Las lecturas crudas pareadas se muestran en color azul y las lecturas crudas no pareadas se muestran en color negro. La región variable del gen *msp1a* se localiza en el extremo izquierdo de cada figura.

La proteína MSP1a de la cepa MEX-17-017-01 (Puente de Ixtla, Morelos) tiene una estructura de dos repetidas: 12 14; la cual es similar a la estructura reportada en la literatura (Tabla 9). La estructura de repetidas no muestra la presencia de la repetida 13 (Tabla 9). La Figura 15B muestra que el ensamblaje de la secuencia completa del gen *msp1a* tiene una elevada calidad y la estructura de repetidas es correcta. Además, la repetida 12 solamente se ha reportado en cinco estructuras de repetidas de cepas mexicanas. Por otra parte, China, Filipinas y México reportaron 19 estructuras de repetidas que contienen a la repetida 14. Por lo tanto, la cepa

MEX-17-017-01 (Puente de Ixtla, Morelos) está altamente relacionada con cepas chinas, filipinas y mexicanas de *A. marginale* en las regiones de Asia y Norteamérica.

El genoma final de la cepa MEX-28-037-02 (Soto la Marina, Tamaulipas) muestra una superinfección de diferentes cepas mexicanas de *A. marginale* (Figura 14). Los contigs 68 y 895 contienen las secuencias parciales de la región variable del gen *msp1a* (Figura 14). El contig 68 contiene la estructura parcial de dos repetidas: F M; la cual no coincide con la estructura completa que fue reportada en la literatura (Tabla 9). Por otra parte, el contig 895 contiene la estructura parcial de tres repetidas: 10 11 9; la cual no coincide con la estructura reportada en la literatura (Tabla 9). Por lo tanto, estos resultados confirmaron que el genoma final de la cepa MEX-28-037-02 (Soto la Marina, Tamaulipas) muestra una superinfección de diferentes cepas mexicanas de *A. marginale*. En consecuencia, el genoma final de la cepa MEX-28-037-02 (Soto la Marina, Tamaulipas) fue excluido de los análisis genómicos posteriores.

La proteína MSP1a de la cepa MEX-30-184-02 (Tlapacoyan, Veracruz) tiene una estructura de dos repetidas: T C; la cual no coincide con la estructura reportada en la literatura (Tabla 9). La Figura 15C muestra que el ensamblaje de la secuencia completa del gen *msp1a* tiene una elevada calidad y la estructura de repetidas es correcta. Estos datos sugieren dos posibles opciones: (i) la cepa mexicana de *A. marginale* fue etiquetada de manera incorrecta y corresponde a otra cepa diferente, por ejemplo, la cepa MEX-30-184-03 de Tlapacoyan, Veracruz

(Tabla 4); y (ii) la estructura de repetidas que fue reportada en la literatura contiene una elevada tasa de error, probablemente debido a la complejidad de ensamblar las secuencias consenso de manera manual a partir de los cromatogramas [49]. Además, Estados Unidos, México y Venezuela reportaron 16 estructuras de repetidas que inician con la repetida T y contienen a la repetida C. Por lo tanto, la cepa MEX-30-184-02 (Tlapacoyan, Veracruz) está altamente relacionada con las cepas mexicanas y americanas de *A. marginale* en la región de Norteamérica.

La proteína MSP1a de la cepa MEX-30-193-01 (Veracruz, Veracruz) tiene una estructura de cuatro repetidas: α β β Γ ; la cual es idéntica a la estructura reportada en la literatura (Tabla 9). Además, la estructura de repetidas α β β Γ , se ha reportada en 13 cepas brasileñas y mexicanas. Por lo tanto, la cepa MEX-30-193-01 (Veracruz, Veracruz) está altamente relacionada con las cepas brasileñas de *A. marginale* en la región de Sudamérica.

La proteína MSP1a de la cepa MEX-31-096-01 (Tizimín, Yucatán) tiene una estructura de dos repetidas: T C; la cual no coincide con la estructura completa que fue reportada en la literatura (Tabla 9). La Figura 15D muestra que el ensamblaje de la secuencia completa del gen *msp1a* tiene una elevada calidad y la estructura de repetidas es correcta. Estos datos sugieren que la estructura de repetidas que fue reportada en la literatura contiene una elevada tasa de error, probablemente debido a la complejidad de ensamblar las secuencias consenso de manera manual a partir de los cromatogramas [49]. Además, la estructura de repetidas es idéntica a la

estructura reportada para la cepa MEX-30-184-02 de Tlapacoyan, Veracruz (Tabla 9). Por lo tanto, la cepa MEX-31-096-01 está altamente relacionada con las cepas mexicanas y americanas de *A. marginale* en la región de Norteamérica.

En México y a nivel mundial, las cepas de *A. marginale* se han identificado principalmente por el análisis de la estructura de repetidas de la región variable de MSP1a [34,44]. Sin embargo, las secuencias de los genes *mSP1a* se han obtenido mediante protocolos no automatizados que incrementan la tasa de error al ensamblar: (i) las regiones variables con los patrones de repeticiones; y (ii) secuencias consenso a partir de dos o más genes *mSP1a* que tienen un peso molecular similar. Por otra parte, la secuenciación masiva del gDNA y el ensamblaje de los genomas de *A. marginale*, proporcionan las secuencias completas del gen *mSP1a* con una mayor calidad de ensamblaje de la estructura de repetidas, mientras que al obtener secuencias parciales del gen *mSP1a* se sugiere que existe una superinfección de *A. marginale*. En este trabajo, las estructuras de repetidas de MSP1a de cinco cepas mexicanas de *A. marginale* no coincidieron con las estructuras reportadas en la literatura (Tabla 9). Y además, dos cepas mexicanas mostraron una superinfección de *A. marginale*. Estos datos sugieren que probablemente un porcentaje de estructuras de repetidas de MSP1a que están reportadas en la literatura y/o en bases de datos, contienen secuencias con errores que afectan la identificación y el análisis de la posible distribución geográfica de las cepas de *A. marginale* a nivel mundial. En consecuencia, el presente trabajo promueve el ensamblaje de genomas de *A. marginale* para realizar análisis de la estructura de repetidas de MSP1a con mayor precisión.

A. marginale es un patógeno que infecta a los eritrocitos del bovino, y además, se estima que también puede infectar a las células endoteliales [104,105]. Por otra parte, *A. marginale* se transmite de manera horizontal por garrapatas de la familia *Ixodidae*, que incluye algunas especies de los géneros *Rhipicephalus* y *Dermacentor*. La garrapata *Rhipicephalus (Boophilus) microplus* es considerada el vector biológico más importante en las regiones tropicales y subtropicales a nivel mundial [106]. La región variable de MSP1a se encuentra expuesta de manera extracelular, y tiene una función de ligando o adhesina que reconoce a un receptor no identificado en los eritrocitos del bovino y algunas células de las garrapatas [107]. Por lo tanto, la proteína MSP1a tiene una función importante en la inmunidad del bovino y en la invasión y transmisibilidad por vectores biológicos [49]. Además, la proteína MSP1a es probable que evolucione con mayor rapidez en comparación con otras proteínas, porque está sujeta a la presión selectiva que ejerce el sistema inmune del bovino [61].

Hasta la fecha, los trabajos sobre la biogeografía y la evolución de *A. marginale* son escasos. En el año 2009, Estrada-Peña y colaboradores reportaron que la filogeografía de las secuencias de MSP1a de *A. marginale* está asociada con cuatro grupos de regiones ecológicas mundiales (ecorregiones) (Tabla 10 y Figura 16) [61]. Los resultados mostraron que las secuencias de microsatélites, y las primeras (RI) y últimas (RL) repetidas de la estructura de MSP1a, están asociadas con grupos de ecorregiones que tienen características ambientales específicas que afectan el rendimiento de las garrapatas como vectores de *A. marginale* (Tablas 11 y 12, respectivamente).

Tabla 10. Características generales de las cuatro ecorregiones reportadas por Estrada-Peña A., *et al.* 2009 [61].

	Grupos de ecorregiones			
	1	2	3	4
Principales regiones geográficas en el mundo	Grandes áreas de África central y Sudamérica central, principalmente Argentina y el sur de Brasil	Vastas áreas del corredor mesoamericano, el norte de Sudamérica y un pequeño territorio del este de Sudáfrica	Sudáfrica central y regiones dispersas del sur de Estados Unidos y México	Grandes áreas de Estados Unidos
Valores de NDVI*	Valores medios a altos con una disminución estacional entre junio y septiembre	Valores altos a lo largo del año sin variación estacional	Valores más bajos y pocos cambios a lo largo del año	Valores muy particulares, muy bajos entre noviembre y marzo y aumentan para alcanzar niveles máximos alrededor de julio
Valores de temperatura	La temperatura registrada más elevada	Similar a la ecorregión 1	Temperaturas muy bajas	La temperatura más fría de todas las ecorregiones
Valores de precipitación	Alrededor de 1,000 mm/año	Alrededor de 1,500 mm/año	Precipitación mínima	Alrededor de 800 mm/año

* Índice mensual de la diferencia de la vegetación normalizada (NDVI del inglés *Normalized Difference Vegetation Index*). NDVI es una variable que refleja el estrés de la vegetación. Datos obtenidos para el periodo 1986-2006 [61].

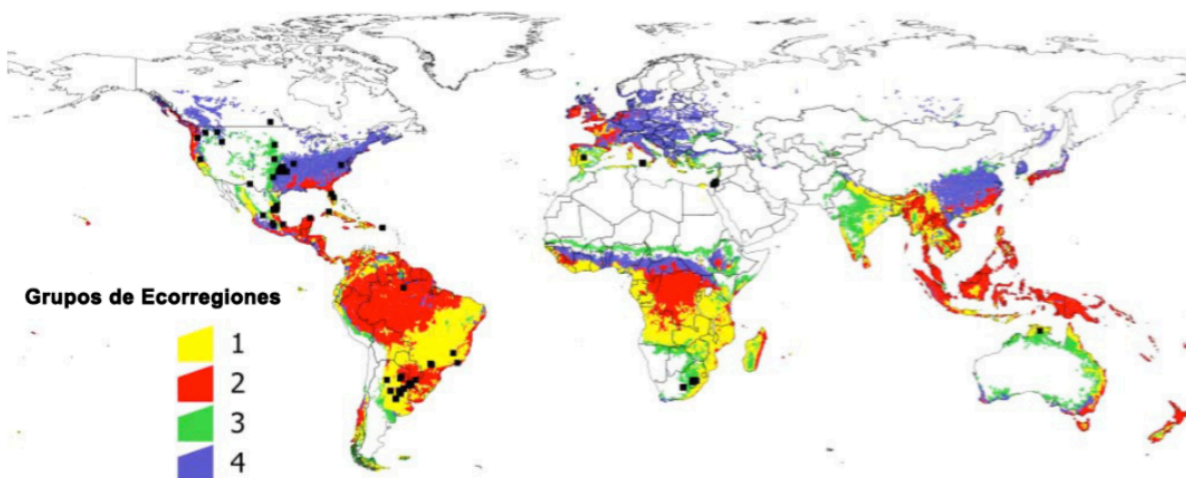


Figura 16. Distribución de las cuatro ecorregiones a nivel mundial. Las características de la vegetación se calcularon mediante una clasificación de los valores del índice mensual de la diferencia de la vegetación normalizada (NDVI del inglés *Normalized Difference Vegetation Index*) entre 1986-2006 a una resolución de 0.1°. Las islas pequeñas fueron eliminadas para mejorar la presentación. Los cuatro grupos de ecorregiones se numeraron de manera arbitraria: ecorregión 1 en color amarillo, ecorregión 2 en color rojo, ecorregión 3 en color verde y ecorregión 4 en color azul (Editada de Estrada-Peña A., *et al.* 2009 [61]).

Tabla 11. Estructura de los microsatélites de MSP1a de *A. marginale* y distribución en los cuatro grupos de ecorregiones (Editada de Estrada-Peña A., *et al.* 2009 [61]).

Genotipo	Número de cepas	Estructura de los microsatélites			Frecuencia del genotipo por grupo de ecorregión			
		m	n	Distancia SD-ATG (nucleótidos)	1	2	3	4
A	2	1	7	19	0.00	1.00	0.00	0.00
B	5	1	9	23	1.00	0.00	0.00	0.00
C	7	2	5	19	0.57	0.14	0.29	0.00
D	3	2	6	21	0.33	0.67	0.00	0.00
E	12	2	7	23	0.75	0.00	0.25	0.00
F	3	3	4	21	0.00	0.00	1.00	0.00
G	78	3	5	23	0.15	0.14	0.56	0.14
H	3	3	6	25	0.00	0.67	0.33	0.00
I	NI*	4	6	29	---	---	---	---

Las secuencias de microsatélites de MSP1a fueron analizadas en 115 cepas de *A. marginale*.

* NI: No incluida en el estudio porque la cepa de *A. marginale* no estaba georreferenciada de manera adecuada.

Tabla 12. Primeras (RI) y últimas (RL) repetidas de la estructura de MSP1a de *A. marginale* que están asociadas a las cuatro ecorregiones (Editada de Estrada-Peña A., *et al.* 2009 [61]).

Repetidas de MSP1a	Repetidas únicas para una ecorregión	Repetidas presentes en más de una ecorregión
RI / Ecoregion 1	4, 8, 16, 56, 60, 64, 67, Γ , π , τ	A, B, D, T, 13, 23, α
RI / Ecoregion 2	28, 48, 53, E, F, ε	A, B, L, T, 13, 23, α
RI / Ecoregion 3	1, 3, 5, 6, 27, 33, 34, 39, M, O, Q, U	A, D
RI / Ecoregion 4	I, J, K	A, B, L, α
RL / Ecoregion 1	8, 9, 12, 15, 59, 61, 66	B, C, M, 18, 27, Γ
RL / Ecoregion 2	10, 31, 52, π , β	F, H, M, 27, Γ
RL / Ecoregion 3	3, 7, 35, 37, 38, 44, E, N, P, Q, U, ρ	B, F, H, 18, Γ
RL / Ecoregion 4	Ninguna	B, C, H

Las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México) y MEX-30-193-01 (Veracruz, Veracruz) tienen una estructura del microsatélite: m=2, n=7 y la distancia SD-ATG=23. Por lo tanto, las cuatro cepas mexicanas pertenecen al genotipo E, el cual muestra una frecuencia de 0.75 y 0.25 para las ecorregiones 1 y

3, respectivamente (Tabla 13). Las estructuras de repetidas de MSP1a de las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes) y MEX-14-010-01 (Atitalaquia, Hidalgo) inician con las repetidas 4 y τ , respectivamente, las cuales solamente se han reportado en cepas que pertenecen a la ecorregión 1 (Tabla 12). Además, las estructuras de repetidas de MSP1a de las cepas MEX-15-099-01 (Texcoco, Estado de México) y MEX-30-193-01 (Veracruz, Veracruz) inician con la repetida α , la cual se ha reportado en cepas que pertenecen a la ecorregión 1, pero no se ha reportado en cepas que pertenecen a la ecorregión 3 (Tabla 12). Por otra parte, la estructura de repetidas de MSP1a de la cepa MEX-01-001-01 (Aguascalientes, Aguascalientes) termina con la repetida 9 (Tabla 9), la cual solamente se ha reportado en cepas que pertenecen a la ecorregión 1 (Tabla 12). Además, las estructuras de repetidas de MSP1a que terminan con las repetidas 18 (MEX-14-010-01 de Atitalaquia, Hidalgo) y Γ (MEX-15-099-01 de Texcoco, Estado de México y MEX-30-193-01 de Veracruz, Veracruz) se han reportado en cepas que pertenecen a las ecorregiones 1 y 3 (Tablas 9 y 12). Estos datos sugieren que la cepa MEX-01-001-1 (Aguascalientes, Aguascalientes) está altamente relacionada con la ecorregión 1, mientras que las cepas MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México) y MEX-30-193-01 (Veracruz, Veracruz) están relacionadas con la ecorregión 1 y posiblemente con la ecorregión 3.

Tabla 13. Características de los microsátélites de MSP1a de las siete cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Estructura de los microsátélites			Genotipo*	Frecuencia del genotipo por grupo de ecorregión*			
		m	n	Distancia SD-ATG (nucleótidos)		1	2	3	4
MEX-01-001-01	Aguascalientes, Aguascalientes	2	7	23	E	0.75	0.00	0.25	0.00
MEX-14-010-01	Atitalaquia, Hidalgo	2	7	23	E	0.75	0.00	0.25	0.00
MEX-15-099-01	Texcoco, Edo de México	2	7	23	E	0.75	0.00	0.25	0.00
MEX-17-017-01	Puente de Ixtla, Morelos	3	5	23	G	0.15	0.14	0.56	0.14
MEX-30-184-02	Tlapacoyan, Veracruz	3	5	23	G	0.15	0.14	0.56	0.14
MEX-30-193-01	Veracruz, Veracruz	2	7	23	E	0.75	0.00	0.25	0.00
MEX-31-096-01	Tizimín, Yucatán	3	5	23	G	0.15	0.14	0.56	0.14

* Datos obtenidos de Estrada-Peña A., *et al.* 2009 [61].

Las cepas MEX-17-017-01 (Puente de Ixtla, Morelos), MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) tienen una estructura del microsátélite: m=3, n=5 y la distancia SD-ATG=23. Por lo tanto, las tres cepas mexicanas pertenecen al genotipo G, el cual muestra una frecuencia de 0.15, 0.14, 0.56 y 0.14 para las cuatro ecorregiones, respectivamente (Tabla 13). La estructura de repetidas de MSP1a de la cepa MEX-17-017-01 (Puente de Ixtla, Morelos) inicia con la repetida 12 (Tabla 9), la cual no se ha reportado en cepas de las cuatro ecorregiones (Tabla 12). Además, las estructuras de repetidas de MSP1a de las cepas MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) inician con la repetida T (Tabla 9), la cual se ha reportado en cepas que pertenecen a las ecorregiones 1 y 2 (Tabla 12). Por otra parte, la estructura de repetidas de MSP1a de la cepa MEX-17-017-01 (Puente de Ixtla, Morelos) termina con la repetida 14 (Tabla 9), la cual no se ha reportado en cepas de las cuatro ecorregiones (Tabla 12). Además, las estructuras de repetidas de MSP1a de las cepas MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán)

terminan con la repetida C (Tabla 9), la cual se ha reportado en cepas que pertenecen a las ecorregiones 1 y 4 (Tabla 12). Estos datos sugieren que la cepa MEX-17-017-01 (Puente de Ixtla, Morelos) no está relacionada con alguna de las cuatro ecorregiones reportadas, mientras que las cepas MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) están relacionadas con la ecorregión 1 y posiblemente con otras ecorregiones.

Los resultados anteriores sugieren que las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México), MEX-30-184-02 (Tlapacoyan, Veracruz), MEX-30-193-01 (Veracruz, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán), probablemente se transmiten por vectores biológicos (garrapatas) que prefieren los ambientes con valores medios a altos de NVDI, es decir, una vegetación de arbustos y pastizales a bosques templados y tropicales, temperaturas elevadas y una precipitación anual de alrededor de 1,000 mm (Tabla 10 y Figura 16). Cabe mencionar que las cuatro ecorregiones se clasificaron utilizando los datos obtenidos durante el periodo de 1986-2006. Los factores ambientales son dinámicos y probablemente sufrieron modificaciones hasta la fecha.

Los genomas finales de las siete cepas mexicanas de *A. marginale* muestran una elevada cobertura final de ensamblaje entre 19 y 154X (Tabla 14), es decir, la secuenciación masiva del gDNA de cada cepa mexicana generó la información genómica suficiente para ensamblar los genomas con una amplitud mínima de 19

veces, lo cual disminuye considerablemente la tasa de error en el ensamblaje de los genomas.

Tabla 14. Cobertura final de ensamblaje de los genomas finales de las siete cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Cobertura final de ensamblaje del genoma	Número de acceso GenBank
MEX-01-001-01	Aguascalientes, Aguascalientes	~36X	QLIV00000000.1
MEX-14-010-01	Atitalaquia, Hidalgo	~22X	VTSO00000000.1
MEX-15-099-01	Texcoco, Estado de México	~30X	VTWW00000000.1
MEX-17-017-01	Puente de Ixtla, Morelos	~58X	VTCX00000000.1
MEX-30-184-03	Tlapacoyan, Veracruz	~27X	VTCY00000000.1
MEX-30-193-01	Veracruz, Veracruz	~154X	VTCZ00000000.1
MEX-31-096-01	Tizimín, Yucatán	~19X	VTWV00000000.1

Por último, los genomas finales de las siete cepas mexicanas de *A. marginale* fueron depositados en la base de datos GenBank (Apéndices 1-7). Los números de acceso GenBank se muestran en la Tabla 14. Además, la información de los genomas finales fue publicada en la revista científica *Microbiology Resource Announcements* de la American Society for Microbiology [108,109] (Apéndices 8 y 9).

6.5 Anotación de los genomas

La anotación automática de los siete genomas finales de las cepas mexicanas de *A. marginale* se realizó al utilizar el servidor RAST (Tabla 15). Como resultado, se obtuvieron más de 1,100 CDS por cada cepa. Además, para cada cepa se anotaron ~400 y ~700 CDS como proteínas hipotéticas y con una posible función

biológica, respectivamente. Estos datos muestran que el número de CDS anotadas es similar en los siete genomas finales de las cepas mexicanas de *A. marginale*.

Tabla 15. Características de la anotación automática de los genomas finales de las siete cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Genes*	CDS*	rRNAs**	tRNAs***
MEX-01-001-01	Aguascalientes, Aguascalientes	1,218	1,178	3	37
MEX-14-010-01	Atitalaquia, Hidalgo	1,190	1,150	3	37
MEX-15-099-01	Texcoco, Estado de México	1,185	1,145	3	37
MEX-17-017-01	Puente de Ixtla, Morelos	1,203	1,163	3	37
MEX-30-184-03	Tlapacoyan, Veracruz	1,205	1,165	3	37
MEX-30-193-01	Veracruz, Veracruz	1,178	1,138	3	37
MEX-31-096-01	Tizimín, Yucatán	1,204	1,164	3	37

* Datos obtenidos con el servidor RAST.

** Datos obtenidos con el servidor RNAmmer.

*** Datos obtenidos con el servidor ARAGORN.

Posteriormente, se realizó la anotación automática de los genes rRNAs y tRNAs de los genomas finales de las siete cepas mexicanas de *A. marginale* (Tabla 15). Como resultado, se detectaron las copias únicas de los tres genes rRNAs (16, 23 y 5S) por cada cepa. También, se detectaron las secuencias de 37 genes tRNAs en los siete genomas finales. Con base en el dato anterior, se propone que los 37 genes tRNAs están conservados en los genomas de *A. marginale*. Cabe mencionar que se detectaron los 37 genes tRNAs en diez genomas de *A. marginale* que están depositados en la base de datos GenBank. Sin embargo, en el genoma completo de la cepa Dawn y los cuatro borradores de genomas de las cepas Florida (204 contigs), Okeechobee (403 contigs), South Idaho (358 contigs) y Washington Okanogan (332 contigs), se detectaron 35, 35, 38, 47 y 52 genes tRNAs, respectivamente. El genoma completo de la cepa Dawn tiene una longitud total de 1,196,760 pb pero contiene 86,632 nucleótidos N (cualquiera de los cuatro

nucleótidos). Por lo tanto, la falta de información genómica pudo evitar la detección de dos genes tRNAs. Por otra parte, el borrador de genoma de la cepa Florida contiene 35 genes tRNAs. Este dato es contradictorio con la detección de 37 genes tRNAs en el genoma completo de *A. marginale* Florida y en el borrador de genoma de *A. marginale* Florida Relapse, porque los tres genomas pertenecen al mismo organismo. Por último, los tres borradores de genomas de las cepas americanas Okeechobee, South Idaho y Washington Okanogan, tienen una cantidad excesiva de contigs (más de 200 contigs por cada cepa), lo cual probablemente ocasionó una duplicación de los genes tRNAs. Los resultados anteriores muestran que los tres genes rRNAs y los 37 genes tRNAs están conservados en los genomas de *A. marginale*. Además, se sugiere que el genoma completo de la cepa Dawn y los cuatro borradores de genomas de las cepas Florida, Okeechobee, South Idaho y Washington Okanogan no contienen información genómica confiable.

El siguiente paso fue agrupar en categorías de familias (también llamadas subsistemas) a las proteínas anotadas de los siete genomas finales de las cepas mexicanas de *A. marginale* al utilizar el servidor RAST (Tabla 16). Como resultado, las proteínas no se agruparon en los procesos de fotosíntesis, motilidad y quimiotaxis, metabolismo secundario, y metabolismo de compuestos aromáticos. Además, entre 17 y 18 proteínas se agruparon en los procesos de virulencia, enfermedad y defensa. En este subsistema no se identificaron posibles proteínas de adhesinas, toxinas, superantígenos, bacteriocinas o péptidos antimicrobianos que son sintetizados de manera ribosomal. Sin embargo, en este subsistema se identificaron posibles proteínas con función de betalactamasas; resistencia a

cobalto-zinc-cadmio; resistencia a fluoroquinolonas; tolerancia al cobre; y operones de virulencia tipo *Mycobacterium* que están involucrados en la síntesis de proteínas y la transcripción del DNA.

Tabla 16. Categorías de familias (subsistemas) de las proteínas anotadas de las siete cepas mexicanas de *A. marginale* obtenidas con el servidor RAST.

Categoría de familias de proteínas (subsistemas)	MEX-01-001 Ags	MEX-14-010 Atitalaquia	MEX-15-099 Texcoco	MEX-17-017 Puente de Ixtila	MEX-30-184 Tlapacoyan	MEX-30-193 Veracruz	MEX-31-096 Tizimín
Cofactores, vitaminas, grupos prostéticos, pigmentos	120	121	123	125	121	120	123
Pared celular y cápsula	37	37	37	37	37	37	37
Virulencia, enfermedad y defensa	17	17	17	18	18	17	18
Metabolismo de potasio	3	3	3	3	3	3	3
Fotosíntesis	0	0	0	0	0	0	0
Misceláneo	4	4	4	4	4	4	4
Fagos, profagos, elementos transponibles, plásmidos	1	1	1	1	1	1	1
Transporte de membrana	13	13	13	13	14	13	14
Adquisición y metabolismo de hierro	3	3	3	3	3	3	3
Metabolismo de RNA	51	52	50	50	52	50	52
Nucleósidos y nucleótidos	43	42	43	46	42	43	42
Metabolismo de proteínas	185	184	182	188	187	184	188
División celular y ciclo celular	24	24	24	24	26	24	26
Motilidad y quimiotaxis	0	0	0	0	0	0	0
Regulación y señalización celular	1	1	1	1	1	1	1
Metabolismo secundario	0	0	0	0	0	0	0
Metabolismo del DNA	53	50	50	51	45	49	45
Ácidos grasos, lípidos e isoprenoídes	54	54	56	56	54	56	54
Metabolismo del nitrógeno	5	5	5	5	5	5	5
Latencia y esporulación	1	1	1	1	1	1	1
Respiración	63	63	63	63	62	71	63
Respuesta al estrés	26	26	26	26	26	26	26
Metabolismo de compuestos aromáticos	0	0	0	0	0	0	0
Aminoácidos y derivados	51	51	51	51	51	51	51
Metabolismo del azufre	4	4	4	4	4	4	4
Metabolismo de fósforo	9	9	9	9	9	9	9
Carbohidratos	45	41	43	43	45	43	45
Total de proteínas anotadas	813	806	809	822	811	815	815

Entre 13 y 14 proteínas se agruparon en el proceso de transporte de membrana. En este subsistema no se identificaron posibles proteínas de los sistemas de secreción tipo I hasta VIII, o de los transportadores ABC. Sin embargo, en este subsistema se identificaron posibles proteínas de los transportadores de translocación a través de la membrana citoplasmática; transportadores de magnesio; sistemas de transporte Ton y Tol; y transportadores TRAP.

También, 26 proteínas se agruparon en el proceso de la respuesta al estrés. En este subsistema no se identificaron posibles proteínas que estuvieran involucradas con los tipos de estrés osmótico, desecación y ácido. Sin embargo, en este subsistema se identificaron posibles proteínas que están involucradas con los tipos de estrés oxidativo, choque térmico de frío y calor, periplásmico y desintoxicación.

Los resultados anteriores muestran que las proteínas anotadas de los siete genomas finales se agrupan de manera similar en los subsistemas de RAST. Además, existen proteínas que se agruparon en los procesos de virulencia, defensa, transporte de membrana y estrés. Las proteínas de estos subsistemas podrían ser utilizadas para la infección persistente de los eritrocitos del ganado bovino y/o las células de las garrapatas.

Por último, las proteínas anotadas en los siete genomas finales de las cepas mexicanas de *A. marginale*, se agruparon en conjuntos predeterminados de proteínas ortólogas de la base de datos COG (versión 2003-2014) (Tabla 17).

Tabla 17. Conjuntos de proteínas ortólogas de las siete cepas mexicanas de *A. marginale* obtenidos con la base de datos COG.

Conjunto de proteínas ortólogas	01-001 Ags	14-010 Atitalaquia	15-099 Texcoco	17-017 Puente de Ixtla	30-184 Tlapacoyan	30-193 Veracruz	31-096 Tizimín
ALMACENAMIENTO Y PROCESAMIENTO DE LA INFORMACIÓN							
[J] – Traducción, estructura ribosomal y biogénesis	131	131	128	132	133	131	134
[A] – Procesamiento y modificación del RNA	0	0	0	0	0	0	0
[K] – Transcripción	26	26	26	26	26	26	26
[L] – Replicación, recombinación y reparación	51	51	51	50	50	50	50
[B] – Estructura y dinámica de la cromatina	0	0	0	0	0	0	0
PROCESOS CELULARES Y SEÑALIZACIÓN							
[D] – Control del ciclo celular división celular y partición de cromosomas	8	8	8	7	8	7	8
[Y] – Estructura nuclear	0	0	0	0	0	0	0
[V] – Mecanismos de defensa	5	5	5	5	5	5	5
[T] – Mecanismos de transducción de señales	14	14	14	14	14	14	14
[M] – Biogénesis de la pared/membrana/envoltura celular	14	14	14	14	15	14	15
[N] – Motilidad celular	2	2	2	2	2	2	2
[Z] - Citoesqueleto	0	0	0	0	0	0	0
[W] – Estructuras extracelulares	0	0	0	0	0	0	0
[U] – Tráfico intracelular, secreción y transporte vesicular	23	23	23	23	23	23	23
[O] – Modificación postraduccional, recambio proteico, chaperonas	53	53	52	52	53	53	53
[X] – Mobiloma: profagos, transposones	5	5	5	5	5	5	5
METABOLISMO							
[C] – Producción y conversión de energía	58	57	57	57	58	57	58
[G] – Transporte y metabolismo de carbohidratos	18	17	18	18	17	18	19
[E] – Transporte y metabolismo de aminoácidos	30	31	32	32	33	33	34
[F] – Transporte y metabolismo de nucleótidos	38	37	38	38	37	39	37
[H] – Transporte y metabolismo de coenzimas	52	52	52	52	52	51	52
[I] – Transporte y metabolismo de lípidos	24	24	23	23	23	24	23
[P] – Transporte y metabolismo de iones inorgánicos	24	24	23	23	23	24	24
[Q] – Biosíntesis, transporte y catabolismo de metabolitos secundarios	7	7	7	7	7	7	6
NO CARACTERIZADAS							
[R] – Únicamente predicción de la función general	32	32	32	32	33	33	34
[S] – Función desconocida	17	17	17	16	17	16	17

No existe un conjunto específico de virulencia que podría ser utilizado para la infección persistente. Por lo tanto, estos conjuntos muestran información limitada sobre la patogenicidad de las siete cepas mexicanas de *A. marginale*. Por otra parte, dos proteínas se agruparon en el conjunto de motilidad celular (conjunto “N”). El dato anterior es contradictorio con el resultado del subsistema de RAST que muestra que no existen proteínas de motilidad o quimiotaxis en el genoma de las siete cepas mexicanas de *A. marginale* (Tabla 16). Los resultados de la Tabla 17 muestran que las proteínas anotadas en los siete genomas finales se agrupan de manera similar en los conjuntos de la base de datos COG.

Los servidores ResFinder, antiSMASH y CARD no identificaron proteínas que están involucradas con la resistencia a antimicrobianos en los genomas de las siete cepas mexicanas. Por lo tanto, los antibióticos son utilizados para controlar los brotes agudos de la anaplasmosis bovina. Sin embargo, los antibióticos son deficientes para prevenir o eliminar al patógeno cuando ha logrado infectar a los eritrocitos del bovino.

Los servidores VirulenceFinder, PathogenFinder y PAIDB no identificaron proteínas que están involucradas con la patogenicidad o virulencia. Por otra parte, el servidor VFPB (del inglés *Virulence Factors of Pathogenic Bacteria*) identificó dos proteínas que están involucradas con la adherencia/invasión (Asp14 y OmpA), 18 proteínas del Sistema de Secreción Tipo IV (SST4) y dos proteínas efectoras del SST4 (AnkA y Ats-1), en los genomas de las siete cepas mexicanas. Estos resultados sugieren que las cepas mexicanas de *A. marginale* utilizan el SST4 de

manera eficiente para infectar a los eritrocitos del ganado bovino y algunas células de las garrapatas.

6.6 Análisis filogenético

Los árboles filogenéticos con las probabilidades bayesianas posteriores y 1,000 repeticiones bootstrap, muestran la misma topología, y fueron obtenidas a partir de 246 genes concatenados del genoma core de las siete cepas mexicanas, dos cepas australianas, dos cepas brasileñas, nueve cepas americanas y una cepa puertorriqueña de *A. marginale*, y una cepa de *A. centrale* como grupo externo (outgroup), al utilizar el modelo de sustitución de nucleótidos TIM1+I+G. El árbol filogenético con las probabilidades bayesianas posteriores (Figura 17) muestra que *A. marginale* Oklahoma como una cepa ancestral. Además, las cepas americanas y puertorriqueña (líneas y letras rojas), y las cepas mexicanas MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) forman un clado independiente y evolutivamente distante en relación con las cepas australianas (líneas y letras verdes), brasileñas (líneas y letras rosas) y cinco cepas mexicanas (líneas y letras azules).

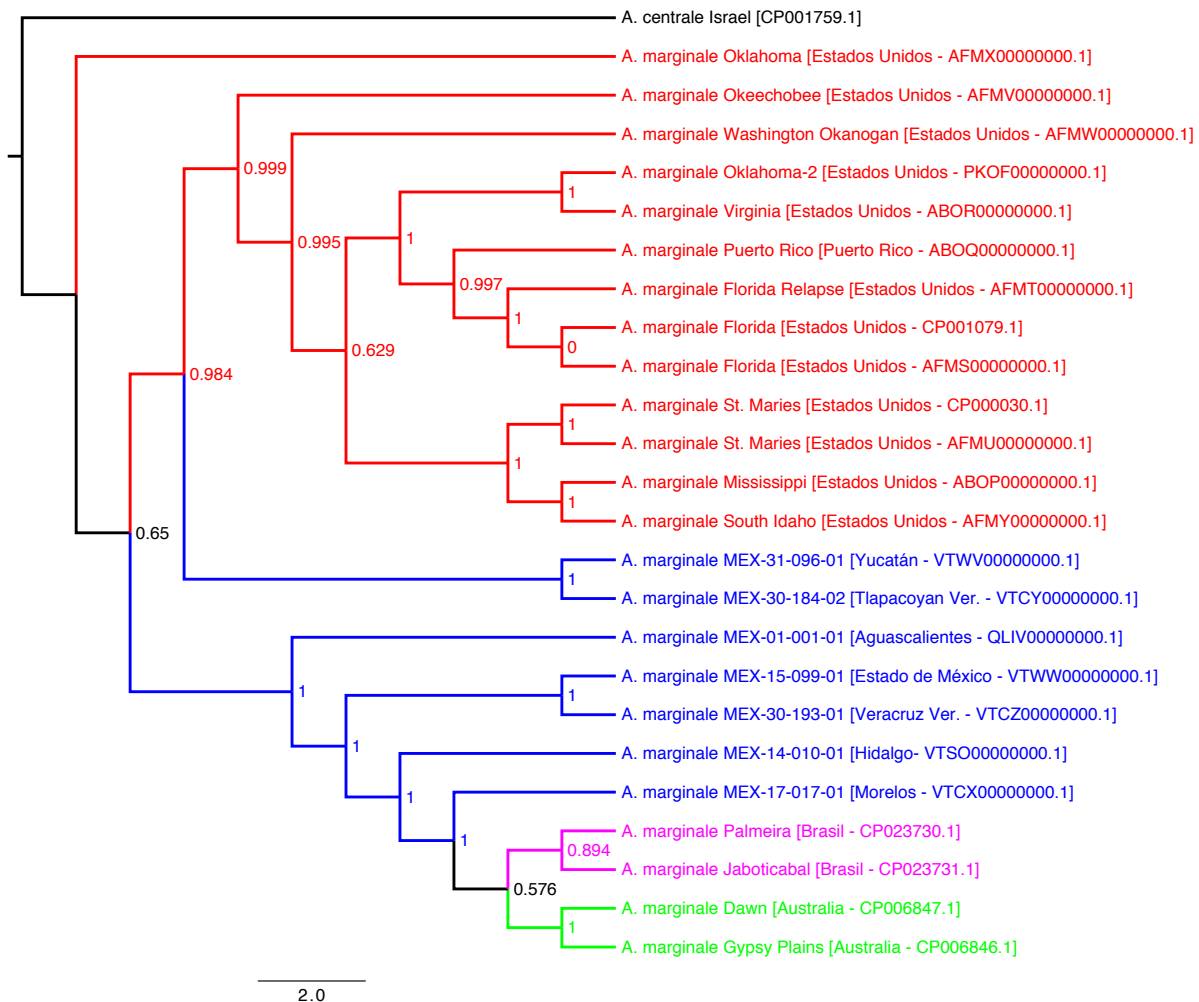


Figura 17. Relaciones filogenéticas de las cepas mexicanas (azul), australianas (verde), brasileñas (rosa), americanas y puertorriqueña (rojo) de *A. marginale*, y la cepa Israel de *A. centrale* (negro). Los valores de las probabilidades bayesianas posteriores se muestran en los nodos.

El árbol filogenético muestra que las cepas mexicanas MEX-30-184-02 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) están estrechamente relacionadas y se encuentran en una posición filogenética cercana a las cepas americanas (Figura 17). Además, tienen la misma estructura de repetidas (T C) de MSP1a (Tabla 9), la cual está altamente relacionada con cepas de México y Estados Unidos (Sección 6.4). Por lo tanto, las dos cepas mexicanas pertenecen a un subclado con una distribución principal en la región de Norteamérica.

La cepa MEX-01-001-01 (Aguascalientes, Aguascalientes) se encuentra en una posición filogenética independiente (Figura 17). La cepa MEX-01-001-01 (Aguascalientes, Aguascalientes) tiene una estructura de repetidas (4 9 10 11 9) de MSP1a (Tabla 9) que está altamente relacionada con cepas mexicanas (Sección 6.4). Por lo tanto, la cepa MEX-01-001-01 (Aguascalientes, Aguascalientes) pertenece a un subclado con una distribución específica en la región de Norteamérica.

Las cepas MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México), MEX-17-017-01 (Puente de Ixtla, Morelos) y MEX-30-193-01 (Veracruz, Veracruz), tienen una posición filogenética cercana a las cepas australianas y brasileñas (Figura 17). La estructura de repetidas de MSP1a de las cepas MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México) y MEX-30-193-01 (Veracruz, Veracruz): τ 22-2 13 18; α β Γ ; α β β Γ (Tabla 9); respectivamente, están altamente relacionadas con cepas de Argentina y Brasil (Sección 6.4). Mientras que la cepa MEX-17-017-01 (Puente de Ixtla, Morelos) tiene una estructura de repetidas (12 14) de MSP1a (Tabla 9) que está altamente relacionada con cepas de Filipinas y China en la región de Asia (Sección 6.4).

Estos resultados sugieren que las cepas mexicanas están relacionadas genéticamente con cepas de otras regiones geográficas a nivel mundial. Sin embargo, la topología del árbol filogenético sería más específica al agregar la información genómica de diferentes cepas que provienen de Canadá,

Centroamérica y el Caribe, Sudamérica (Venezuela y Argentina), Europa, África y Asia.

6.7 Análisis pangenómico

El análisis pangenómico entre las siete cepas mexicanas de *A. marginale* muestra que el genoma core contiene 883 CDS (Figura 18). Además, las cepas MEX-15-099-01 (Texcoco, Estado de México), MEX-30-184-03 (Tlapacoyan, Veracruz), MEX-30-193-01 (Veracruz, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán) tienen menos de 22 CDS únicas (Tabla 18), mientras que las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 y MEX-17-017-01 (Puente de Ixtla, Morelos) tienen más de 45 CDS únicas (Tabla 18). Los datos de la Tabla 18 y el árbol filogenético (Figura 17) muestran que las cepas MEX-30-184-03 (Tlapacoyan, Veracruz) y MEX-31-096-01 (Tizimín, Yucatán), y por otra parte, las cepas MEX-15-099-01 (Texcoco, Estado de México) y MEX-30-193-01 (Veracruz, Veracruz) están estrechamente relacionadas pero no son idénticas.

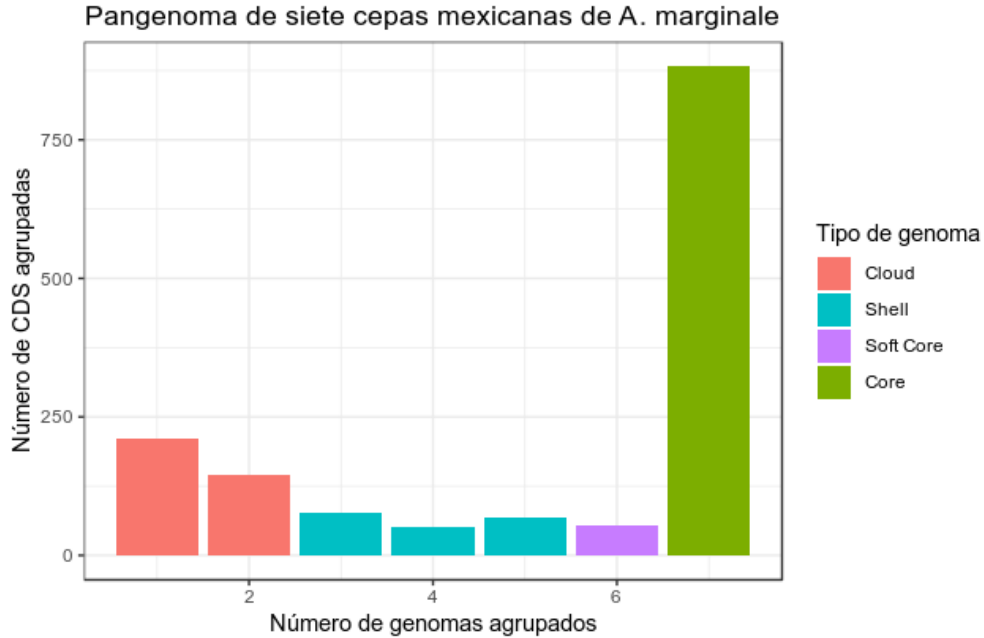


Figura 18. Número de secuencias codificantes (CDS) que comparten las siete cepas mexicanas de *A. marginale*. La gráfica muestra el tipo de genoma: i) Cloud (rojo) que contiene entre 0 y menos de 15% del total de CDS; ii) Shell (azul) que contiene entre 15 y menos de 95% del total de CDS; iii) Soft Core (morado) que contiene entre 95 y menos de 100% del total de CDS; y iv) Core (verde) que contiene el 100% del total de CDS.

Tabla 18. Secuencias codificantes (CDS) únicas en los genomas finales de las siete cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	CDS únicas (siete cepas mexicanas)	CDS únicas (24 genomas de <i>A. marginale</i>)
MEX-01-001-01	Aguascalientes, Aguascalientes	49	14
MEX-14-010-01	Atitalaquia, Hidalgo	48	20
MEX-15-099-01	Texcoco, Estado de México	18	7
MEX-17-017-01	Puente de Ixtla, Morelos	45	17
MEX-30-184-03	Tlapacoyan, Veracruz	20	11
MEX-30-193-01	Veracruz, Veracruz	21	7
MEX-31-096-01	Tizimín, Yucatán	11	2

El análisis pangenómico entre los 24 genomas de las cepas mexicanas, americanas, australianas, brasileñas y puertorriqueña de *A. marginale*, muestra que el genoma core contiene 534 CDS (Figura 19). El análisis pangenómico entre los 24

genomas de *A. marginale* muestra que las CDS únicas y del genoma core disminuyen significativamente con respecto al análisis pangénomico entre los siete genomas de cepas mexicanas. Estos datos muestran que las cepas mexicanas contiene 349 CDS del genoma core que no comparten con las cepas americanas, australianas, brasileñas y puertorriqueña de *A. marginale*. Y además, los datos sugieren que existen CDS que están conservadas en las cepas mexicanas pero no están presentes en las cepas de otros países, a pesar de las relaciones filogenéticas cercanas.

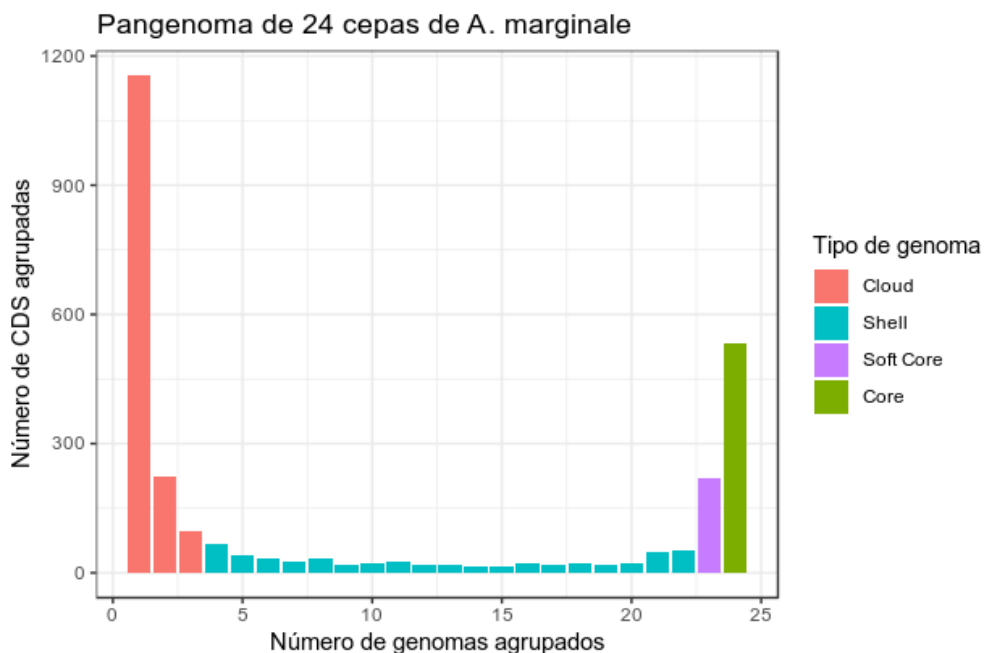


Figura 19. Número de secuencias codificantes (CDS) que comparten 24 genomas de la especie *A. marginale*. La gráfica muestra el tipo de genoma: i) Cloud (rojo) que contiene entre 0 y menos de 15% del total de CDS; ii) Shell (azul) que contiene entre 15 y menos de 95% del total de CDS; iii) Soft Core (morado) que contiene entre 95 y menos de 100% del total de CDS; y iv) Core (verde) que contiene el 100% del total de CDS.

6.8 Genómica comparativa

Los valores de ANIm (basados en MUMmer) de 24 genomas de cepas mexicanas, americanas, australianas, brasileñas y puertorriqueña de *A. marginale*, muestran una cobertura mayor del 76%, y una identidad mayor del 98%. Estos datos confirman que los 24 genomas de *A. marginale* pertenecen a la misma especie. Por otra parte, los genomas de las cepas australianas, brasileñas, norteamericana (St. Maries) y dos cepas mexicanas (MEX-30-184-02 de Tlapacoyan, Veracruz y MEX-31-096-01 de Tizimín, Yucatán), tienen una estructura altamente conservada (Figura 20). Sin embargo, los genomas de cinco cepas mexicanas: MEX-01-001-01 (Aguascalientes, Aguascalientes), MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estado de México), MEX-17-017-01 (Puente de Ixtla, Morelos) y MEX-30-193-01 (Veracruz, Veracruz) muestran reordenamientos e inversiones de segmentos genómicos (Figura 20). Brevemente, las cepas MEX-14-010-01 (Atitalaquia, Hidalgo), MEX-15-099-01 (Texcoco, Estados de México), MEX-17-017-01 (Puente de Ixtla, Morelos) y MEX-30-193-01 (Veracruz, Veracruz), muestran la inversión de un fragmento genómico (en color azul) entre las 300 y 400 Kb. También, las cepas MEX-01-001-01 (Aguascalientes, Aguascalientes) y MEX-14-010-01 (Atitalaquia, Hidalgo) muestran la inversión y movimiento de un fragmento genómico (en color rojo oscuro) entre las 500 y 700 Kb. Además, solamente la cepa MEX-17-017-01 (Puente de Ixtla, Morelos) muestra la inversión de un fragmento genómico (en color marrón) en la posición ~1 Mb (Figura 20). Por lo tanto, se sugiere que los genomas de las cepas mexicanas muestran una gran cantidad de modificaciones estructurales en comparación con

cepas de otros países, lo cual podría ser por presión selectiva y condiciones geográficas que existen en México.

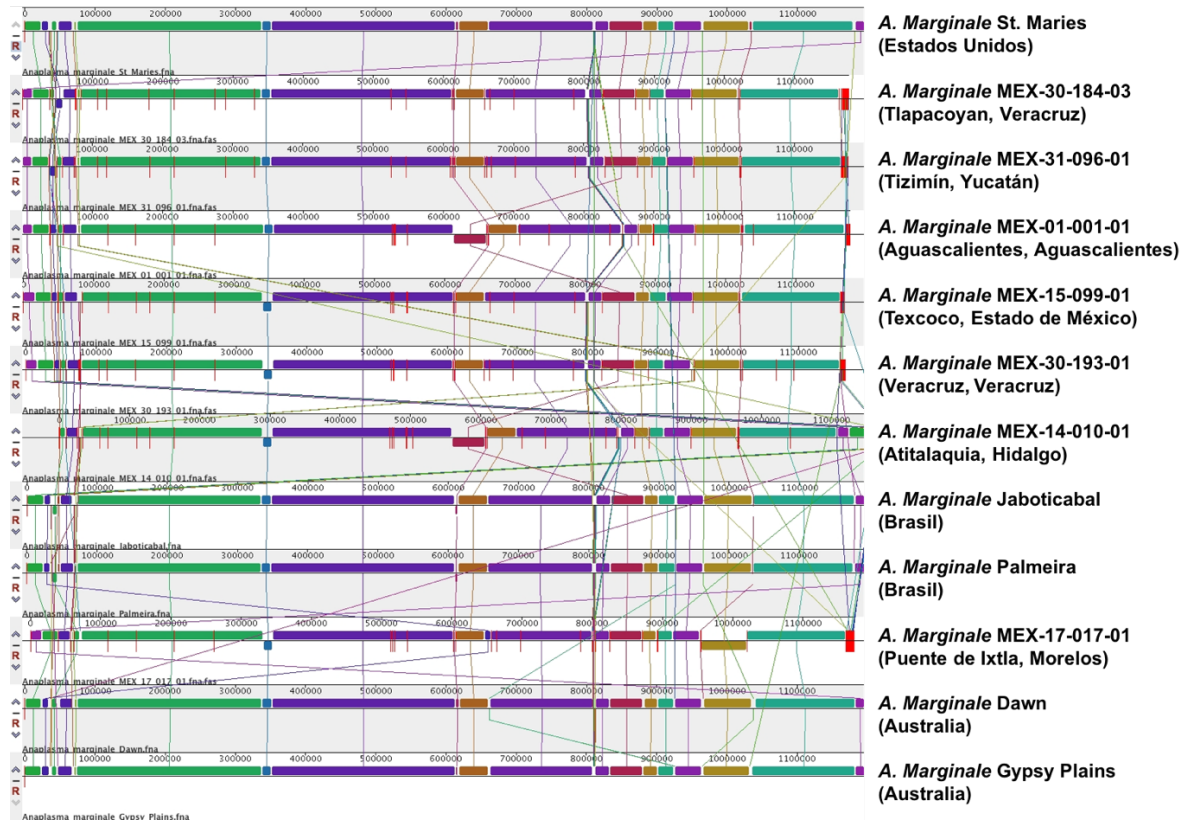


Figura 20. Estructura del genoma (sintenia) de las siete cepas mexicanas, una cepa americana, dos cepas australianas y dos cepas brasileñas de *A. marginale*.

Por último, el mapa circular de la genómica comparativa muestra que los genomas de dos cepas australianas, dos cepas brasileñas, siete cepas mexicanas y dos cepas norteamericanas están altamente conservados (Figura 21). No obstante, los genomas contienen regiones que muestran una elevada diversidad y variabilidad genética, por ejemplo, las secuencias de los genes *msp2* y *msp3* (flechas y letras moradas en la Figura 21). Estos resultados sugieren que los genomas de cepas de *A. marginale* que provienen de diferentes regiones

geográficas a nivel mundial están altamente conservados, pero muestran regiones genómicas que son específicas para cada cepa. Además, los genomas de las cepas mexicanas de *A. marginale* tienen un mayor número de eventos evolutivos.

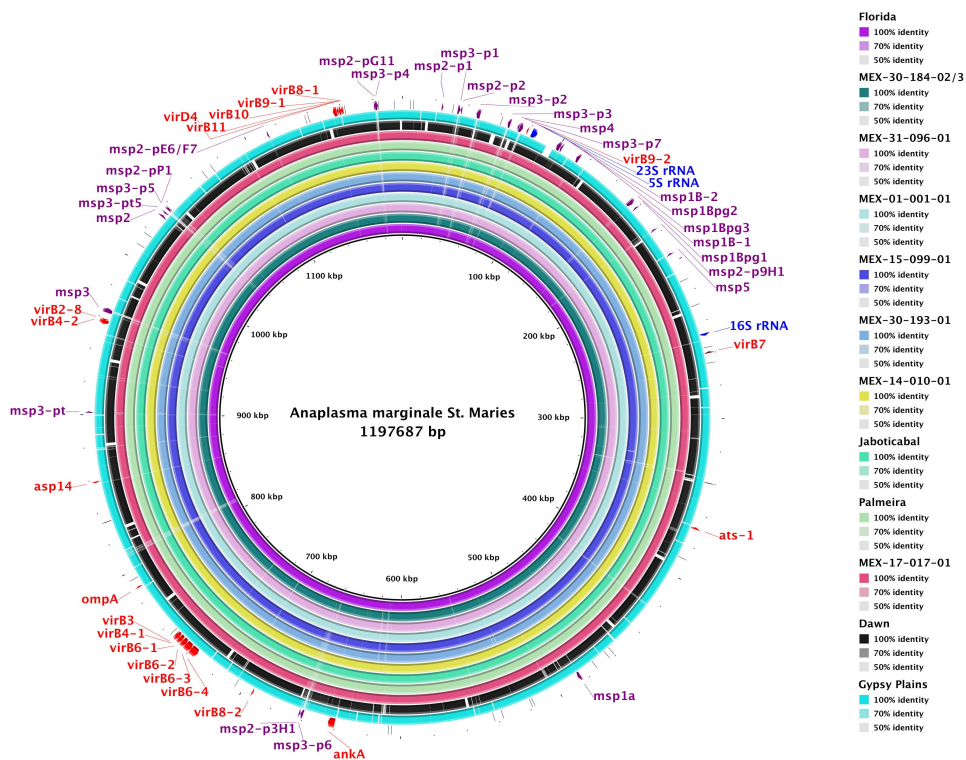


Figura 21. Mapa circular de la genómica comparativa de las siete cepas mexicanas, dos cepas americanas, dos cepas australianas y dos cepas brasileñas de *A. marginale*. El anillo externo muestra los genes rRNAs (azul), tRNAs (líneas negras), *msp* (morado) y que están involucrados con la patogenicidad o virulencia (rojo).

6.9 Identificación de los posibles candidatos vacunales

La anotación automática de las CDS que pertenecen a los genomas finales de las siete cepas mexicanas de *A. marginale* (Tabla 19), muestra que los programas Prokka y GeneMark generaron entre 944 y 981 CDS. Además, los algoritmos Classic RAST y RASTtk del servidor RAST generaron entre 1,138 y 1,178 CDS. Los

archivos FAA (del inglés *Fasta Amino Acid*) obtenidos con el algoritmo Classic RAST, fueron utilizados para los análisis posteriores de localización subcelular y la predicción de antígenos protectores y epítomos de tipo B, debido a que tienen más de 1,100 CDS (número similar a los genomas de *A. marginale* que están reportados en la base de datos GenBank (Tabla 2)) y fueron utilizados para los análisis pangenómicos (sección 6.7) y genómica comparativa (sección 6.8).

Tabla 19. Secuencias codificantes (CDS) que fueron anotadas en los genomas finales de las siete cepas mexicanas de *A. marginale*, al utilizar cuatro diferentes algoritmos.

Cepa	Región geográfica	Classic RAST	RASTtk	Prokka	GeneMark
MEX-01-001-01	Aguascalientes, Aguascalientes	1,178	1,167	981	948
MEX-14-010-01	Atitalaquia, Hidalgo	1,150	1,159	964	955
MEX-15-099-01	Texcoco, Estado de México	1,145	1,163	972	945
MEX-17-017-01	Puente de Ixtla, Morelos	1,163	1,154	970	953
MEX-30-184-02	Tlapacoyan, Veracruz	1,165	1,171	975	944
MEX-30-193-01	Veracruz, Veracruz	1,138	1,144	965	944
MEX-31-096-01	Tizimin, Yucatán	1,164	1,166	977	947

La localización subcelular de las proteínas de membrana externa y extracelulares que pertenecen a las siete cepas mexicanas de *A. marginale* (Tabla 20), muestra que el servidor PSORTb identificó entre 9 y 18 proteínas, SOSUIGramN identificó entre 29 y 39 proteínas, y CELLO identificó entre 63 y 118 proteínas.

Tabla 20. Proteínas de membrana externa y extracelulares identificadas en las siete cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	PSORTb		CELLO		SOSUIGramN	
		Membrana externa	Extracelular	Membrana externa	Extracelular	Membrana externa	Extracelular
MEX-01-001-01	Aguascalientes, Aguascalientes	15	13	118	74	33	39
MEX-14-010-01	Atitalaquia, Hidalgo	17	10	115	67	30	34
MEX-15-099-01	Texcoco, Estado de México	15	9	116	69	33	36
MEX-17-017-01	Puente de Ixtla, Morelos	18	11	116	63	32	39
MEX-30-184-02	Tlapacoyan, Veracruz	17	12	115	69	32	35
MEX-30-193-01	Veracruz, Veracruz	18	10	116	67	31	33
MEX-31-096-01	Tizimín, Yucatán	16	12	116	73	29	38

Los genomas de las siete cepas mexicanas de *A. marginale*, contienen entre 48 y 57 proteínas de membrana externa y extracelulares que fueron identificadas por al menos dos servidores (Tabla 21), y fueron utilizadas para los análisis posteriores de predicción de antígenos protectores y epítomos de tipo B, debido al número excesivo de proteínas identificadas por los tres servidores. Además, las proteínas de membrana externa y extracelulares al ser identificadas por dos o tres servidores diferentes, aumenta el nivel de certeza de la localización subcelular.

Tabla 21. Antígenos protectores identificados en las siete cepas mexicanas de *A. marginale*.

Cepa	Región geográfica	Proteínas de membrana externa y extracelulares*	Antígenos protectores**
MEX-01-001-01	Aguascalientes, Aguascalientes	57	18
MEX-14-010-01	Atitalaquia, Hidalgo	48	14
MEX-15-099-01	Texcoco, Estado de México	50	12
MEX-17-017-01	Puente de Ixtla, Morelos	57	15
MEX-30-184-03	Tlapacoyan, Veracruz	56	18
MEX-30-193-01	Veracruz, Veracruz	48	12
MEX-31-096-01	Tizimín, Yucatán	53	14

* Proteínas identificadas por dos o más servidores diferentes (PSORTb, CELLO y SOSUIGramN).

** Proteínas identificadas por el servidor VaxiJen al utilizar un valor de corte igual a 0.7.

La predicción de los antígenos protectores que pertenecen a las siete cepas mexicanas de *A. marginale* (Tabla 21), muestra que el servidor VaxiJen identificó entre 12 y 18 proteínas al utilizar un valor de corte igual a 0.7. Los antígenos protectores al ser identificados por un valor de corte estricto, aumenta el nivel de certeza de ser reconocidos por el sistema inmune del hospedero. Posteriormente, 12 antígenos protectores (Tabla 22) fueron seleccionados porque están conservados en las siete cepas mexicanas, no muestran homología con las proteínas de eucariotas, y además, también fueron anotados por el algoritmo RASTtk del servidor RAST y los programas Prokka y GenMark. Los 12 antígenos protectores seleccionados tienen el mayor potencial antigénico predicho para ser considerados como posibles candidatos vacunales contra la anaplasmosis bovina en México.

Tabla 22. 12 antígenos protectores conservados en las siete cepas mexicanas de *A. marginale*.

Proteína*	Longitud (aa)	Anotación**	Antígenos protectores	Epítomos de tipo B (18 aa)#	Péptido señal ^o	Región transmembranal ⁺
peg.54	452	FIG00385025: proteína hipotética	1.0268	18	Sí	12 - 34
peg.33	203	FIG00385151: proteína hipotética	0.9856	5	No	No
peg.684	222	FIG00384993: proteína hipotética	0.9341	7	No	84 - 101
peg.12	146	Probable proteína de unión a hierro de la familia HesB_IscA_SufA	0.8612	6	No	No
peg.445	416	FIG00385650: proteína hipotética (Ats-1)	0.8388	18	No	No
peg.139	310	FIG00385229: proteína hipotética	0.7560	15	No	No
peg.125	906	FIG00385106: proteína hipotética	0.7259	25	No	11 - 33
peg.622	332	FIG00384989: proteína hipotética (OMP11)	0.7210	7	Sí	No
peg.926	210	Proteína hipotética	0.7084	8	No	No
peg.389	1,213	Proteína reguladora de alginato AlgP	0.7078	55	No	No
peg.624	340	Proteína de membrana externa p28 (OMP13)	0.7070	13	No	36 - 58
peg.952	313	Proteína de opacidad y antígenos de superficie relacionados (MSP2)	0.7002	7	No	No

* Proteínas de la cepa MEX-01-001-01 (Aguascalientes. Ags.) como referencia.

** Datos obtenidos con el algoritmo Classic RAST del servidor RAST.

^ Datos obtenidos con el servidor VaxiJen.

Datos obtenidos con el servidor ABCpred al utilizar un valor de corte de 0.79.

^o Datos obtenidos con el servidor SignalP-5.0.

⁺ Datos obtenidos con el servidor TMHMM 2.0.

Para finalizar, se identificaron 47 epítomos de tipo B que están conservados en los genomas de las siete cepas mexicanas y pueden ser utilizados para el diseño de MAPs (del inglés *Multiple Antigenic Peptides*) [46], como una alternativa para el diseño de vacunas contra la anaplasmosis bovina en México.

Tabla 23. 47 epítomos tipo B que están conservados en las siete cepas mexicanas de *A. marginale*.

Proteína	Secuencia	Antígeno protector*
peg.54	TEQTGGDVRKSGAGAAG	2.6704
peg.54	KRGGDKKIAKGRGTEKSG	2.2433
peg.926	TQAQSSATVGGGGAQQTP	2.0799
peg.590	ETSPFDWDWEGSGTKGSK	2.0207
peg.139	DGSQQPGEQTVQAQSTDG	1.8503
peg.624	VGDQRGLKGMALRWRLAV	1.8151
peg.624	TTNAAATTNAAAGSSAG	1.8044
peg.445	IKNDGSARSGSDATGPRS	1.7814
peg.952	GTDTEEHAGGGPTLLSTT	1.6044
peg.389	PASAEKAGRSASTSSQKD	1.6005
peg.445	RIQNTPSERARGNDVSAS	1.5739
peg.389	SVAQQHGDVGTGSRADAS	1.5683
peg.139	RTEGEQVGKTSAPPEESA	1.5428
peg.445	SVDSSGSEVGDMDVLGLS	1.5218
peg.624	DVRVSDVHISYLGAEIGL	1.4739
peg.12	VSVQSGGCAGFKYEFATG	1.4463
peg.389	KNSAQTRESRAEFANRLN	1.3202
peg.54	ARVEGARFLNVDLKPGE	1.2951
peg.125	VPVDTTSSSPRGEKEKF	1.2856
peg.622	DQDVRHVRNLNGVTIAANS	1.2781
peg.125	LDLPGKNNRLSVSGVTSE	1.2582
peg.125	QTSKVSGHVSVSGSIQTQ	1.2468
peg.125	PSDLEGRASVALSSRGMR	1.2424
peg.12	FKYEFATGSYEDDADSSD	1.2158
peg.12	DSGFDDMDGDEEDNDEVE	1.1931
peg.389	TPRNSQQVSENAAKAEP	1.1925
peg.389	AQSESTVRETGATTPAAK	1.1853
peg.445	NDVSASYGQSGKLLTIA	1.1801
peg.125	IDLASFEGLDGSIEVHVK	1.1705
peg.389	SSVSASSDFEFRTGEQYY	1.1653
peg.684	CGHVWFFSAEEGTAKSSS	1.1627
peg.389	TAAKTSAQQASVAQQHGD	1.1618
peg.389	SAQSESTVRETGVTTTAA	1.1547
peg.622	FRTAPLNVLPMLWFAGQD	1.1452
peg.54	RFLNVDLKPGDEALEKSE	1.1271
peg.139	DVTTTQDQLPESVNGSAA	1.1263
peg.622	QAKVGASYRFGDNVHGFA	1.1164
peg.139	INAHDNVSISSDDFSIHG	1.1088
peg.33	MSSHEDAESDKSALAIAC	1.0845
peg.33	HHLDARDDPHLYDPWRFN	1.0835
peg.389	TQQAERVIGGHSGAQKAM	1.0739
peg.926	GEASWSTTYARPHSTTGG	1.0638
peg.590	MIGIALGGIGGYRIKIPA	1.0545
peg.139	PIIEVSAPTETANSDETA	1.0355
peg.445	ARTRDGFERTLKVTLDE	1.031
peg.445	EVGDDMVLGLSAAQSSGA	1.0236
peg.622	TDGGDANAMVIVKNEGVQ	1.0084

* Datos obtenidos con el servidor VaxiJen.

7. CONCLUSIONES

Con base en los resultados obtenidos en el presente trabajo, se concluye lo siguiente:

- Las cepas mexicanas de *A. marginale* y los hemoplasmas ('*Candidatus Mycoplasma haemobos*' y *Mycoplasma wenyonii*) coinfectan a los eritrocitos de los bovinos.
- Los borradores de genomas de dos cepas mexicanas mostraron una superinfección de diferentes cepas de *A. marginale*.
- Los borradores de genomas de siete cepas mexicanas de *A. marginale* fueron ensamblados con una elevada calidad y reportados en la base de datos GenBank de NCBI.
- Las estructuras de repetidas de MSP1a de cinco cepas mexicanas no coinciden con las estructuras reportadas previamente en la literatura.
- En México existe una amplia diversidad de cepas de *A. marginale*, las cuales están altamente relacionadas con cepas de diferentes regiones geográficas a nivel mundial.
- Las cepas mexicanas de *A. marginale* contienen antígenos protectores que pueden ser utilizados como posibles candidatos vacunales contra la anaplasmosis bovina en México.

8. PERSPECTIVAS

- Secuenciar el genoma de las siete cepas mexicanas de *A. marginale* con la tecnología NGS de tercera generación (por ejemplo, con la plataforma PacBio RS II) con la finalidad de ensamblar los genomas completos en un cromosoma.
- Secuenciar, ensamblar y analizar los genomas de cepas africanas, asiáticas y europeas de *A. marginale* para tener una mayor cobertura de diferentes cepas de *A. marginale* a nivel mundial.
- Evaluar *in vitro* el potencial antigénico e inmunogénico de los candidatos vacunales de las cepas mexicanas de *A. marginale*, mediante la amplificación y clonación de los candidatos vacunales o la síntesis de los epítomos tipo B en MAPs (Péptidos antigénicos múltiples).

9. BIBLIOGRAFÍA

1. Battilani M, De Arcangeli S, Balboni A, Dondi F. Genetic diversity and molecular epidemiology of *Anaplasma*. *Infect Genet Evol.* 2017;49:195–211.
2. Dumler JS, Barbet AF, Bekker C, Dasch GA, Palmer GH, Ray SC, *et al.* Reorganization of genera in the families Rickettsiaceae and Anaplasmataceae in the order Rickettsiales: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combi. *Int J Syst Evol Microbiol.* 2001;51(6):2145–65.
3. Rikihisa Y, Zhang C, Christensen BM. Molecular Characterization of *Aegyptianella pullorum* (Rickettsiales, Anaplasmataceae). *J Clin Microbiol.* 2003;41(11):5294–7.
4. Ybañez AP, Matsumoto K, Kishimoto T, Inokuma H. Molecular analyses of a potentially novel *Anaplasma* species closely related to *Anaplasma phagocytophilum* detected in sika deer (*Cervus nippon yesoensis*) in Japan. *Vet Microbiol.* 2012;157(1–2):232–6.
5. Tate CM, Howerth EW, Mead DG, Dugan VG, Luttrell MP, Sahara AI, *et al.* *Anaplasma odocoilei* sp. nov. (family Anaplasmataceae) from white-tailed deer (*Odocoileus virginianus*). *Ticks Tick Borne Dis.* 2013;4(1–2):110–9.
6. Li H, Zheng Y-C, Ma L, Jia N, Jiang B-G, Jiang R-R, *et al.* Human infection with a novel tick-borne *Anaplasma* species in China: a surveillance study. *Lancet Infect Dis.* 2015;15(6):663–70.
7. Rikihisa Y. The tribe Ehrlichieae and ehrlichial diseases. *Clin Microbiol Rev.* 1991;4(3):286–308.
8. Kieser ST, Eriks IS, Palmer GH. Cyclic rickettsemia during persistent *Anaplasma marginale* infection of cattle. *Infect Immun.* 1990;58(4):1117–9.
9. Kocan KM, Stiller D, Goff WL, Claypool PL, Edwards W, Ewing SA, *et al.* Development of *Anaplasma marginale* in male *Dermacentor andersoni* transferred from parasitemic to susceptible cattle. *Am J Vet Res.* 1992;53(4):499–507.
10. Villar M, López V, Ayllón N, Cabezas-Cruz A, López JA, Vázquez J, *et al.* The intracellular bacterium *Anaplasma phagocytophilum* selectively manipulates the levels of vertebrate host proteins in the tick vector *Ixodes scapularis*. *Parasit Vectors.* 2016;9(1):467.
11. Ueti MW, Knowles DP, Davitt CM, Scoles GA, Baszler T V, Palmer GH. Quantitative differences in salivary pathogen load during tick transmission underlie strain-specific variation in transmission efficiency of *Anaplasma marginale*. *Infect Immun.* 2009;77(1):70–5.
12. Macleod J, Gordon WS. Studies in Tick-borne Fever of Sheep. I. Transmission by the Tick, *Ixodes ricinus*, with a Description of the Disease Produced. *Parasitology.* 1933;25(2):273–83.
13. Rar V, Golovljova I. *Anaplasma*, *Ehrlichia*, and *Candidatus Neoehrlichia* bacteria: pathogenicity, biodiversity, and molecular genetic characteristics, a review. *Infect Genet Evol.* 2011;11(8):1842–61.
14. Kocan KM, de la Fuente J, Guglielmone AA, Melendez RD. Antigenes and Alternatives for Control of *Anaplasma marginale* Infection in Cattle. *Clin*

- Microbiol Rev. 2003;16(4):698–712.
15. Aubry P, Geale DW. A Review of Bovine Anaplasmosis. *Transbound Emerg Dis.* 2011;58(1):1–30.
 16. Harrison A, Bastos ADS, Medger K, Bennett NC. Eastern rock sengis as reservoir hosts of *Anaplasma bovis* in South Africa. *Ticks Tick Borne Dis.* 2013;4(6):503–5.
 17. Dugat T, Lagrée A-C, Maillard R, Boulouis H-J, Haddad N. Opening the black box of *Anaplasma phagocytophilum* diversity: current situation and future perspectives. *Front Cell Infect Microbiol.* 2015;5:61.
 18. Woldehiwet Z. The natural history of *Anaplasma phagocytophilum*. *Vet Parasitol.* 2010;167(2–4):108–22.
 19. de la Fuente J, Lew A, Lutz H, Meli ML, Hofmann-Lehmann R, Shkap V, *et al.* Genetic diversity of *Anaplasma* species major surface proteins and implications for anaplasmosis serodiagnosis and vaccine development. *Anim Heal Res Rev.* 2005;6(1):75–89.
 20. Theiler A. Gall-sickness of South Africa. (Anaplasmosis of Cattle.). *J Comp Pathol Ther.* 1910;23:98–115.
 21. Kocan KM, de la Fuente J, Blouin EF, Coetzee JF, Ewing SA. The natural history of *Anaplasma marginale*. *Vet Parasitol.* 2010;167(2–4):95–107.
 22. Kocan KM, de la Fuente FJ, Cabezas-Cruz A. The genus *Anaplasma*: new challenges after reclassification. *Rev Sci Tech l’OIE.* 2015;34(2):577–86.
 23. de la Fuente J, Garcia-Garcia JC, Blouin EF, Saliki JT, Kocan KM. Infection of Tick Cells and Bovine Erythrocytes with One Genotype of the Intracellular Ehrlichia *Anaplasma marginale* Excludes Infection with Other Genotypes. *Clin Vaccine Immunol.* 2002;9(3):658–68.
 24. Palmer GH, Knowles DP, Rodriguez J-L, Gnad DP, Hollis LC, Marston T, *et al.* Stochastic transmission of multiple genotypically distinct *Anaplasma marginale* strains in a herd with high prevalence of *Anaplasma* infection. *J Clin Microbiol.* 2004;42(11):5381–4.
 25. Castañeda-Ortiz EJ, Ueti MW, Camacho-Nuez M, Mosqueda JJ, Mousel MR, Johnson WC, *et al.* Association of *Anaplasma marginale* Strain Superinfection with Infection Prevalence within Tropical Regions. Ganta R, editor. *PLoS One.* 2015;10(3):e0120748.
 26. Horobin RW, Walter KJ. Understanding Romanowsky staining. I: The Romanowsky-Giemsa effect in blood smears. *Histochemistry.* 1987;86(3):331–6.
 27. Torioni de Echaide S, Knowles DP, McGuire TC, Palmer GH, Suarez CE, McElwain TF. Detection of cattle naturally infected with *Anaplasma marginale* in a region of endemicity by nested PCR and a competitive enzyme-linked immunosorbent assay using recombinant major surface protein 5. *J Clin Microbiol.* 1998;36(3):777–82.
 28. Kuttler KL. Dynamics of *Anaplasma marginale* in splenectomised calves treated with either imidocarb or oxytetracycline. *Trop Anim Health Prod.* 1986;18(2):91–6.
 29. Vega y Murguía CA, Jiménez-Ocampo R, García-Ortiz MA, Preciado De la Torre JF, Rojas Ramírez EE, Rodríguez-Camarillo SD. Cómo evitar las pérdidas por la anaplasmosis bovina. CENID-PAVET, INIFAP. Folleto técnico

- No. 9. 2010;1–44 p.
30. Thullner F, Willadsen P, Kemp D. Acaricide rotation strategy for managing resistance in the tick *Rhipicephalus (Boophilus) microplus* (Acarina: Ixodidae): laboratory experiment with a field strain from Costa Rica. *J Med Entomol*. 2007;44(5):817–21.
 31. Pipano E. Live vaccines against hemoparasitic diseases in livestock. *Vet Parasitol*. 1995;57(1–3):213–31.
 32. Rodríguez Camarillo SD, García Ortiz MÁ, Rojas Ramírez EE, Cantó Alarcón GJ, Preciado de la Torre JF, Rosario Cruz R, *et al*. *Anaplasma marginale* Yucatan (Mexico) Strain. Assessment of low virulence and potential use as a live vaccine. *Ann N Y Acad Sci*. 2008;1149(1):98–102.
 33. Palmer GH, McElwain TF. Molecular basis for vaccine development against anaplasmosis and babesiosis. *Vet Parasitol*. 1995;57(1–3):233–53.
 34. Quiroz-Castañeda RE, Amaro-Estrada I, Rodríguez-Camarillo SD. *Anaplasma marginale*: Diversity, Virulence, and Vaccine Landscape through a Genomics Approach. *Biomed Res Int*. 2016;2016:9032085.
 35. Carter PD, Bock RE, DeVos AJ. Benefits of Improved Vaccine for Anaplasmosis in Australia. In: Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics. 2006.
 36. Rogers RJ, Dimmock CK, de Vos AJ, Rodwell BJ. Bovine leucosis virus contamination of a vaccine produced in vivo against bovine babesiosis and anaplasmosis. *Aust Vet J*. 1988;65(9):285–7.
 37. Bell-Sakyi L, Palomar AM, Bradford EL, Shkap V. Propagation of the Israeli vaccine strain of *Anaplasma centrale* in tick cell lines. *Vet Microbiol*. 2015;179(3–4):270–6.
 38. Ocampo Espinoza V, Vázquez JES, Aguilar MD, Ortiz MAG, Alarcón GJC, Rodríguez SD. *Anaplasma marginale*: lack of cross-protection between strains that share MSP1a variable region and MSP4. *Vet Microbiol*. 2006;114(1–2):34–40.
 39. Zhao L, Mahony D, Cavallaro AS, Zhang B, Zhang J, Deringer JR, *et al*. Immunogenicity of Outer Membrane Proteins VirB9-1 and VirB9-2, a Novel Nanovaccine against *Anaplasma marginale*. Russell B, editor. *PLoS One*. 2016;11(4):e0154295.
 40. de la Fuente J, Van Den Bussche RA, Kocan KM. Molecular phylogeny and biogeography of North American isolates of *Anaplasma marginale* (Rickettsiaceae: Ehrlichieae). *Vet Parasitol*. 2001;97(1):65–76.
 41. Vega LEO, Rodríguez SD, Alarcón GJC, Flores RL, Ocampo RJ, Ortiz MAG, *et al*. *Anaplasma marginale* field challenge: protection by an inactivated immunogen that shares partial sequence of *msp1alpha* variable region with the challenge strain. *Vaccine*. 2007;25(3):519–25.
 42. Riding G, Hope M, Waltisbuhl D, Willadsen P. Identification of novel protective antigens from *Anaplasma marginale*. *Vaccine*. 2003;21(17–18):1874–83.
 43. Lopez JE, Palmer GH, Brayton KA, Dark MJ, Leach SE, Brown WC. Immunogenicity of *Anaplasma marginale* type IV secretion system proteins in a protective outer membrane vaccine. *Infect Immun*. 2007;75(5):2333–42.
 44. de la Fuente J, Ruybal P, Mtshali MS, Naranjo V, Shuqing L, Mangold AJ, *et al*. Analysis of world strains of *Anaplasma marginale* using major surface

- protein 1a repeat sequences. *Vet Microbiol.* 2007;119(2–4):382–90.
45. Seib KL, Zhao X, Rappuoli R. Developing vaccines in the era of genomics: a decade of reverse vaccinology. *Clin Microbiol Infect.* 2012;18 Suppl 5:109–16.
 46. Tam JP. Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. *Proc Natl Acad Sci U S A.* 1988;85(15):5409–13.
 47. Robinson TP, Wint GRW, Conchedda G, Van Boeckel TP, Ercoli V, Palamara E, *et al.* Mapping the Global Distribution of Livestock. Baylis M, editor. *PLoS One.* 2014;9(5):e96084.
 48. Rodríguez SD, García Ortiz MA, Jiménez Ocampo R, Vega y Murguía CA. Molecular epidemiology of bovine anaplasmosis with a particular focus in Mexico. *Infect Genet Evol.* 2009;9(6):1092–101.
 49. Jiménez Ocampo R, Vega y Murguía CA, Oviedo Ortega N, Rojas Ramírez EE, García Ortiz MÁ, Preciado de la Torre JF, *et al.* Diversidad genética de la región variable de los genes *msp1a* y *msp4* en cepas de *Anaplasma marginale* de México. *Rev Mex Ciencias Pecu.* 2012;3(3):373–87.
 50. García Ortiz MÁ, Aboytes Torres R, Hernández Salgado G, Cantó Alarcón GJ, Rodríguez Camarillo SD. *Anaplasma marginale*: Diferentes grados de virulencia en dos aislados mexicanos. *Vet Méx.* 2000;31(2):157–60.
 51. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20.
 52. Cox MP, Peterson DA, Biggs PJ. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics.* 2010;11(1):485.
 53. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol.* 2012;19(5):455–77.
 54. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–10.
 55. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* 2013;29(8):1072–5.
 56. Catanese HN, Brayton KA, Gebremedhin AH. RepeatAnalyzer: a tool for analysing and managing short-sequence repeat data. *BMC Genomics.* 2016;17(1):422.
 57. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9.
 58. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–9.
 59. Carver T, Harris SR, Otto TD, Berriman M, Parkhill J, McQuillan JA. BamView: visualizing and interpretation of next-generation sequencing read alignments. *Brief Bioinform.* 2013;14(2):203–12.
 60. de La Fuente J, Garcia-Garcia JC, Blouin EF, Rodríguez SD, García MA, Kocan KM. Evolution and function of tandem repeats in the major surface protein 1a of the ehrlichial pathogen *Anaplasma marginale*. *Anim Heal Res Rev.* 2001;2(2):163–73.
 61. Estrada-Peña A, Naranjo V, Acevedo-Whitehouse K, Mangold AJ, Kocan KM,

- de la Fuente J. Phylogeographic analysis reveals association of tick-borne pathogen, *Anaplasma marginale*, MSP1a sequences with ecological traits affecting tick vector performance. *BMC Biol.* 2009;7(1):57.
62. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics.* 2008;9(1):75.
 63. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 2007;35(9):3100–8.
 64. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 2004;32(1):11–6.
 65. Tatusov RL, Galperin MY, Natale DA, Koonin E V. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000;28(1):33–6.
 66. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, *et al.* Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother.* 2012;67(11):2640–4.
 67. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, *et al.* antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 2015;43(W1):W237–43.
 68. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 2017;45(D1):D566–D573.
 69. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, *et al.* Real-Time Whole-Genome Sequencing for Routine Typing, Surveillance, and Outbreak Detection of Verotoxigenic *Escherichia coli*. *J Clin Microbiol.* 2014;52(5):1501–10.
 70. Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. PathogenFinder—distinguishing friend from foe using bacterial whole genome sequence data. Barbour JD, editor. *PLoS One.* 2013;8(10):e77302.
 71. Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.* 2016;44(D1):D694–7.
 72. Yoon SH, Park Y-K, Kim JF. PAIDB v2.0: exploration and analysis of pathogenicity and resistance islands. *Nucleic Acids Res.* 2015;43(Database issue):D624–30.
 73. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30(14):2068–9.
 74. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics.* 2015;31(22):3691–3.
 75. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods.* 2012;9(8):772.
 76. Guindon S, Gascuel O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Rannala B, editor. *Syst Biol.* 2003;52(5):696–704.
 77. Contreras-Moreira B, Vinuesa P. GET_HOMOLOGUES, a versatile software

- package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol.* 2013;79(24):7696–701.
78. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. *Genome Res.* 2004;14(7):1394–403.
 79. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics.* 2011;12(1):402.
 80. Besemer J, Borodovsky M. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* 2005;33(Web Server issue):W451–4.
 81. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, *et al.* PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics.* 2010;26(13):1608–15.
 82. Yu C-S, Lin C-J, Hwang J-K. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.* 2004;13(5):1402–6.
 83. Imai K, Asakawa N, Tsuji T, Akazawa F, Ino A, Sonoyama M, *et al.* SOSUI-GramN: high performance prediction for sub-cellular localization of proteins in gram-negative bacteria. *Bioinformation.* 2008;2(9):417–21.
 84. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics.* 2007;8(1):4.
 85. Saha S, Raghava GPS. BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties. In Springer, Berlin, Heidelberg; 2004. p. 197–204.
 86. Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins.* 2006;65(1):40–8.
 87. Zhang Q, Wang P, Kim Y, Haste-Andersen P, Beaver J, Bourne PE, *et al.* Immune epitope database analysis resource (IEDB-AR). *Nucleic Acids Res.* 2008;36(Web Server):W513--W518.
 88. Cossío-Bayúgar R, Rodríguez SD, García-Ortiz MA, García-Tapia D, Aboytes-Torres R. Bovine anaplasmosis prevalence in northern Veracruz state, Mexico. *Prev Vet Med.* 1997;32(3–4):165–70.
 89. García Tapia D, López Rojas M, Cossío Bayúgar R, García Vázquez Z, García Ortiz MA, Dimínguez Jalil P, *et al.* Seroprevalencia de anaplasmosis en explotaciones bovinas de 18 municipios de la zona norte de Veracruz. *Tec Pecu Mex.* 1996;34(1):38–45.
 90. Fragoso SH. La anaplasmosis bovina en México. Garrapatas y enfermedades que transmiten. In: Segundo Seminario Internacional de Parasitología Animal. 1991. p. 153–160.
 91. Jiménez Ocampo R, Rodríguez Camarillo SD, Rosario Cruz R, Orozco Vega LE, de la Fuente J. *Anaplasma marginale*: Analysis of variable fragment sequences in *msp1 α* and *msp4* genes in four new Mexican strains. *Tec Pecu en Mex.* 2008;46(1):69–78.
 92. de la Fuente J, Van Den Bussche RA, Garcia-Garcia JC, Rodríguez SD,

- García MA, Guglielmone AA, *et al.* Phylogeography of New World isolates of *Anaplasma marginale* based on major surface protein sequences. *Vet Microbiol.* 2002;88(3):275–85.
93. Munderloh UG, Liu Y, Wang M, Chen C, Kurtti TJ. Establishment, maintenance and description of cell lines from the tick *Ixodes scapularis*. *J Parasitol.* 1994;80(4):533–43.
 94. Bastos C V, Passos LMF, Vasconcelos MMC, Ribeiro MFB. *In vitro* establishment and propagation of a Brazilian strain of *Anaplasma marginale* with appendage in IDE8 (*Ixodes scapularis*) cells. *Braz J Microbiol.* 2009;40(2):399–403.
 95. Blouin EF, Kocan KM. Morphology and Development of *Anaplasma marginale* (Rickettsiales: Anaplasmataceae) in Cultured *Ixodes scapularis* (Acari: Ixodidae) Cells. *J Med Entomol.* 1998;35(5):788–97.
 96. Silveira JAG, Silvestre BT, Bastos C V, Ribeiro MFB. Isolation and attempted cultivation of an *Anaplasma marginale* strain from Brazilian brown brocket deer (*Mazama gouazoubira*, Fisher, 1814) in the tick cell line IDE8. *Ticks Tick Borne Dis.* 2016;7(6):1102–8.
 97. Luna-Castro G, Orozco-Vega L, Vega y Murguía C, Rodríguez-Camarillo S, Ramírez-Noguera P. El cultivo *in vitro* de *Anaplasma marginale*: ¿Evolución o involución? In: Moreno-Chan R, editor. *Ciencia Veterinaria Volumen 10.* UNAM; 2007. p. 174–208.
 98. Machuca Figueroa S, Granjeno Colín G, Rodríguez Camarillo SD, Vega y Murguía CA. Propiedades de crecimiento de las líneas celulares DH82 y RF/6A bajo condiciones normales de laboratorio. *Rev Mex Ciencias Pecu.* 2016;7(1):85–104.
 99. Cobaxin Cárdenas ME, Amaro Estrada I, Quiroz Castañeda RE, Preciado de la Torre JF, Rodríguez Camarillo SD, Vega y Murguía CA. Determinación del tiempo de duplicación *in vitro* de la línea celular BUVEC-E6E7 de origen bovino. *Rev Mex Ciencias Pecu.* 2018;9(1):162–70.
 100. Alleman R, Kamper MS, Viseshakul N, Barbet AF. Analysis of the *Anaplasma marginale* genome by pulsed-field electrophoresis. *J Gen Microbiol.* 1993;139(10):2439–44.
 101. Hallin PF, Nielsen N, Devine KM, Binnewies TT, Willenbrock H, Ussery DW. Genome update: base skews in 200+ bacterial chromosomes. *Microbiology.* 2005;151(Pt 3):633–7.
 102. Ueti MW, Tan Y, Broschat SL, Castañeda Ortiz EJ, Camacho-Nuez M, Mosqueda JJ, *et al.* Expansion of Variant Diversity Associated with a High Prevalence of Pathogen Strain Superinfection under Conditions of Natural Transmission. Morrison RP, editor. *Infect Immun.* 2012;80(7):2354–60.
 103. Vallejo Esquerria E, Herndon DR, Alpirez Mendoza F, Mosqueda J, Palmer GH. *Anaplasma marginale* superinfection attributable to pathogen strains with distinct genomic backgrounds. Morrison RP, editor. *Infect Immun.* 2014;82(12):5286–92.
 104. Munderloh UG, Lynch MJ, Herron MJ, Palmer AT, Kurtti TJ, Nelson RD, *et al.* Infection of endothelial cells with *Anaplasma marginale* and *A. phagocytophilum*. *Vet Microbiol.* 2004;101(1):53–64.
 105. Carreño AD, Alleman AR, Barbet AF, Palmer GH, Noh SM, Johnson CM. *In*

- vivo* Endothelial Cell Infection by *Anaplasma marginale*. Vet Pathol. 2007;44(1):116–8.
106. Estrada-Peña A, Bouattour A, Camicas J-L, Guglielmone A, Horak I, Jongejan F, *et al.* The Known Distribution and Ecological Preferences of the Tick Subgenus *Boophilus* (Acari: *Ixodidae*) in Africa and Latin America. Exp Appl Acarol. 2006;38(2–3):219–35.
 107. Garcia-Garcia JC, de la Fuente J, Blouin EF, Johnson TJ, Halbur T, Onet VC, *et al.* Differential expression of the *msp1 α* gene of *Anaplasma marginale* occurs in bovine erythrocytes and tick cells. Vet Microbiol. 2004;98(3–4):261–72.
 108. Quiroz Castañeda RE, Amaro-Estrada I, Martínez-Ocampo F, Rodríguez-Camarillo S, Dantán González E, Cobaxin-Cárdenas M, *et al.* Draft Genome Sequence of *Anaplasma marginale* Strain Mex-01-001-01, a Mexican Strain That Causes Bovine Anaplasmosis. Cameron Thrash J, editor. Microbiol Resour Announc. 2018;7(16).
 109. Martínez-Ocampo F, Quiroz-Castañeda RE, Amaro-Estrada I, Cobaxin Cárdenas M, Dantán-González E, Rodríguez-Camarillo S. Draft Genome Sequences of *Anaplasma marginale* Strains MEX-15-099-01 and MEX-31-096-01, Two Mexican Isolates with Different Degrees of Virulence. Microbiol Resour Announc. 2019;8(45):7–8.

10. APÉNDICES


Apéndice 1

Genoma reportado:

Secuencia del genoma de *Anaplasma marginale* MEX-01-001-01
(Aguascalientes, Aguascalientes) que está depositado en la base de datos
GenBank

Anaplasma marginale strain MEX-01-001-01, whole genome shotgun sequencing project

GenBank: QLIV00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

Go to: 

```
LOCUS       QLIV01000000          34 rc  DNA   linear   BCT 14-SEP-2018
DEFINITION  Anaplasma marginale strain MEX-01-001-01, whole genome shotgun
sequencing project.
ACCESSION   QLIV00000000
VERSION     QLIV00000000.1
DBLINK     BioProject: PRJNA476110
           BioSample: SAMN09425886
KEYWORDS    WGS.
SOURCE     Anaplasma marginale
ORGANISM    Anaplasma marginale
           Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
           Anaplasmataceae; Anaplasma.
REFERENCE   1 (bases 1 to 34)
AUTHORS     Martinez Ocampo,F., Quiroz Castaneda,R.E., Amaro Estrada,I. and
Rodriguez Camarillo,S.
TITLE       Direct Submission
JOURNAL     Submitted (14-JUN-2018) Unidad de Anaplasmosis, CENID-PAVET,
INIFAP, Km. 11.5 de la Carretera Federal Cuernavaca-Cuautla, Col.
Progreso, Jiutepec, Morelos 62550, Mexico
COMMENT     The Anaplasma marginale whole genome shotgun (WGS) project has the
project accession QLIV00000000. This version of the project (01)
has the accession number QLIV01000000, and consists of sequences
QLIV01000001-QLIV01000034.
           Bacteria and source DNA available from Unidad de Anaplasmosis,
           CENID-PAVET, INIFAP.
           This project was originally submitted as Anaplasma marginale strain
           INIFAP09. The organism name was changed to Anaplasma marginale
           strain MEX-01-001-01 in September 2018.
           Annotation was added by the NCBI Prokaryotic Genome Annotation
           Pipeline (released 2013). Information about the Pipeline can be
           found here: https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/

##Genome-Assembly-Data-START##
Assembly Date      :: 13-JUN-2018
Assembly Method    :: SPAdes v. 3.11.1
Genome Representation :: Full
Expected Final Version :: Yes
Genome Coverage    :: 36.0x
Sequencing Technology :: Illumina HiSeq
##Genome-Assembly-Data-END##

##Genome-Annotation-Data-START##
Annotation Provider      :: NCBI
Annotation Date          :: 06/15/2018 16:07:58
Annotation Pipeline      :: NCBI Prokaryotic Genome
                          Annotation Pipeline
Annotation Method        :: Best-placed reference protein
                          set; GeneMarkS+
Annotation Software revision :: 4.5
Features Annotated      :: Gene; CDS; rRNA; tRNA; ncRNA;
                          repeat_region
Genes (total)           :: 1,053
CDS (total)             :: 1,009
Genes (coding)          :: 969
CDS (coding)            :: 969
Genes (RNA)             :: 44
rRNAs                   :: 1, 1, 1 (5S, 16S, 23S)
complete rRNAs         :: 1, 1, 1 (5S, 16S, 23S)
tRNAs                   :: 37
ncRNAs                  :: 4
Pseudo Genes (total)   :: 40
Pseudo Genes (ambiguous residues) :: 0 of 40
Pseudo Genes (frameshifted) :: 13 of 40
Pseudo Genes (incomplete) :: 27 of 40
Pseudo Genes (internal stop) :: 4 of 40
Pseudo Genes (multiple problems) :: 4 of 40
##Genome-Annotation-Data-END##
FEATURES             Location/Qualifiers
     source           1..34
                     /organism="Anaplasma marginale"
                     /mol_type="genomic DNA"
                     /strain="MEX-01-001-01"
                     /isolation_source="blood"
                     /host="Bos taurus"
                     /db_xref="taxon:770"
                     /country="Mexico: Morelos"
                     /collection_date="2015"
WGS                   QLIV01000001-QLIV01000034
//
```


Apéndice 2

Genoma reportado:

Secuencia del genoma de *Anaplasma marginale* MEX-14-010-01
(Atitalaquia, Hidalgo) que está depositado en la base de datos GenBank

Anaplasma marginale strain MEX-14-010- 01, whole genome shotgun sequencing project

GenBank: VTSO00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

Go to: 

```
LOCUS       VTSO01000000          46 rc  DNA   linear  BCT 04-SEP-2019
DEFINITION  Anaplasma marginale strain MEX-14-010- 01, whole genome shotgun
sequencing project.
ACCESSION   VTSO00000000
VERSION     VTSO00000000.1
DBLINK      BioProject: PRJNA562203
            BioSample: SAMN12636411
KEYWORDS    WGS.
SOURCE      Anaplasma marginale
  ORGANISM  Anaplasma marginale
            Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
            Anaplasmataceae; Anaplasma.
REFERENCE   1 (bases 1 to 46)
AUTHORS     Quiroz Castaneda,R.E., Amaro Estrada,I., Martinez Ocampo,F. and
            Rodriguez Camarillo,S.D.
TITLE       Direct Submission
JOURNAL     Submitted (27-AUG-2019) Unidad de Anaplasmosis, CENID-SAI, INIFAP,
            Km. 11.5 de la Carretera Federal Cuernavaca-Cuautla, Col. Progreso,
            Jiutepec, Morelos 62550, Mexico
COMMENT     The Anaplasma marginale whole genome shotgun (WGS) project has the
            project accession VTSO00000000. This version of the project (01)
            has the accession number VTSO01000000, and consists of sequences
            VTSO01000001-VTSO01000046.
            Bacteria and source DNA available from Anaplasmosis Unit CENID-SAI,
            INIFAP.
            The annotation was added by the NCBI Prokaryotic Genome Annotation
            Pipeline (PGAP). Information about PGAP can be found here:
            https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/

##Genome-Assembly-Data-START##
Assembly Date      :: 15-JUN-2019
Assembly Method    :: SPAdes v. 3.11.1
Genome Representation :: Full
Expected Final Version :: Yes
Genome Coverage    :: 22.0x
Sequencing Technology :: Illumina MiSeq
##Genome-Assembly-Data-END##

##Genome-Annotation-Data-START##
Annotation Provider      :: NCBI
Annotation Date          :: 08/30/2019 19:47:25
Annotation Pipeline      :: NCBI Prokaryotic Genome
                        Annotation Pipeline (PGAP)
Annotation Method        :: Best-placed reference protein
                        set; GeneMarkS-2+
Annotation Software revision :: 4.9
Features Annotated      :: Gene; CDS; rRNA; tRNA; ncRNA;
                        repeat_region
Genes (total)           :: 1,038
CDSs (total)            :: 994
Genes (coding)          :: 936
CDSs (with protein)     :: 936
Genes (RNA)             :: 44
rRNAs                   :: 1, 1, 1 (5S, 16S, 23S)
complete rRNAs          :: 1, 1, 1 (5S, 16S, 23S)
tRNAs                   :: 37
ncRNAs                  :: 4
Pseudo Genes (total)    :: 58
CDSs (without protein)  :: 58
Pseudo Genes (ambiguous residues) :: 0 of 58
Pseudo Genes (frameshifted) :: 18 of 58
Pseudo Genes (incomplete) :: 38 of 58
Pseudo Genes (internal stop) :: 6 of 58
Pseudo Genes (multiple problems) :: 4 of 58
##Genome-Annotation-Data-END##

FEATURES             Location/Qualifiers
     source            1..46
                        /organism="Anaplasma marginale"
                        /mol_type="genomic DNA"
                        /strain="MEX-14-010- 01"
                        /host="Bos taurus"
                        /db_xref="taxon:770"
                        /country="Mexico: Atitalaquia Hidalgo"
                        /collection_date="2017-09-04"
WGS                   VTSO01000001-VTSO01000046
//
```

Apéndice 3

Genoma reportado:

Secuencia del genoma de *Anaplasma marginale* MEX-15-099-01
(Texcoco, Estado de México) que está depositado en la base de datos GenBank

Anaplasma marginale strain MEX-15-099-01, whole genome shotgun sequencing project

GenBank: VTWW00000000.1

i This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

[Go to:](#)

LOCUS VTWW01000000 32 rc DNA linear BCT 23-SEP-2019
DEFINITION Anaplasma marginale strain MEX-15-099-01, whole genome shotgun sequencing project.
ACCESSION VTWW00000000
VERSION VTWW00000000.1
DBLINK BioProject: [PRJNA563358](#)
BioSample: [SAMN12672923](#)
KEYWORDS WGS.
SOURCE Anaplasma marginale
ORGANISM [Anaplasma marginale](#)
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Anaplasmataceae; Anaplasma.
REFERENCE 1 (bases 1 to 32)
AUTHORS Martinez-Ocampo,F., Quiroz-Castaneda,R.E., Amaro Estrada,I., Rodriguez Camarillo,S.D., Dantan Gonzalez,E. and Preciado De La Torre,J.F.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-2019) Unidad de Anaplasmosis, CENID-SAI, INIFAP, Km. 11.5 de la Carretera Federal Cuernavaca-Cuautla, Col. Progreso, Jiutepec, Morelos 62550, Mexico
COMMENT The Anaplasma marginale whole genome shotgun (WGS) project has the project accession VTWW00000000. This version of the project (01) has the accession number VTWW01000000, and consists of sequences VTWW01000001-VTWW01000032.
Bacteria and source DNA available from Anaplasmosis Unit CENID-SAI, INIFAP.
The annotation was added by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). Information about PGAP can be found here: https://www.ncbi.nlm.nih.gov/genome/annotation_prok/

##Genome-Assembly-Data-START##
Assembly Date :: 15-JUL-2019
Assembly Method :: SPAdes v. 3.11.1
Genome Representation :: Full
Expected Final Version :: Yes
Genome Coverage :: 30.0x
Sequencing Technology :: Illumina MiSeq
##Genome-Assembly-Data-END##

##Genome-Annotation-Data-START##
Annotation Provider :: NCBI
Annotation Date :: 09/05/2019 11:05:34
Annotation Pipeline :: NCBI Prokaryotic Genome Annotation Pipeline (PGAP)
Annotation Method :: Best-placed reference protein set; GeneMarkS-2+
Annotation Software revision :: 4.9
Features Annotated :: Gene; CDS; rRNA; tRNA; ncRNA; repeat_region
Genes (total) :: 1,040
CDSs (total) :: 996
Genes (coding) :: 940
CDSs (with protein) :: 940
Genes (RNA) :: 44
rRNAs :: 1, 1, 1 (5S, 16S, 23S)
complete rRNAs :: 1, 1, 1 (5S, 16S, 23S)
tRNAs :: 37
ncRNAs :: 4
Pseudo Genes (total) :: 56
CDSs (without protein) :: 56
Pseudo Genes (ambiguous residues) :: 0 of 56
Pseudo Genes (frameshifted) :: 17 of 56
Pseudo Genes (incomplete) :: 37 of 56
Pseudo Genes (internal stop) :: 5 of 56
Pseudo Genes (multiple problems) :: 3 of 56
##Genome-Annotation-Data-END##
FEATURES Location/Qualifiers
source 1..32
/organism="Anaplasma marginale"
/mol_type="genomic DNA"
/strain="MEX-15-099-01"
/isolation_source="blood"
/host="Bos taurus"
/db_xref="taxon:770"
/country="Mexico: Texcoco, Estado de Mexico"
/collection_date="2003-04-01"
WGS [VTWW01000001-VTWW01000032](#)
//


Apéndice 4

Genoma reportado:

Secuencia del genoma de *Anaplasma marginale* MEX-17-017-01
(Puente de Ixtla, Morelos) que está depositado en la base de datos GenBank

Anaplasma marginale strain MEX-17-017- 01, whole genome shotgun sequencing project

GenBank: VTCX00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

[Go to:](#) 

```
LOCUS       VTCX01000000          41 rc    DNA     linear   BCT 01-OCT-2019
DEFINITION  Anaplasma marginale strain MEX-17-017- 01, whole genome shotgun
            sequencing project.
ACCESSION   VTCX00000000
VERSION     VTCX00000000.1
DBLINK      BioProject: PRJNA562211
            BioSample: SAMN12636419
KEYWORDS    WGS.
SOURCE      Anaplasma marginale
            ORGANISM Anaplasma marginale
            Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
            Anaplasmataceae; Anaplasma.
REFERENCE   1 (bases 1 to 41)
AUTHORS     Quiroz Castaneda,R.E., Amaro Estrada,I., Martinez Ocampo,F. and
            Rodriguez Camarillo,S.D.
TITLE       Direct Submission
JOURNAL     Submitted (26-AUG-2019) Unidad de Anaplasmosis, CENID-SAI, INIFAP,
            Km. 11.5 de la Carretera Federal Cuernavaca-Cuatla, Col. Progreso,
            Jiutepec, Morelos 62550, Mexico
COMMENT     The Anaplasma marginale whole genome shotgun (WGS) project has the
            project accession VTCX00000000. This version of the project (01)
            has the accession number VTCX01000000, and consists of sequences
            VTCX01000001-VTCX01000041.
            Bacteria and source DNA available from Anaplasmosis Unit CENID-SAI,
            INIFAP.
            The annotation was added by the NCBI Prokaryotic Genome Annotation
            Pipeline (PGAP). Information about PGAP can be found here:
            https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/

##Genome-Assembly-Data-START##
Assembly Method      :: SPAdes v. 3.11.1
Genome Representation :: Full
Expected Final Version :: Yes
Genome Coverage      :: 58.0x
Sequencing Technology :: Illumina MiSeq
##Genome-Assembly-Data-END##

##Genome-Annotation-Data-START##
Annotation Provider   :: NCBI
Annotation Date       :: 08/27/2019 13:20:45
Annotation Pipeline   :: NCBI Prokaryotic Genome
                        Annotation Pipeline (PGAP)
Annotation Method     :: Best-placed reference protein
                        set; GeneMarkS-2+
Annotation Software revision :: 4.9
Features Annotated    :: Gene; CDS; rRNA; tRNA; ncRNA;
                        repeat_region
Genes (total)         :: 1,045
CDSs (total)          :: 1,001
Genes (coding)        :: 947
CDSs (with protein)   :: 947
Genes (RNA)           :: 44
rRNAs                 :: 1, 1, 1 (5S, 16S, 23S)
complete rRNAs       :: 1, 1, 1 (5S, 16S, 23S)
tRNAs                 :: 37
ncRNAs                :: 4
Pseudo Genes (total) :: 54
CDSs (without protein) :: 54
Pseudo Genes (ambiguous residues) :: 0 of 54
Pseudo Genes (frameshifted) :: 22 of 54
Pseudo Genes (incomplete) :: 31 of 54
Pseudo Genes (internal stop) :: 7 of 54
Pseudo Genes (multiple problems) :: 6 of 54
##Genome-Annotation-Data-END##

FEATURES             Location/Qualifiers
     source           1..41
                     /organism="Anaplasma marginale"
                     /mol_type="genomic DNA"
                     /strain="MEX-17-017- 01"
                     /isolation_source="infected blood"
                     /host="Bos taurus"
                     /db_xref="taxon:770"
                     /country="Mexico: Puente de Ixtla Morelos"
                     /collection_date="2009-09-01"
WGS                   VTCX01000001-VTCX01000041
//
```


Apéndice 5

Genoma reportado:

Secuencia del genoma de *Anaplasma marginale* MEX-30-184-02
(Tlapacoyan, Veracruz) que está depositado en la base de datos GenBank

Anaplasma marginale strain MEX-30-184-02, whole genome shotgun sequencing project

GenBank: VTCY00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

[Go to:](#) 

```
LOCUS       VTCY01000000          40 rc  DNA   linear   BCT 01-OCT-2019
DEFINITION  Anaplasma marginale strain MEX-30-184-02, whole genome shotgun
sequencing project.
ACCESSION   VTCY000000000
VERSION     VTCY000000000.1
DBLINK      BioProject: PRJNA562217
            BioSample: SAMN12636426
KEYWORDS    WGS.
SOURCE      Anaplasma marginale
            ORGANISM Anaplasma marginale
            Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
            Anaplasmataceae; Anaplasma.
REFERENCE   1 (bases 1 to 40)
AUTHORS     Amaro Estrada,I., Quiroz Castaneda,R.E., Martinez Ocampo,F. and
            Rodriguez Camarillo,S.D.
TITLE       Direct Submission
JOURNAL     Submitted (26-AUG-2019) Unidad de Anaplasmosis, CENID-SAI, INIFAP,
            Km. 11.5 de la Carretera Federal Cuernavaca-Cuautla, Col. Progreso,
            Jiutepec, Morelos 62550, Mexico
COMMENT     The Anaplasma marginale whole genome shotgun (WGS) project has the
            project accession VTCY000000000. This version of the project (01)
            has the accession number VTCY01000000, and consists of sequences
            VTCY01000001-VTCY01000040.
            Bacteria and source DNA available from Anaplasmosis Unit CENID-SAI,
            INIFAP.
            The annotation was added by the NCBI Prokaryotic Genome Annotation
            Pipeline (PGAP). Information about PGAP can be found here:
            https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/

            ##Genome-Assembly-Data-START##
            Assembly Date      :: 15-JUN-2019
            Assembly Method    :: SPAdes v. 3.11.1
            Genome Representation :: Full
            Expected Final Version :: Yes
            Genome Coverage    :: 27.0x
            Sequencing Technology :: Illumina MiSeq
            ##Genome-Assembly-Data-END##

            ##Genome-Annotation-Data-START##
            Annotation Provider      :: NCBI
            Annotation Date          :: 08/27/2019 13:21:58
            Annotation Pipeline      :: NCBI Prokaryotic Genome
            Annotation Pipeline (PGAP)
            Annotation Method        :: Best-placed reference protein
            set; GeneMarkS-2+
            Annotation Software revision :: 4.9
            Features Annotated       :: Gene; CDS; rRNA; tRNA; ncRNA;
            repeat_region
            Genes (total)             :: 1,050
            CDSs (total)              :: 1,006
            Genes (coding)            :: 952
            CDSs (with protein)       :: 952
            Genes (RNA)               :: 44
            rRNAs                    :: 1, 1, 1 (5S, 16S, 23S)
            complete rRNAs           :: 1, 1, 1 (5S, 16S, 23S)
            tRNAs                     :: 37
            ncRNAs                    :: 4
            Pseudo Genes (total)     :: 54
            CDSs (without protein)    :: 54
            Pseudo Genes (ambiguous residues) :: 0 of 54
            Pseudo Genes (frameshifted) :: 20 of 54
            Pseudo Genes (incomplete) :: 35 of 54
            Pseudo Genes (internal stop) :: 5 of 54
            Pseudo Genes (multiple problems) :: 6 of 54
            ##Genome-Annotation-Data-END##

FEATURES             Location/Qualifiers
     source           1..40
                     /organism="Anaplasma marginale"
                     /mol_type="genomic DNA"
                     /strain="MEX-30-184-02"
                     /isolation_source="infected blood"
                     /host="Bos taurus"
                     /db_xref="taxon:770"
                     /country="Mexico: Tlapacoyan Veracruz"
                     /collection_date="2014-02-10"
WGS                 VTCY01000001-VTCY01000040
//
```


Apéndice 6

Genoma reportado:

Secuencia del genoma de *Anaplasma marginale* MEX-30-193-01
(Veracruz, Veracruz) que está depositado en la base de datos GenBank

Anaplasma marginale strain MEX-30-193-01, whole genome shotgun sequencing project

GenBank: VTCZ00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

[Go to:](#) 

```
LOCUS       VTCZ01000000          41 rc  DNA   linear  BCT 01-OCT-2019
DEFINITION  Anaplasma marginale strain MEX-30-193-01, whole genome shotgun
            sequencing project.
ACCESSION   VTCZ00000000
VERSION     VTCZ00000000.1
DBLINK      BioProject: PRJNA562228
            BioSample: SAMN12636558
KEYWORDS    WGS.
SOURCE      Anaplasma marginale
            ORGANISM  Anaplasma marginale
            Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
            Anaplasmataceae; Anaplasma.
REFERENCE   1 (bases 1 to 41)
AUTHORS     Amaro Estrada,I., Quiroz Castaneda,R.E., Martinez Ocampo,F. and
            Rodriguez Camarillo,S.D.
TITLE       Direct Submission
JOURNAL     Submitted (26-AUG-2019) Unidad de Anaplasmosis, CENID-SAI, INIFAP,
            Km. 11.5 de la Carretera Federal Cuernavaca-Cuautla, Col. Progreso,
            Jiutepec, Morelos 62550, Mexico
COMMENT     The Anaplasma marginale whole genome shotgun (WGS) project has the
            project accession VTCZ00000000. This version of the project (01)
            has the accession number VTCZ01000000, and consists of sequences
            VTCZ01000001-VTCZ01000041.
            Bacteria and source DNA available from Anaplasmosis Unit CENID-SAI,
            INIFAP.
            The annotation was added by the NCBI Prokaryotic Genome Annotation
            Pipeline (PGAP). Information about PGAP can be found here:
            https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/

##Genome-Assembly-Data-START##
Assembly Date      :: 15-JUN-2019
Assembly Method    :: SPAdes v. 3.11.1
Genome Representation :: Full
Expected Final Version :: Yes
Genome Coverage    :: 154.0x
Sequencing Technology :: Illumina MiSeq
##Genome-Assembly-Data-END##

##Genome-Annotation-Data-START##
Annotation Provider      :: NCBI
Annotation Date          :: 08/27/2019 13:31:03
Annotation Pipeline      :: NCBI Prokaryotic Genome
                        Annotation Pipeline (PGAP)
Annotation Method        :: Best-placed reference protein
                        set; GeneMarkS-2+
Annotation Software revision :: 4.9
Features Annotated       :: Gene; CDS; rRNA; tRNA; ncRNA;
                        repeat_region
Genes (total)            :: 1,037
CDSs (total)             :: 993
Genes (coding)           :: 935
CDSs (with protein)      :: 935
Genes (RNA)              :: 44
rRNAs                   :: 1, 1, 1 (5S, 16S, 23S)
complete rRNAs           :: 1, 1, 1 (5S, 16S, 23S)
tRNAs                   :: 37
ncRNAs                  :: 4
Pseudo Genes (total)    :: 58
CDSs (without protein)  :: 58
Pseudo Genes (ambiguous residues) :: 0 of 58
Pseudo Genes (frameshifted) :: 18 of 58
Pseudo Genes (incomplete) :: 38 of 58
Pseudo Genes (internal stop) :: 5 of 58
Pseudo Genes (multiple problems) :: 3 of 58
##Genome-Annotation-Data-END##

FEATURES             Location/Qualifiers
     source           1..41
                     /organism="Anaplasma marginale"
                     /mol_type="genomic DNA"
                     /strain="MEX-30-193-01"
                     /isolation_source="infected blood"
                     /host="Bos taurus"
                     /db_xref="taxon:770"
                     /country="Mexico: Veracruz Veracruz"
                     /collection_date="2011-11-25"
WGS                   VTCZ01000001-VTCZ01000041
//
```


Apéndice 7

Genoma reportado:

Secuencia del genoma de *Anaplasma marginale* MEX-31-096-01
(Tizimín, Yucatán) que está depositado en la base de datos GenBank

Anaplasma marginale strain MEX-31-096-01, whole genome shotgun sequencing project

GenBank: VTWV00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

[Go to:](#) 

```
LOCUS       VTWV01000000          43 rc    DNA     linear   BCT 23-SEP-2019
DEFINITION  Anaplasma marginale strain MEX-31-096-01, whole genome shotgun
sequencing project.
ACCESSION   VTWV00000000
VERSION     VTWV00000000.1
DBLINK      BioProject: PRJNA563357
            BioSample: SAMN12672922
KEYWORDS    WGS.
SOURCE      Anaplasma marginale
  ORGANISM  Anaplasma marginale
            Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
            Anaplasmataceae; Anaplasma.
REFERENCE   1 (bases 1 to 43)
AUTHORS     Martinez-Ocampo,F., Quiroz-Castaneda,R.E., Amaro Estrada,I.,
            Rodriguez Camarillo,S.D., Dantan Gonzalez,E. and Preciado De La
            Torre,J.F.
TITLE       Direct Submission
JOURNAL     Submitted (03-SEP-2019) Unidad de Anaplasmosis, CENID-SAI, INIFAP,
            Km. 11.5 de la Carretera Federal Cuernavaca-Cuautla, Col. Progreso,
            Jiutepec, Morelos 62550, Mexico
COMMENT     The Anaplasma marginale whole genome shotgun (WGS) project has the
            project accession VTWV00000000. This version of the project (01)
            has the accession number VTWV01000000, and consists of sequences
            VTWV01000001-VTWV01000043.
            Bacteria and source DNA available from Anaplasmosis Unit CENID-SAI,
            INIFAP.
            The annotation was added by the NCBI Prokaryotic Genome Annotation
            Pipeline (PGAP). Information about PGAP can be found here:
            https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/

##Genome-Assembly-Data-START##
Assembly Date      :: 15-JUL-2019
Assembly Method    :: SPAdes v. 3.11.1
Genome Representation :: Full
Expected Final Version :: Yes
Genome Coverage    :: 19.0x
Sequencing Technology :: Illumina MiSeq
##Genome-Assembly-Data-END##

##Genome-Annotation-Data-START##
Annotation Provider      :: NCBI
Annotation Date          :: 09/05/2019 10:22:19
Annotation Pipeline      :: NCBI Prokaryotic Genome
                        Annotation Pipeline (PGAP)
Annotation Method        :: Best-placed reference protein
                        set; GeneMarkS-2+
Annotation Software revision :: 4.9
Features Annotated       :: Gene; CDS; rRNA; tRNA; ncRNA;
                        repeat_region
Genes (total)            :: 1,047
CDSs (total)             :: 1,003
Genes (coding)           :: 947
CDSs (with protein)      :: 947
Genes (RNA)              :: 44
rRNAs                   :: 1, 1, 1 (5S, 16S, 23S)
complete rRNAs          :: 1, 1, 1 (5S, 16S, 23S)
tRNAs                   :: 37
ncRNAs                  :: 4
Pseudo Genes (total)    :: 56
CDSs (without protein)  :: 56
Pseudo Genes (ambiguous residues) :: 0 of 56
Pseudo Genes (frameshifted) :: 18 of 56
Pseudo Genes (incomplete) :: 38 of 56
Pseudo Genes (internal stop) :: 6 of 56
Pseudo Genes (multiple problems) :: 6 of 56
##Genome-Annotation-Data-END##
FEATURES             Location/Qualifiers
     source            1..43
                        /organism="Anaplasma marginale"
                        /mol_type="genomic DNA"
                        /strain="MEX-31-096-01"
                        /isolation_source="blood"
                        /host="Bos taurus"
                        /db_xref="taxon:770"
                        /country="Mexico: Tizimin, Yucatan"
                        /collection_date="2010-06-22"
WGS                   VTWV01000001-VTWV01000043
//
```


Apéndice 8

Genoma reportado:

Secuencia del genoma de '*Candidatus Mycoplasma haemobos*' INIFAP01
que está depositado en la base de datos GenBank

Candidatus Mycoplasma haemobos strain INIFAP01, whole genome shotgun sequencing project

GenBank: LWUJ00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

[Go to:](#) 

```
LOCUS      LWUJ01000000          18 rc   DNA    linear   BCT 17-MAY-2016
DEFINITION Candidatus Mycoplasma haemobos strain INIFAP01, whole genome
            shotgun sequencing project.
ACCESSION  LWUJ00000000
VERSION   LWUJ00000000.1
DBLINK    BioProject: PRJNA319681
            BioSample: SAMN04903359
KEYWORDS   WGS.
SOURCE    Candidatus Mycoplasma haemobos
  ORGANISM Candidatus Mycoplasma haemobos
            Bacteria; Tenericutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE  1 (bases 1 to 18)
  AUTHORS  Quiroz-Castaneda,R.E. and Martinez-Ocampo,F.
  TITLE    Direct Submission
  JOURNAL  Submitted (26-APR-2016) Anaplasmosis, CENID-PAVET, INIFAP,
            Progreso, Coyoacan, Mexico D.F 04010, Mexico
COMMENT   The Candidatus Mycoplasma haemobos whole genome shotgun (WGS)
            project has the project accession LWUJ00000000. This version of
            the project (01) has the accession number LWUJ01000000, and
            consists of sequences LWUJ01000001-LWUJ01000018.
            Annotation was added by the NCBI Prokaryotic Genome Annotation
            Pipeline (released 2013). Information about the Pipeline can be
            found here: https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/
            ##Genome-Assembly-Data-START##
            Assembly Method      :: SPAdes v. 3.1.1
            Genome Representation :: Full
            Expected Final Version :: Yes
            Genome Coverage      :: 48.0x
            Sequencing Technology :: Illumina MiSeq
            ##Genome-Assembly-Data-END##

            ##Genome-Annotation-Data-START##
            Annotation Provider   :: NCBI
            Annotation Date      :: 04/27/2016 16:49:38
            Annotation Pipeline  :: NCBI Prokaryotic Genome
            Annotation Pipeline  :: Annotation Pipeline
            Annotation Method     :: Best-placed reference protein
            set; GeneMarkS+
            Annotation Software revision :: 3.1
            Features Annotated    :: Gene; CDS; rRNA; tRNA; ncRNA;
            repeat_region
            Genes (total)        :: 1,166
            CDS (total)          :: 1,129
            Genes (coding)       :: 1,109
            CDS (coding)         :: 1,109
            Genes (RNA)          :: 37
            rRNAs                 :: 1, 1, 1 (5S, 16S, 23S)
            complete rRNAs       :: 1, 1, 1 (5S, 16S, 23S)
            tRNAs                 :: 31
            ncRNAs                :: 3
            Pseudo Genes (total) :: 20
            Pseudo Genes (ambiguous residues) :: 0 of 20
            Pseudo Genes (frameshifted) :: 1 of 20
            Pseudo Genes (incomplete) :: 19 of 20
            Pseudo Genes (internal stop) :: 0 of 20
            ##Genome-Annotation-Data-END##
FEATURES             Location/Qualifiers
     source           1..18
                     /organism="Candidatus Mycoplasma haemobos"
                     /mol_type="genomic DNA"
                     /strain="INIFAP01"
                     /host="Bos taurus"
                     /db_xref="taxon:432608"
                     /country="Mexico: Chihuahua"
                     /collection_date="24-May-2008"
                     /collected_by="Anaplasmosis Unit, CENID-PAVET"
                     /identified_by="Fernando Martinez-Ocampo"
WGS                  LWUJ01000001-LWUJ01000018
//
```


Apéndice 9

Genoma reportado:

Secuencia del genoma de *Mycoplasma wenyonii* INIFAP02
que está depositado en la base de datos GenBank

Mycoplasma wenyonii strain INIFAP02, whole genome shotgun sequencing project

GenBank: QKVO00000000.1

 This entry is the master record for a whole genome shotgun sequencing project and contains no sequence data.

Go to: 

```
LOCUS       QKVO01000000                37 rc    DNA     linear   BCT 26-JUN-2018
DEFINITION  Mycoplasma wenyonii strain INIFAP02, whole genome shotgun
            sequencing project.
ACCESSION   QKVO00000000
VERSION     QKVO00000000.1
DBLINK      BioProject: PRJNA475119
            BioSample: SAMN09379922
KEYWORDS    WGS.
SOURCE      Mycoplasma wenyonii
  ORGANISM  Mycoplasma wenyonii
            Bacteria; Tenericutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 37)
  AUTHORS   Martinez Ocampo,F., Quiroz Castaneda,R.E. and Rojas Lopez,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2018) Anaplasmosis, CENID-PAVET, INIFAP,
            PROGRESO, COYOACAN, MEXICO D.F 04010, Mexico
COMMENT     The Mycoplasma wenyonii whole genome shotgun (WGS) project has the
            project accession QKVO00000000. This version of the project (01)
            has the accession number QKVO01000000, and consists of sequences
            QKVO01000001-QKVO01000037.
            Bacteria and source DNA available from Anaplasma Unit, CENID-PAVET,
            INIFAP.
            Annotation was added by the NCBI Prokaryotic Genome Annotation
            Pipeline (released 2013). Information about the Pipeline can be
            found here: https://www.ncbi.nlm.nih.gov/genome/annotation\_prok/
            ##Genome-Assembly-Data-START##
            Assembly Date          :: 15-MAY-2018
            Assembly Method        :: SPAdes v. 3.11.1
            Genome Representation  :: Full
            Expected Final Version :: Yes
            Genome Coverage        :: 7.0x
            Sequencing Technology  :: Illumina HiSeq
            ##Genome-Assembly-Data-END##

            ##Genome-Annotation-Data-START##
            Annotation Provider    :: NCBI
            Annotation Date        :: 06/12/2018 23:06:28
            Annotation Pipeline    :: NCBI Prokaryotic Genome
                                   Annotation Pipeline
            Annotation Method      :: Best-placed reference protein
                                   set; GeneMarkS+
            Annotation Software revision :: 4.5
            Features Annotated     :: Gene; CDS; rRNA; tRNA; ncRNA;
                                   repeat_region
            Genes (total)          :: 669
            CDS (total)            :: 634
            Genes (coding)         :: 629
            CDS (coding)           :: 629
            Genes (RNA)            :: 35
            rRNAs                  :: 1, 1, 1 (5S, 16S, 23S)
            complete rRNAs         :: 1, 1, 1 (5S, 16S, 23S)
            tRNAs                  :: 31
            ncRNAs                 :: 1
            Pseudo Genes (total)   :: 5
            Pseudo Genes (ambiguous residues) :: 0 of 5
            Pseudo Genes (frameshifted) :: 0 of 5
            Pseudo Genes (incomplete) :: 4 of 5
            Pseudo Genes (internal stop) :: 1 of 5
            ##Genome-Annotation-Data-END##
FEATURES             Location/Qualifiers
     source           1..37
                     /organism="Mycoplasma wenyonii"
                     /mol_type="genomic DNA"
                     /strain="INIFAP02"
                     /host="Bos taurus"
                     /db_xref="taxon:65123"
                     /country="Mexico: Morelos"
                     /collection_date="2015-06"
WGS                  QKVO01000001-QKVO01000037
//
```

Apéndice 10

Artículo publicado:

Quiroz Castañeda RE, Amaro-Estrada I, **Martínez-Ocampo F**, Rodríguez-Camarillo SD, Dantán González E, Cobáxin-Cárdenas M, Preciado-de la Torre JF. (2018) Draft genome sequence of *Anaplasma marginale* strain MEX-01-001-01, a Mexican strain that causes bovine anaplasmosis. Microbiol Resour Announc. 7(16):e01101-18.



Draft Genome Sequence of *Anaplasma marginale* Strain Mex-01-001-01, a Mexican Strain That Causes Bovine Anaplasmosis

Rosa Estela Quiroz Castañeda,^a Itzel Amaro-Estrada,^a Fernando Martínez-Ocampo,^b Sergio Rodríguez-Camarillo,^a Edgar Dantán González,^b Mayra Cobaxin-Cárdenas,^a Jesús Francisco Preciado-de la Torre^a

^aUnidad de Anaplasmosis, Centro Nacional de Investigación Disciplinaria en Parasitología Veterinaria, INIFAP, Carretera Federal Cuernavaca-Cuautla No. 8534, Jiutepec, Morelos, Mexico

^bLaboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, Mexico

ABSTRACT Bovine anaplasmosis is an arthropod-borne hemolytic disease caused by *Anaplasma marginale*. While only a few *Anaplasma marginale* strains have been reported, no Mexican strains have been reported. Due to the genetic diversity of *A. marginale*, the genome of the strain Mex-01-001-01, isolated in Mexico, represents a new source of information.

Bovine anaplasmosis is an arthropod-borne hemolytic disease caused by *Anaplasma marginale* (order *Rickettsiales*, family *Anaplasmataceae*), an obligate intraerythrocytic bacterium (1). Bovine anaplasmosis is endemic in Mexico and Central and South America and enzootic in most Latin American countries (2). Bovine anaplasmosis causes important economic losses, mainly due to high morbidity and mortality in cattle (3, 4). The genetic diversity of *A. marginale* has been reported in several countries, including Mexico (5–10).

Here, we report the draft genome of *A. marginale* strain Mex-01-001-01, isolated from blood of sick cattle from Aguascalientes, Mexico.

We used 200 μ l of bovine blood to extract genomic DNA using the UltraClean DNA BloodSpin kit (Mo Bio Laboratories). Two micrograms of genomic DNA was sequenced with the MiSeq system (Illumina) (Genomics Core of Arizona State University). We obtained a random data set of 1,100,300 paired-end reads of 300 bp which were reported in the SRA database.

The Illumina adapter sequences were removed from Illumina paired-end reads using the ILLUMINACLIP trimming step of Trimmomatic version 0.36, using the default settings (11). Low-quality bases were removed using the dynamictrim algorithm of SolexaQA++ version 3.1.7.1 (12), with a Phred quality score Q of 13. The resulting paired-end reads were *de novo* assembled using SPAdes version 3.11.1 (13) using the options (i) only runs assembly module (–only-assembler), (ii) reduce the number of mismatches (–careful), and (iii) k-mer lengths between 21 and 127. The contigs of *A. marginale* Mex-01-001-01 were differentiated from those contigs that belong to other organisms (i.e., contigs of bovine genomes) based on the GC content of each *A. marginale* contig using a Python script (https://github.com/FernandoMtzMx/GC_content_MultiFasta) (the *A. marginale* genomes reported have GC contents of 46% to 52%). Also, we aligned the sequences of each contig with the nr/nt database using BLASTN (14). Those contigs with an alignment coverage higher than 50% and identity higher than 70% corresponded to *A. marginale*.

The features of the draft genome were evaluated with QUAST version 4.6.2 using default settings (15).

The draft genome of *A. marginale* Mex-01-001-01 has 34 contigs with a total length of 1,179,425 bp and an N_{50} contig length of 65,428 bp, a GC content of 49.79%, and

Received 6 August 2018 Accepted 2 October 2018 Published 25 October 2018

Citation Quiroz Castañeda RE, Amaro-Estrada I, Martínez-Ocampo F, Rodríguez-Camarillo S, Dantán González E, Cobaxin-Cárdenas M, Preciado-de la Torre JF. 2018. Draft genome sequence of *Anaplasma marginale* strain Mex-01-001-01, a Mexican strain that causes bovine anaplasmosis. *Microbiol Resour Announc* 7:e01101-18. <https://doi.org/10.1128/MRA.01101-18>.

Editor J. Cameron Thrash, Louisiana State University

Copyright © 2018 Quiroz Castañeda et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rosa Estela Quiroz Castañeda, requiroz79@yahoo.com.mx. R.E.Q.C. and I.A.-E. contributed equally to this work.


Apéndice 11

Artículo publicado:

Martínez-Ocampo F, Quiroz-Castañeda RE, Amaro-Estrada I, Cobáxin-Cárdenas M, Dantán González E, Rodríguez-Camarillo SD. (2019) Draft genome sequences of *Anaplasma marginale* strains MEX-15-099-01 and MEX-31-096-01, Two Mexican Isolates with Different Degrees of Virulence. Microbiol Resour Announc. 8(45):e01184-19.



Draft Genome Sequences of *Anaplasma marginale* Strains MEX-15-099-01 and MEX-31-096-01, Two Mexican Isolates with Different Degrees of Virulence

Fernando Martínez-Ocampo,^a  Rosa Estela Quiroz-Castañeda,^b Itzel Amaro-Estrada,^b Mayra Cobaxin Cárdenas,^b  Edgar Dantán-González,^a Sergio Rodríguez-Camarillo^b

^aLaboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, Mexico

^bUnidad de Anaplasmosis, Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad, INIFAP, Jiutepec, Morelos, Mexico

ABSTRACT *Anaplasma marginale* is an intraerythrocytic bacterium that causes bovine anaplasmosis and is endemic in Mexico. In this work, we report two draft genome sequences of Mexican isolates from different geographical regions and with different degrees of virulence.

Anaplasma marginale is an intraerythrocytic rickettsial Gram-negative bacterium that causes clinical signs of bovine anaplasmosis, namely, fever, anemia, jaundice, weakness, and even respiratory distress. *A. marginale* is endemic in tropical and subtropical areas of the world, including Mexico (1). Only one genome sequence, that of the Mexican strain MEX-01-001-01 from Aguascalientes, has actually been reported (2). In 2000, García et al. evaluated the level of virulence of two Mexican strains, the high-virulence strain MEX-15-099-01 isolated from Texcoco, State of Mexico, and the low-virulence strain MEX-31-096-01 isolated from Tizimin, Yucatan (3). Also, in 2008, Rodríguez-Camarillo et al. reported that strain MEX-31-096-01 is used as a live vaccine (4). Here, we report the draft genome sequences of *A. marginale* strains MEX-15-099-01 and MEX-31-096-01, which were isolated from the blood of sick cattle and blood of asymptomatic cattle, respectively.

We used 200 μ l of bovine blood for each isolate to extract genomic DNA using the UltraClean DNA BloodSpin kit (Mo Bio Laboratories). Genomic DNA (2 μ g) for each isolate was sequenced using the MiSeq platform (Illumina), and the libraries were prepared by fragmenting the genomic DNA and ligating specialized adapters to both fragment ends (Arizona State University DNA Sequencing Core). We obtained two data sets of 605,370 (MEX-15-099-01) and 1,671,658 (MEX-31-096-01) 300-bp paired-end reads, which were reported to the SRA database.

The Illumina adapter sequences were removed from paired-end reads using the ILLUMINACLIP trimming step of Trimmomatic (version 0.36), with default settings (5). Low-quality bases were removed using the dynamictrim algorithm of the SolexaQA++ suite (version 3.1.7.1) (6) with a Phred quality score (Q) of <13. The resulting paired-end reads were *de novo* assembled using SPAdes (version 3.11.1) (7) with the following options: (i) only run the assembly module (-only-assembler), (ii) reduce the number of mismatches (-careful), and (iii) k-mer lengths between 21 and 127. Contigs of two Mexican strains were differentiated from contigs that belong to other organisms (i.e., bovine genomes) based on G+C content of each contig assessed using a Python script (https://github.com/FernandoMtzMx/GC_content_MultiFasta) (the *A. marginale* genomes reported in databases have a G+C content between 46 and 52%). Also, we aligned the sequences of each contig assembled with the NCBI Nucleotide (nr/nt) database using the BLASTN suite (8) and "*Anaplasma marginale*" as the organism name. Contigs with an alignment coverage of greater than 50% and an identity of greater than

Citation Martínez-Ocampo F, Quiroz-Castañeda RE, Amaro-Estrada I, Cobaxin Cárdenas M, Dantán-González E, Rodríguez-Camarillo S. 2019. Draft genome sequences of *Anaplasma marginale* strains MEX-15-099-01 and MEX-31-096-01, two Mexican isolates with different degrees of virulence. Microbiol Resour Announc 8:e01184-19. <https://doi.org/10.1128/MRA.01184-19>.

Editor Christina A. Cuomo, Broad Institute

Copyright © 2019 Martínez-Ocampo et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rosa Estela Quiroz-Castañeda, requiroz79@yahoo.com.mx.

Received 30 September 2019

Accepted 22 October 2019

Published 7 November 2019


Apéndice 12

Artículo publicado:

Martínez-Ocampo F, Quiroz-Castañeda RE, Amaro-Estrada I, Dantán-González E, Preciado de la Torre JF, Rodríguez-Camarillo SD. (2020) Whole-Genome Sequencing of Mexican Strains of *Anaplasma marginale*: An Approach to the Causal Agent of Bovine Anaplasmosis. Int J Genomics. 2020:5902029.

Research Article

Whole-Genome Sequencing of Mexican Strains of *Anaplasma marginale*: An Approach to the Causal Agent of Bovine Anaplasmosis

Fernando Martínez-Ocampo ¹, Rosa Estela Quiroz-Castañeda ², Itzel Amaro-Estrada ²,
Edgar Dantán-González ¹, Jesús Francisco Preciado de la Torre,²
and Sergio Rodríguez-Camarillo²

¹Laboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, C.P. 62209 Cuernavaca, Morelos, Mexico

²Unidad de Anaplasmosis, Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad, INIFAP, C.P. 62574 Jiutepec, Morelos, Mexico

Correspondence should be addressed to Rosa Estela Quiroz-Castañeda; requiroz79@yahoo.com.mx

Academic Editor: Marco Gerdol

Copyright © 2020 Fernando Martínez-Ocampo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Anaplasma marginale is the main etiologic agent of bovine anaplasmosis, and it is extensively distributed worldwide. We have previously reported the first genome sequence of a Mexican strain of *A. marginale* (Mex-01-001-01). In this work, we report the genomic analysis of one strain from Hidalgo (MEX-14-010-01), one from Morelos (MEX-17-017-01), and two strains from Veracruz (MEX-30-184-02 and MEX-30-193-01). We found that the genome average size is 1.16-1.17 Mbp with a GC content close to 49.80%. The genomic comparison reveals that most of the *A. marginale* genomes are highly conserved and the phylogeny showed that Mexican strains cluster with Brazilian strains. The genomic information contained in the four draft genomes of *A. marginale* from Mexico will contribute to understanding the molecular landscape of this pathogen.

1. Introduction

Bovine anaplasmosis is an infectious, tick-borne disease caused mainly by *Anaplasma marginale*; typical signs include anemia, fever, abortion, weight loss, decreased milk production, jaundice, and potentially death. Although a sick bovine may recover when antibiotics are administered, it usually remains as a carrier for life, being a risk of infection for susceptible cattle. *Anaplasma marginale* is an obligate intracellular Gram-negative bacterium with a genetic composition that is highly diverse among geographical isolates [1]. Currently, there are no fully effective vaccines against bovine anaplasmosis; therefore, the economic losses due to the disease are present. Whole-genome sequencing (WGS) is an applicable tool for many pathogenic bacterial studies since 1995, when the first bacterial genomes were determined [2, 3]. Vaccine

formulation became a hard task for pathogens as diverse as *Anaplasma marginale*, and almost all efforts have been directed toward Outer Membrane Proteins (Omp), Type IV Secretion System (T4SS), and Major Surface Proteins (Msp) [4–8]. Up to date, there are several genomes reported from *A. marginale*, but only one is from a Mexican strain [9]. New data could be useful for focusing in alternative antigens that induce specific and protective responses against bovine anaplasmosis. In this work, we present draft genomes from four *Anaplasma marginale* Mexican strains. In addition, a first approach for comparative analyses between them and Brazilian, Australian, and North American strains is shown. In order to advance in the identification of potential vaccine molecules, pathogenicity, transmission and infection mechanisms, and genetic diversity of *Anaplasma marginale*, further analyses are necessary.


Apéndice 13

Artículo publicado:

Rodríguez-Camarillo SD, Quiroz-Castañeda RE, Aguilar-Díaz H, Vara-Pastrana JE, Pescador-Pérez D, Amaro-Estrada I, **Martínez-Ocampo F.** (2020) Immunoinformatic Analysis to Identify Proteins to Be Used as Potential Targets to Control Bovine Anaplasmosis. Int J Microbiol. 2020:8882031.

Research Article

Immunoinformatic Analysis to Identify Proteins to Be Used as Potential Targets to Control Bovine Anaplasmosis

Sergio D. Rodríguez-Camarillo,¹ Rosa E. Quiroz-Castañeda ,¹ Hugo Aguilar-Díaz,² José E. Vara-Pastrana,³ Diego Pescador-Pérez,³ Itzel Amaro-Estrada,¹ and Fernando Martínez-Ocampo⁴

¹Unidad de Anaplasmosis, Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad, INIFAP, C.P. 62574, Jiutepec, Morelos, Mexico

²Unidad de Artropodología, Centro Nacional de Investigación Disciplinaria en Salud Animal e Inocuidad, INIFAP, C.P. 62574, Jiutepec, Morelos, Mexico

³Universidad Politécnica del Estado de Morelos, Paseo Cuauhnahuac 566, Lomas Del Texcal, C.P. 62574, Jiutepec, Morelos, Mexico

⁴Laboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, C.P. 62209, Cuernavaca, Morelos, Mexico

Correspondence should be addressed to Rosa E. Quiroz-Castañeda; requiroz79@yahoo.com.mx

Academic Editor: Todd R. Callaway

Copyright © 2020 Sergio D. Rodríguez-Camarillo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Omics sciences and new technologies to sequence full genomes provide valuable data that are revealed only after detailed bioinformatic analysis is performed. In this work, we analyzed the genomes of seven Mexican *Anaplasma marginale* strains and the data from a transcriptome analysis of the tick *Rhipicephalus microplus*. The aim of this analysis was to identify protein sequences with predicted features to be used as potential targets to control the bacteria or tick-vector transmission. We chose three amino acid sequences different to all proteins previously reported in *A. marginale* that have been used as potential vaccine candidates, and also, we report, for the first time, the presence of a peroxinectin protein sequence in the transcriptome of *R. microplus*, a protein associated with the immune response of ticks. The bioinformatics analyses revealed the presence of B-cell epitopes in all the amino acid sequences chosen, which opens the way for their likely use as single or arranged peptides to develop new strategies for the control and prevention of bovine anaplasmosis transmitted by ticks.

1. Introduction

Ticks and tick-borne pathogens constitute a major challenge for the cattle industry due to their impact on production losses [1]. Bovine anaplasmosis is a disease caused by *Anaplasma marginale*, an important tick-transmitted, intraerythrocytic Gram-negative bacterium, that is endemic in Mexico [2]. The disease has a worldwide distribution and causes serious economic losses, particularly in beef cattle, as they are more exposed to *A. marginale* transmitted by the tick-vector *Rhipicephalus microplus* [3]. The control of bovine anaplasmosis does not only depend on controlling the

pathogen itself, but also the vector which transmits it as they have coevolved with the host [4].

So far, there are no commercial vaccines against bovine anaplasmosis, and those that have been prepared from freed bacteria (initial bodies) have a limited use due to a wide antigenic diversity of the pathogen, while those prepared from live attenuated organisms carry the risk for the cotransmission of other blood-borne pathogens [2, 5]. With regards to tick vaccination, recent studies have shown that antitick recombinant vaccines such as aquaporin, subolesin, and tick gut glycoprotein Bm86 have a synergistic effect in reducing the engorgement of tick larvae *in vitro* [6, 7, 8]. So,

Apéndice 14

Artículo publicado:

Martínez-Ocampo F, Rodríguez-Camarillo SD, Amaro-Estrada I, Quiroz-Castañeda RE. (2016) Draft Genome Sequence of "*Candidatus* Mycoplasma haemobos," a Hemotropic Mycoplasma Identified in Cattle in Mexico. *Genome Announc.* 4(4):e00656-16.

Draft Genome Sequence of “*Candidatus Mycoplasma haemobos*,” a Hemotropic Mycoplasma Identified in Cattle in Mexico

Fernando Martínez-Ocampo, Sergio D. Rodríguez-Camarillo, Itzel Amaro-Estrada, Rosa Estela Quiroz-Castañeda

Unidad de Anaplasmosis, Centro Nacional de Investigación Disciplinaria en Parasitología Veterinaria, INIFAP, Jiutepec, Morelos, Mexico

We present here the draft genome sequence of the first “*Candidatus Mycoplasma haemobos*” strain found in cattle in Mexico. This hemotropic mycoplasma causes acute and chronic disease in animals. This genome is a starting point for studying the role of this mycoplasma in coinfections and synergistic mechanisms associated with the disease.

Received 17 May 2016 Accepted 19 May 2016 Published 7 July 2016

Citation Martínez-Ocampo F, Rodríguez-Camarillo SD, Amaro-Estrada I, Quiroz-Castañeda RE. 2016. Draft genome sequence of “*Candidatus Mycoplasma haemobos*,” a hemotropic mycoplasma identified in cattle in Mexico. *Genome Announc* 4(4):e00656-16. doi:10.1128/genomeA.00656-16.

Copyright © 2016 Martínez-Ocampo et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rosa Estela Quiroz-Castañeda, quiroz.rosa@inifap.gob.mx.

Hemotropic mycoplasmas (hemoplasmas) are the causative agent of infectious anemia in a variety of mammalian species (1). Bovine hemoplasmas are cell wall-less bacteria that can be found attached to surface-bound or free erythrocytes in the plasma of cattle (2). “*Candidatus Mycoplasma haemobos*” (syn. “*Candidatus Mycoplasma haemobovis*”) may cause anemia and depression to various degrees, edema, reproductive problems, and various other clinical signs in cattle (3). These bacteria have been reported in cattle from several countries around the world and may act in a synergistic manner when coinfections are present (4–8). In this study, we report the draft genome sequence of “*Candidatus Mycoplasma haemobos*” INIFAP01, an uncultivated hemoplasma found in blood of sick cattle from Chihuahua, Mexico. Genomic DNA was obtained using the Ultra-Clean DNA BloodSpin kit (Mo Bio Laboratories). Five micrograms of genomic DNA were sequenced in the HiSeq 2000 system (Illumina). We obtained a random data set of 45,165,900 paired-end reads of 300 bases long. Quality-based trimming was performed with a DynamicTrim (SolexaQA++) Perl script, and genome assembly was accomplished *de novo* using the SPAdes (version 3.1.1) program. The draft genome has 18 contigs, with a calculated 935,638-bp total length, N_{50} contig size of 256,799 bp, G+C content of 30.46% and $\sim 48\times$ coverage.

We identified the 16S rRNA gene of “*Ca. Mycoplasma haemobos*” INIFAP01 using the RNAmmer 1.2 server (<http://www.cbs.dtu.dk/services/RNAmmer/>) and compared it with 30 16S rRNA gene sequences of the genus *Mycoplasma* and two 16S rRNA genes of the genus *Ureaplasma* (outgroup). All sequences were aligned with the MUSCLE server (http://phylogeny.lirmm.fr/phylo.cgi/one_task.cgi?task_type=muscle), and a phylogenetic analysis was performed in the MEGA (version 6.1) program with a neighbor-joining algorithm using 1,000 replicates for bootstrapping. Phylogenetic analysis reveals that “*Ca. Mycoplasma haemobos*” INIFAP01 is closely related to the *Mycoplasma haemobos* species. The 16S rRNA gene of “*Ca. Mycoplasma haemobos*” INIFAP01 is 1,460 bp in length.

The 18 contigs were analyzed with the RAST server (<http://rast.nmpdr.org/>), identifying 1,216 open reading frames (ORFs) and

1,184 coding sequences (CDSs). This study was approved by the Centro Nacional de Investigación Disciplinaria en Parasitología Veterinaria (CENID-PAVET) branch of the Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP) Animal Experimentation and Ethics Committee and conducted considering ethic and methodological aspects, in agreement with the Mexican regulations related to use, housing, and transport of experimental animals, no. NOM-062-ZOO-1999.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. [LWUJ000000000](https://www.ncbi.nlm.nih.gov/nuclink/LWUJ000000000). The version described in this paper is version LWUJ01000000.

ACKNOWLEDGMENTS

We thank Luis Fernando Lozano-Aguirre Beltrán for bioinformatics support.

We declare no conflicts of interest regarding the publication of this paper.

FUNDING INFORMATION

This work, including the efforts of Itzel Amaro-Estrada, was funded by Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (1134353011). This work, including the efforts of Rosa Estela Quiroz-Castañeda, was funded by Consejo Nacional de Ciencia y Tecnología (CONACYT) (248855). This work, including the efforts of Sergio Dario Rodríguez-Camarillo, was funded by Consejo Nacional de Ciencia y Tecnología (CONACYT) (168167).

REFERENCES

- Willi B, Boretti FS, Tasker S, Meli ML, Wengi N, Reusch CE, Lutz H, Hofmann-Lehmann R. 2007. From haemobartonella to hemoplasma: molecular methods provide new insights. *Vet Microbiol* 125:197–209. [http://dx.doi.org/10.1016/j.vetmic.2007.06.027](https://doi.org/10.1016/j.vetmic.2007.06.027).
- Giroto A, Zangirólamo AF, Bogado ALG, Souza ASLE, da Silva GCF, Garcia JL, Vilas Boas LA, Biondo AW, Vidotto O. 2012. Molecular detection and occurrence of “*Candidatus Mycoplasma haemobos*” in dairy cattle of southern Brazil. *Rev Bras Parasitol Vet* 21:342–344. [http://dx.doi.org/10.1590/S1984-29612012000300034](https://doi.org/10.1590/S1984-29612012000300034).
- Hornok S, Micsutka A, Meli ML, Lutz H, Hofmann-Lehmann R. 2011.

Apéndice 15

Artículo publicado:

Quiroz-Castañeda RE, **Martínez-Ocampo F**, Dantán-González E. (2018) Draft Genome Sequence of *Mycoplasma wenyonii*, a Second Hemotropic Mycoplasma Species Identified in Mexican Bovine Cattle. Microbiol Resour Announc. 7(9):e00875-18



Draft Genome Sequence of *Mycoplasma wenyonii*, a Second Hemotropic *Mycoplasma* Species Identified in Mexican Bovine Cattle

Rosa Estela Quiroz-Castañeda,^a Fernando Martínez-Ocampo,^b Edgar Dantán-González^b

^aUnidad de Anaplasmosis, Centro Nacional de Investigación Disciplinaria en Parasitología Veterinaria, INIFAP, Jiutepec, Morelos, Mexico

^bLaboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, Mexico

ABSTRACT The hemotropic mycoplasma (hemoplasma) *Mycoplasma wenyonii* is an animal pathogen that affects bovine cattle health. Here, we present the draft genome sequence of the hemoplasma *M. wenyonii* strain INIFAP02 found in cattle from Mexico.

Hemotropic mycoplasmas (hemoplasmas) are epierythrocytic bacteria that have been reported in cattle from several countries and may act synergistically during a coinfection (1–5). Animals infected with hemoplasmas show clinical signs that include severe anemia, ill thrift, fever, lymphadenopathy, depression, decreased milk production, infertility, and reproductive inefficiency (3, 6, 7). To date, “*Candidatus Mycoplasma haemobos*” and *Mycoplasma wenyonii* are the only reported hemoplasmas that infect cattle. Previously, we reported the draft genome sequence of “*Ca. Mycoplasma haemobos*” strain INIFAP01, the first genome sequence of this hemoplasma (8). Here, we report the draft genome sequence of *M. wenyonii* strain INIFAP02, an uncultivated hemoplasma identified in the blood of sick cattle from Chihuahua, Mexico.

Genomic DNA was obtained from blood samples of sick cattle using the UltraClean DNA BloodSpin kit (Mo Bio Laboratories) following the manufacturer’s instructions, and 2 µg of the genomic DNA was sequenced using the Illumina MiSeq system (2 × 300-bp paired-end reads). We obtained a total data set of 605,370 paired-end reads, of which 13,678 paired-end reads belonged to the genome of *M. wenyonii* strain INIFAP02. The remaining paired-end reads belonged to the genomes of bovine and intraerythrocytic microorganisms (data not shown). Illumina adapter sequences were removed using the ILLUMINACLIP trimming step of the Trimmomatic version 0.36 program (9). Poor-quality bases from the Illumina paired-end reads were removed with the DynamicTrim algorithm in the SolexaQA++ software package (10) using a Phred quality threshold (Q) of 13. The paired-end reads were *de novo* assembled using the SPAdes version 3.11.1 program (11) with the following options: (i) only runs assembly module (–only-assembler), (ii) reduce the number of mismatches (–careful), and (iii) *k*-mer lengths between 21 to 127. Contigs not belonging to the genus *Mycoplasma* were eliminated using the nucleotide collection (NR/NT) database of the Nucleotide BLAST (BLASTN) server (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) (12). The draft genome sequence was processed using the QCAST version 4.6.2 program (13). The draft genome of *M. wenyonii* strain INIFAP02 has 37 contigs with a total length of 596,665 bp, an *N*₅₀ contig length of 35,044 bp, a G+C content of 33.43%, and a coverage of ~7×.

We obtained the 16S rRNA gene sequence of *M. wenyonii* strain INIFAP02 using the RNAmmer version 1.2 server (<http://www.cbs.dtu.dk/services/RNAmmer>) (14). The 16S rRNA gene has a length of 1,480 bp and an alignment coverage between 99 and 100%,

Received 4 July 2018 Accepted 8 August 2018 Published 6 September 2018

Citation Quiroz-Castañeda RE, Martínez-Ocampo F, Dantán-González E. 2018. Draft genome sequence of *Mycoplasma wenyonii*, a second hemotropic *Mycoplasma* species identified in Mexican bovine cattle. *Microbiol Resour Announc* 7:e00875-18. <https://doi.org/10.1128/MRA.00875-18>.

Editor Jason Stajich, University of California, Riverside

Copyright © 2018 Quiroz-Castañeda et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rosa Estela Quiroz-Castañeda, requiroz79@yahoo.com.mx.

R.E.Q.-C. and F.M.-O. contributed equally to this work.

Apéndice 16

Artículo publicado:

Popoca-Ursino EC, **Martínez-Ocampo F**, Dantán González E, Sánchez-Salinas E, Ortiz-Hernández ML. (2017) Characterization of methyl parathion degradation by a *Burkholderia zhejiangensis* strain, CEIB S4-3, isolated from agricultural soils. *Biodegradation*. 28(5-6):351-367.

Characterization of methyl parathion degradation by a *Burkholderia zhejiangensis* strain, CEIB S4-3, isolated from agricultural soils

Elida C. Popoca-Ursino · Fernando Martínez-Ocampo · Edgar Dantán-González · Enrique Sánchez-Salinas · Ma. Laura Ortiz-Hernández

Received: 28 April 2016 / Accepted: 4 July 2017
© Springer Science+Business Media B.V. 2017

Abstract Through the use of an enrichment technique, we isolated from the agricultural soils of Morelos in central México a strain of *Burkholderia zhejiangensis* identified as CEIB S4-3, it's could use the pesticide methyl parathion (MP) as the only source of carbon and degrade completely *p*-nitrophenol (PNP). For more efficient MP and PNP degradation by the CEIB S4-3 strain, the absence of an extra carbon source, a large inoculum and an MP concentration up to 50 mg/l are required. Sequence and annotation analysis of the draft genome, showed presence of *mpd* functional gene, which was expressed and its activity on the MP was confirmed. Additionally, the genes coding for enzymes in the benzoquinone pathway (conducted by Gram-negative bacteria) and the benzenetriol pathway (conducted by Gram-positive bacteria) were found, which was corroborated by identification of intermediary metabolites by HPLC. Thus, we propose that

B. zhejiangensis CEIB S4-3 uses both degradation pathways.

Keywords *Burkholderia zhejiangensis* CEIB S4-3 · Biodegradation · Methyl parathion · *p*-Nitrophenol

Introduction

Methyl parathion (*O,O*-dimethyl *O*-4 nitrophenyl phosphorothioate), an organophosphate pesticide, is an extremely toxic compound that is widely used in agriculture worldwide, and it is frequently found in terrestrial and aquatic environments (Papadakis et al. 2015). MP has low persistence in the environment, with reported field half-lives of 1 to 30 days in soils. However, the continuous and excessive application in addition to accidental spills allows further spread. In the environment, there are various abiotic factors that can hydrolyse MP, such as soil pH, temperature and sunlight exposure, but further degradation occurs via microbiota (Wauchope et al. 1992; Singh and Walker 2006). The physicochemical properties of MP make it susceptible to leaching and subsequent contamination of water bodies. On the other hand, because MP is metabolized by both plants and animals, it is not expected to persist or bioconcentrate (Bosco de Salles et al. 2015). However, studies have reported that *Girardinichthys multiradiatus* captured at the Ignacio Ramirez Dam, México, accumulated methyl parathion

Electronic supplementary material The online version of this article (doi:10.1007/s10532-017-9801-1) contains supplementary material, which is available to authorized users.

E. C. Popoca-Ursino · F. Martínez-Ocampo ·
E. Dantán-González · E. Sánchez-Salinas ·
Ma. L. Ortiz-Hernández (✉)
Laboratorio de Investigaciones Ambientales, Centro de
Investigación en Biotecnología, Universidad Autónoma
del Estado de Morelos, Av. Universidad 1001, Col.
Chamilpa, C.P. 62209 Cuernavaca, Morelos, Mexico
e-mail: ortizhl@uaem.mx

Published online: 11 July 2017

 Springer

Apéndice 17

Artículo publicado:

Salgado-Morales R, Rivera-Gómez N, **Martínez-Ocampo F**, Lozano-Aguirre Beltrán LF, Hernández-Mendoza A, Dantán González E. (2017) Draft Genome Sequence of *Photorhabdus luminescens* HIM3 Isolated from an Entomopathogenic Nematode in Agricultural Soils. Genome Announc. 5(35):e00745-17.



Draft Genome Sequence of *Photorhabdus luminescens* HIM3 Isolated from an Entomopathogenic Nematode in Agricultural Soils

Rosalba Salgado-Morales,^a Nancy Rivera-Gómez,^a Fernando Martínez-Ocampo,^a
Luis Fernando Lozano-Aguirre Beltrán,^b Armando Hernández-Mendoza,^c
Edgar Dantán-González^a

Laboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, Mexico^a; Centro de Ciencias Genómicas, Programa de Genómica Evolutiva, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico^b; Centro de Investigación en Dinámica Celular, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, Mexico^c

ABSTRACT In this work, we report the draft genome sequence of *Photorhabdus luminescens* strain HIM3, a symbiotic bacterium associated with the entomopathogenic nematode *Heterorhabditis indica* MOR03, isolated from soil sugarcane in Yauhtepec, Morelos, Mexico. These bacteria have a G+C content of 42.6% and genome size of 5.47 Mb.

The bacterium *Photorhabdus luminescens* is a gammaproteobacterium, is Gram negative, and is a member of the *Enterobacteriaceae* family; it is highly lethal for a wide spectrum of insects and maintains a symbiotic relationship with nematodes of the genus *Heterorhabditis* (1). The nematode carries the bacterium in the intestine during the juvenile infective stage (IJ), which penetrates an insect and regurgitates the symbionts in the hemolymph; these bacteria kill the insect in 24 to 48 h (2, 3). Therefore, *P. luminescens* is a good model for understanding the genetic bases involved in mutualism with nematodes and pathogenicity against insects, as well as being a source of insecticidal toxins (4, 5).

We sequenced the genome of the bacterium *P. luminescens* strain HIM3, isolated from the entomopathogenic nematode *Heterorhabditis indica* MOR03, recovered from soil cultivated with sugarcane in Yauhtepec, Morelos, Mexico (18°55'16.9"N, 99°02'24.0"W). The bacteria were grown in LB broth and incubated for 12 h at 30°C with shaking at 250 rpm; genomic DNA was extracted with the ZR fungal/bacterial miniprep kit (Zymo Research), and 5 µg of genomic DNA was sequenced using the Illumina HiSeq platform (2 × 300 bp paired-end reads). A total of 814,835 reads were obtained, and quality trimming and error correction of the reads were performed using a DynamicTrim (SolexaQA++) (6) Perl script. Reads were assembled using SPAdes version 3.5.0 (7), and the draft genome resulted in 91 contigs with a total length of 5,472,253 bp. The largest contig was 706,123 bp, the N_{50} value was 232,526 bp, and the L_{50} was reached in 8 contigs. The genome sequence of *P. luminescens* subsp. *laumondii* TTO1 was used as a reference. Automated annotation was performed with Rapid Annotations using Subsystems Technology (RAST) (8). The genome of *P. luminescens* HIM3 has a G+C content of 42.6% and includes 4,659 coding sequences; using RNAmmer and Aragorn, 10 rRNAs (5S, 16S, and 23S) and 70 tRNAs, respectively, were obtained (9, 10). Based on the analysis of concatenated sequences (*gyrA*, *gyrB*, *dnaN*, and 16S and 23S rRNA genes) *P. luminescens* HIM3 is closely related to *P. luminescens* DSM 3368 (GenBank accession no. NZ_JXSK00000000). The genome of *P. luminescens* HIM3 contains genes encoding the type III secretion system (T3SS) responsible for the injection of virulence proteins

Received 15 June 2017 Accepted 18 July 2017 Published 31 August 2017

Citation Salgado-Morales R, Rivera-Gómez N, Martínez-Ocampo F, Lozano-Aguirre Beltrán LF, Hernández-Mendoza A, Dantán-González E. 2017. Draft genome sequence of *Photorhabdus luminescens* HIM3 isolated from an entomopathogenic nematode in agricultural soils. *Genome Announc* 5:e00745-17. <https://doi.org/10.1128/genomeA.00745-17>.

Copyright © 2017 Salgado-Morales et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Edgar Dantán-González, edantan@uaem.mx.





Apéndice 18

Artículo publicado:

Salgado-Morales R, **Martínez-Ocampo F**, Obregón-Barboza V, Vilchis-Martínez K, Jiménez-Pérez A, Dantán González E. (2019) Assessing the Pathogenicity of Two Bacteria Isolated from the Entomopathogenic Nematode *Heterorhabditis indica* against *Galleria mellonella* and Some Pest Insects. *Insects*. 10(3):E83.

Article

Assessing the Pathogenicity of Two Bacteria Isolated from the Entomopathogenic Nematode *Heterorhabditis indica* against *Galleria mellonella* and Some Pest Insects

Rosalba Salgado-Morales ^{1,2}, Fernando Martínez-Ocampo ², Verónica Obregón-Barboza ², Kathia Vilchis-Martínez ³, Alfredo Jiménez-Pérez ³ and Edgar Dantán-González ^{2,*}

¹ Doctorado en Ciencias, Instituto de Investigación en Ciencias Básicas y Aplicadas, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Chamilpa, 62209 Cuernavaca, Morelos, Mexico; salgadomoralesr@hotmail.com

² Laboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Chamilpa, 62209 Cuernavaca, Morelos, Mexico; fernando.martinez@uaem.mx (F.M.-O.); veronica.obregon@uaem.mx (V.O.-B.)

³ Centro de Desarrollo de Productos Bióticos, Instituto Politécnico Nacional, Calle Ceprobi No. 8, San Isidro, Yauteppec, 62739 Morelos, Mexico; vilchisk78@gmail.com (K.V.-M.); aljimenez@ipn.mx (A.J.-P.)

* Correspondence: edantan@uaem.mx; Tel.: +52-777-329-7000

Received: 20 December 2018; Accepted: 15 March 2019; Published: 26 March 2019



Abstract: The entomopathogenic nematodes *Heterorhabditis* are parasites of insects and are associated with mutualist symbiosis enterobacteria of the genus *Photorhabdus*; these bacteria are lethal to their host insects. *Heterorhabditis indica* MOR03 was isolated from sugarcane soil in Morelos state, Mexico. The molecular identification of the nematode was confirmed using sequences of the ITS1-5.8S-ITS2 region and the D2/D3 expansion segment of the 28S rRNA gene. In addition, two bacteria HIM3 and NA04 strains were isolated from the entomopathogenic nematode. The genomes of both bacteria were sequenced and assembled de novo. Phylogenetic analysis was confirmed by concatenated gene sequence datasets as *Photorhabdus luminescens* HIM3 (16S rRNA, 23S rRNA, *dnaN*, *gyrA*, and *gyrB* genes) and *Pseudomonas aeruginosa* NA04 (16S rRNA, 23S rRNA and *gyrB* genes). *H. indica* MOR03 infects *Galleria mellonella*, *Tenebrio molitor*, *Heliothis subflexa*, and *Diatraea magnifactella* larvae with LC50 values of 1.4, 23.5, 13.7, and 21.7 IJs/cm², respectively, at 48 h. These bacteria are pathogenic to various insects and have high injectable insecticide activity at 24 h.

Keywords: entomopathogenic nematode; *Heterorhabditis indica*; bacteria; pathogenicity; insects

1. Introduction

The entomopathogenic nematodes (Rhabditida: Heterorhabditidae and Steinernematidae) are parasites of insects that are associated in mutualist symbiosis with enterobacteria of the genera *Photorhabdus* and *Xenorhabdus* [1,2]. The entomopathogenic nematodes, unlike other parasites, cause the death of the host insect in a short time (within 24–48 h of infection), and this lethality is provided by associated bacteria [2].

The infective juvenile stage (IJ) is the only free life form of the entomopathogenic nematodes and functions as a vector carrying the symbiotic bacteria in the intestine to release them in the host insect's hemolymph [3]. The life cycle of the nematode is completed inside the insect cadaver.

The symbiotic bacteria of entomopathogenic nematodes (*Photorhabdus* and *Xenorhabdus*) are gammaproteobacteria, Gram-negative, and members of the family Enterobacteriaceae. Canonically,

Apéndice 19

Capítulo de libro publicado:

Martínez-Ocampo F. (2018). Genomics of Apicomplexa. Farm Animals Diseases, Recent Omic Trends and New Strategies of Treatment. Quiroz Castañeda RE. IntechOpen.

Genomics of Apicomplexa

Fernando Martínez-Ocampo

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.72633>

Abstract

Apicomplexa is a eukaryotic phylum of intracellular parasites with more than 6000 species. Some of these single-celled parasites are important pathogens of livestock. At present, 128 genomes of phylum Apicomplexa have been reported in the GenBank database, of which 17 genomes belong to five genera that are pathogens of farm animals: *Babesia*, *Theileria*, *Eimeria*, *Neospora* and *Sarcocystis*. These 17 genomes are *Babesia bigemina* (five chromosomes), *Babesia divergens* (514 contigs) and *Babesia bovis* (four chromosomes and one apicoplast); *Theileria parva* (four chromosomes and one apicoplast), *Theileria annulata* (four chromosomes), *Theileria orientalis* (four chromosomes and one apicoplast) and *Theileria equi* (four chromosomes and one apicoplast); *Eimeria brunetti* (24,647 contigs), *Eimeria necatrix* (4667 contigs), *Eimeria tenella* (12,727 contigs), *Eimeria acervulina* (4947 contigs), *Eimeria maxima* (4570 contigs), *Eimeria mitis* (65,610 contigs) and *Eimeria praecox* (53,359 contigs); *Neospora caninum* (14 chromosomes); and *Sarcocystis neurona* strains SN1 (2862 contigs) and SN3 (3191 contigs). The study of these genomes allows us to understand their mechanisms of pathogenicity and identify genes that encode proteins as a possible vaccine antigen.

Keywords: Apicomplexa, genomics, parasitic protists, *Babesia*, *Theileria*, *Eimeria*, *Sarcocystis*, *Neospora*

1. Introduction

Apicomplexa (also called Apicomplexia) is a group of protists comprising a eukaryotic phylum of obligate intracellular parasites with more than 6000 described species [1]. Many of these cell single parasites are important pathogens of humans, domestic animals and livestock, with a health and economic relevance worldwide [2–5]. Apicomplexa microorganisms are intracellular eukaryotes thriving within another eukaryotic cell [6].

Apéndice 20

Artículo de divulgación científica publicado:

Hernández Mendoza A, **Martínez Ocampo F**, Dantán González E. (2017) La genómica ambiental, centinela del hábitat. HYPATIA. Revista No. 56.

La genómica ambiental, centinela del hábitat

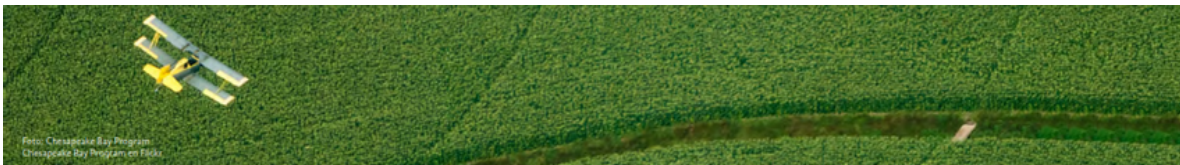


Los ecosistemas han sido contaminados como resultado de las actividades del ser humano. Por ejemplo, en los campos agrícolas se utilizan plaguicidas, productos químicos tóxicos para el ser humano y para otros seres vivos. Otro caso es la minería donde, desafortunadamente, tanto los principales productores de minerales como los productores secundarios, producen millones de toneladas de residuos con altos contenidos de elementos tóxicos que impactan sobre la salud humana y la de los ecosistemas (figura 1).

Por otro lado, en estas actividades la humanidad establece relaciones con los microbios, y no es que debamos preocuparnos porque nos vayan a contagiar, la razón principal es que ellos también ayudan a disminuir el impacto ambiental de las actividades industriales. Los microorganismos con estas capacidades, habitualmente forman comunidades que interactúan entre ellas; y estas asociaciones representan una fuente de herramientas moleculares que se puede explorar continuamente.

¿De dónde obtienen información estos organismos para llevar a cabo estas actividades?

La respuesta está escrita en su genoma. El genoma es todo el conjunto de genes contenidos en una molécula de ADN. El área que estudia a los genomas se le llama genómica. El tamaño de un genoma se mide en pares de bases (pb) y para representarlo se habla de kilobases (kb=1000 pb) o megabases (Mb=1,000,000 pb), que en el caso de las bacterias puede ir desde 500 kb hasta las 10 Mb. Los genes se copian a una molécula mensajera, el RNA, para que finalmente se fabriquen las proteínas que van a ser las responsables de usar y degradar aquellas moléculas contaminantes de nuestro ambiente.



Con todo lo anterior, surgen algunas interrogantes: ¿la genómica podrá ayudar a las ciencias ambientales para entender a estos organismos?, ¿será posible que con ella se puedan obtener y utilizar esas herramientas moleculares para contribuir en la mitigación de los daños ecológicos? La respuesta es contundente: ¡sí!, el impresionante mundo de la genómica cuenta con los elementos necesarios para estos retos.



Cuernavaca Morelos, a 14 de septiembre del 2021.

NOMBRAMIENTO DE COMISIÓN REVISORA DE TESIS DE GRADO

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Por este conducto me permito informarle que ha sido asignado integrante de la Comisión Revisora de tesis que presenta el **C. Fernando Martínez Ocampo**, titulada: **“Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”**, realizada bajo la dirección del DR. EDGAR DANTÁN GONZÁLEZ y la codirección de la DRA. ROSA ESTELA QUIROZ CASTAÑEDA, en el Programa Educativo de Doctorado en Ciencias Naturales.

Agradezco su valiosa participación en esta Comisión y quedo a sus órdenes para cualquier aclaración o duda al respecto.

ATENTAMENTE

“Por una humanidad culta”

DR. RAMÓN SUÁREZ RODRÍGUEZ
Coordinador del Doctorado En Ciencias Naturales



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

RAMON SUAREZ RODRIGUEZ | Fecha:2021-12-16 07:58:26 | Firmante

D9r6eW7r8j9uaAuslrwdDXV4VvNo+Dc0DjyvOYoQCJQUatcAhF2HrMy8H5vC4BcNjrqimnWAo+HI55Z1u7WYIz2mD3z1TyjLqSTVqXPtujqqHrW9UkkH9h7sEuGKLG/qBytN+7PLHUKVkv5hIMvNr3TQuAsudUrpTydDUgvgvghqIYjvwBM7Sr4heqQPpF48iiJQKBUG+YdrMHmi5LBGQ42rtbGPI3ieQyaDvaNCgOOiXsGfW6UaXag0ZefmtMinZAudYBeD9/P28lpmv9KEGK8FSrH307+AA4hNKivgWI2j3IACBoyNxAufSD6t3CXrbARyFpTNQj4rfWtIW2rkoRA==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:



RQKL6JuXV

<https://efirma.uaem.mx/noRepudio/E771cQTJHX3yjCv1xNbxFOB4Bmil1TI7>





UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



Cuernavaca, Morelos, a 14 de diciembre del 2021.

COMITÉ DE REVISIÓN DE TESIS

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Tesis: “Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”,

Alumno que lo presenta a revisión: FERNANDO MARTÍNEZ OCAMPO

Programa: DOCTORADO EN CIENCIAS NATURALES

VOTO

El documento ha sido revisado y reúne los requisitos para editarse como TESIS por lo que es

APROBADO

ATENTAMENTE

DR. EDGAR DANTÁN GONZÁLEZ



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

EDGAR DANTAN GONZALEZ | Fecha:2022-05-09 09:41:09 | Firmante

UjMVUJmOQInsoS4u/sxbNYUieT8uLned143LZEjV+d8QLTFGSLbD61J0iMKqCQG05By/esm+JJTsIG4dmHmT4Ksu3q/IX74gNahkHILrYIFPhQnluQi7+6v/qDrml/AAG51xpJH81iil
PrcFUgnXL6kzNhn8+q5urkaqPEzh0u36cpxzZxZ/hIR3Rpnzcow0lcnZ9dK1Ov1YK+vDsMh1k8hyo16vmc0dyShUAW4UuUwJ72Xj6yY7mycqwnH1E9THC22b2wy/ntslz3Z1G8ptPa
HIkK4aWIBqTG6lrMTSFiCogHyTDAF/+fuHTS4ROR0TdQRxBWBTPj8BsnWryJWPiyg==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o
escaneando el código QR ingresando la siguiente clave:



JTOt0F82b

<https://efirma.uaem.mx/noRepudio/rGDOOudkJNCON9QXX6tvRxcmcluXEFYQ>





Cuernavaca, Morelos, a 14 de diciembre del 2021.

COMITÉ DE REVISIÓN DE TESIS

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Tesis: “Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”,

Alumno que lo presenta a revisión: FERNANDO MARTÍNEZ OCAMPO

Programa: DOCTORADO EN CIENCIAS NATURALES

VOTO

El documento ha sido revisado y reúne los requisitos para editarse como TESIS por lo que es

APROBADO

ATENTAMENTE

DRA. ROSA ESTELA QUIROZ CASTAÑEDA



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

ROSA ESTELA QUIROZ CASTAÑEDA | Fecha:2022-06-08 22:20:58 | Firmante

cgVd1dLIYYPjEIKy9PFRXyISNTAog/Mor4vry6lOzCapltbdgmcl+jaCzIYS+HZr5EpPrXvBXmU/lay5EornIVrODNORFmwJHyE440kEWYvw5fXUdVwdy7ftvfHeo4mvBHtai5BJ34pby
p2DW3iHV1mn5m+aRoDQE+scdJn+dFGVI4XXnZOguwpxQFz4sy5RHIWYVPOIWb8oIRpKSMgY4azNSe/RcWCYc0oagXMHMWxQ1TUoErRz7JXObqCFTqrPgRbSPzakVIHQB
HFH/uycm8MfwrQjCMeqoYtfjG1XGwTTIZyiqFjnSK7pUNrmkqLjptoDubdvelO9TUNJ7cRw==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o
escaneando el código QR ingresando la siguiente clave:



61rCtR0FI

<https://efirma.uaem.mx/noRepudio/CG5Kx2WKCi56bCVcjQqZo661jOtAWPGm>



Cuernavaca, Morelos, a 14 de diciembre del 2021.

COMITÉ DE REVISIÓN DE TESIS

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Tesis: “Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”,

Alumno que lo presenta a revisión: FERNANDO MARTÍNEZ OCAMPO

Programa: DOCTORADO EN CIENCIAS NATURALES

VOTO

El documento ha sido revisado y reúne los requisitos para editarse como TESIS por lo que es

APROBADO

ATENTAMENTE

DR. RAMÓN SUÁREZ RODRÍGUEZ



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

RAMON SUAREZ RODRIGUEZ | Fecha:2022-05-05 17:08:15 | Firmante

i/fRmQcmIFq3UCbqZVnkcoeLUJ1jw3hhdpHk+QqfRP3VsUEW3+ebBrFu5tTQdR/plSOf+qi0UaiA0XWnEH3HCvfnULctxm566tRcv9UDFqETjm3iELoWb+vd92knaWo4VNd6WeT
v3iwGgLq79bTn0Ng7MRKDMclzf9zBCizLwgS70ro6eNeVjDRQ47kZT2zcZut8qj56nxdMDxAyn7WHyufJPHi86OaxdOINMa8ebXMIcXjf1Vk9gITC80KqpUb+dvavg0/k86OGZIVCf
Eetr9YFqYNIDL001wlvLqtq7GvwofkfadrGc+NWnx3hqetwE6rSRNWtCuevEINxD++1g==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o
escaneando el código QR ingresando la siguiente clave:



[gskREmd29](#)

<https://efirma.uaem.mx/noRepudio/J94fbCVsQjXK85o2JIQWfD7qcjt5WjCd>





UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



Cuernavaca, Morelos, a 14 de diciembre del 2021.

COMITÉ DE REVISIÓN DE TESIS

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Tesis: “Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”,

Alumno que lo presenta a revisión: FERNANDO MARTÍNEZ OCAMPO

Programa: DOCTORADO EN CIENCIAS NATURALES

VOTO

El documento ha sido revisado y reúne los requisitos para editarse como TESIS por lo que es

APROBADO

ATENTAMENTE

DR. ARMANDO HERNÁNDEZ MENDOZA



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

ARMANDO HERNANDEZ MENDOZA | Fecha:2022-05-05 10:55:42 | Firmante

CX3cpJk2tbvrrmFKuSYXVquXYgOqVP95Kg54/TVqUEYk6jufyR2TszYjWx2gr0+9YD+ddq4JYa+6wmmq5R6qwP2W5Diq031pm+mVXd9fDPNYrrUuBUoCoiWagMcZ7oa5wyEfv7e
ckDdS7p313iOz9pPhXfKkMtwJ6MvxIMhkGRbNKPEzh+jHQUPIC8P+yGqo0MZMWAasp1MTYwHQOtC4e48RAHICWSu/vzVpniLaNQiVni9clDvaJDrqNWKon6HHWSuSpQ3ud
Q/wKBpoYy16H+J3CbwG1nDobhdy9uTH3llzcOkgCyhp4hxERx0tqGTyqTpPzLEfs6mdPcjufZJzw==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o
escaneando el código QR ingresando la siguiente clave:



[hozMiA5Gc](#)

<https://efirma.uaem.mx/noRepudio/QpgaPMshGAOrL6oVZbUbnISNaYWhjwgl>



Cuernavaca, Morelos, a 14 de diciembre del 2021.

COMITÉ DE REVISIÓN DE TESIS

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Tesis: “Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”,

Alumno que lo presenta a revisión: FERNANDO MARTÍNEZ OCAMPO

Programa: DOCTORADO EN CIENCIAS NATURALES

VOTO

El documento ha sido revisado y reúne los requisitos para editarse como TESIS por lo que es

APROBADO

ATENTAMENTE

DR. JOSÉ AUGUSTO RAMÍREZ TRUJILLO



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

JOSE AUGUSTO RAMIREZ TRUJILLO | Fecha:2022-05-05 18:02:54 | Firmante

g21Ey8qK3fpUIKpEHc2rvMOTLqwm5UGbKrijQLDlvqxN8FzwCMAeJcwKYyMUTK2XPRuiP3FQJRqqj6VxVdAoMA/xqDn1/KGA+RMLJfgqIS9vyqcsHbksTgWFpHrGx8KU4A4uHs1lLe+ZtxYYTiLYoOu6WEE1F1sV4sXaJn9LijEJ9XXiVfxpzuLyR3orov5quKiXyurmRc05ltgP2SDhIUGTft2Qs8L0OfA42/2fL+/HaHBYSa1jWzqgsI4m57nPNBu17B8qGPM0eOclCmPHt1vAQpm2PfvRCwEFaiPMw0ISgJE07+7xRbc5UxDh6kiQoMvTwsbyDYcYGx/WdLGiLXA==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:



[W9P58iANw](#)

<https://efirma.uaem.mx/noRepudio/UJnPc5mMiyjsCQh76tXcE0kiU9Uvj2B3>





UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



Cuernavaca, Morelos, a 14 de diciembre del 2021.

COMITÉ DE REVISIÓN DE TESIS

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Tesis: “Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”,

Alumno que lo presenta a revisión: FERNANDO MARTÍNEZ OCAMPO

Programa: DOCTORADO EN CIENCIAS NATURALES

VOTO

El documento ha sido revisado y reúne los requisitos para editarse como TESIS por lo que es

APROBADO

ATENTAMENTE

DRA. MARÍA DEL RAYO SÁNCHEZ CARBENTE



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

MARIA DEL RAYO SANCHEZ CARBENTE | Fecha:2022-05-05 12:03:47 | Firmante

uyoUCbjoFgel+PyArYiHzOP/ouHRUYkByU6bbU6AsMW+0rpnth/15EoHtCa86Ytu2VURxJ5zYXuvxcJtaelw6G8CeKDzr/wqXMX8z8DQcKzp4njLbbsZdA8bd09r5xDbcPsvRZDI2UI
mt+ak6QIm1Qbkohy/Yohms98+hw8CotzNyOLcf9e5wXJb0E1VpITZvbaMd6rBfpENyYwT9g70clHvK0FqhICutNz5aYtdHAI/CM0ILpXPezaHilDqPzJF4tOrcX0I5YKa84yY686Bb4cZ
DZNV6YHI/gIKwrunCvInNy6ObcF4JwbB84Qd5OisqL0kARztS4OVES3CivDtw==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o
escaneando el código QR ingresando la siguiente clave:



[tvKiTrc3h](#)

<https://efirma.uaem.mx/noRepudio/dYBHd6sBjxJ3ehlviilTc9oobh3B0HCa>





UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



Cuernavaca, Morelos, a 14 de diciembre del 2021.

COMITÉ DE REVISIÓN DE TESIS

Dr. Edgar Dantán González (director de tesis, CEIB UAEM)
Dra. Rosa Estela Quiroz Castañeda (codirector de tesis, CNIDPV INIFAP)
Dr. Ramón Suárez Rodríguez (CEIB UAEM)
Dr. Armando Hernández Mendoza (CIDC UAEM)
Dr. José Augusto Ramírez Trujillo (CEIB UAEM)
Dra. María del Rayo Sánchez Carbente (CEIB UAEM)
Dra. Itzel Amaro Estrada (CENID_SAI INIFAP)

Tesis: “Genómica comparativa de nueve cepas del patógeno *Anaplasma marginale* e identificación de posibles candidatos vacunales contra la anaplasmosis bovina en México”,

Alumno que lo presenta a revisión: FERNANDO MARTÍNEZ OCAMPO

Programa: DOCTORADO EN CIENCIAS NATURALES

VOTO

El documento ha sido revisado y reúne los requisitos para editarse como TESIS por lo que es

APROBADO

ATENTAMENTE

DRA. ITZEL AMARO ESTRADA



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

ITZEL AMARO ESTRADA | Fecha:2022-05-06 13:27:37 | Firmante

dRSKKhG4IopNJppoAmaq7RM2YDQ9gCFd78qjcVnGlaDeFzc6s9z1eyEN+IorbEa32fgnzTc5DP1YdmvHAeyL6XzSjgHALtaD89gY+FVCmJp4WhihsnwgQruX1RtBXDvWroqOR7S0S+xNs8lC3okimfkaPe8oD3rujeUdry1zRqTDJmzqaQZ09YDcX9Snto9xHGzdiwo3uzi5zomOK4Po4YL1PqXAUJAVAKuwgWTZD3QPiyUQ5YcaxAk09g1vUpnqg02e5lbXANM8F5+PLJuftqF4TvJGivrzXCDICoJzWOceW2H0scN+5GkZ05eKoavd2sCTusSZyN3HZYH7xWN8hg==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:



[qevM03FkU](#)

<https://efirma.uaem.mx/noRepudio/E2dhUHPTCWgHO842XzXYBsSGAZ4ahHFH>

