



**UNIVERSIDAD NACIONAL
AUTÓNOMA
DE MÉXICO**



**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS DE
LA SALUD Y LA PRODUCCIÓN ANIMAL
FACULTAD DE MEDICINA VETERINARIA Y ZOOTECNIA**

Análisis de la cinética de crecimiento y genómica comparativa de *Brucella melitensis* 133 y su mutante atenuada LVM31

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS DE LA PRODUCCIÓN Y LA SALUD ANIMAL**

P R E S E N T A:

JULIO MANUEL MÉNDEZ ALEMÁN

Tutor:

José Luis Puente García

Departamento de Microbiología Molecular, Instituto de Biotecnología, UNAM

Comité tutor:

Jorge Ramírez Salcedo

Unidad de Microarreglos de DNA, Instituto de Fisiología Celular, UNAM

Antonio Verdugo Rodríguez

Laboratorio de Microbiología Molecular, Departamento de Microbiología e
Inmunología, Facultad de Medicina Veterinaria y Zootecnia, UNAM

Ciudad Universitaria, Cd. Mx. Noviembre 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*“El hombre del conocimiento no solo
tiene que poder amar a sus
enemigos tiene también que poder
odiar a sus amigos. Puesto que se
recompensa mal a un maestro si se
permanece siempre discípulo.
Ahora les ordeno que me pierdan a
mí y se encuentren a ustedes
mismos ...”*

Así habló Zaratustra

Nietzsche

Dedicatoria

A mi madre y mi padre que me han apoyado a lo largo de estos años, aunque hemos pasado momentos difíciles a lo largo de estos años siempre han estado ahí para ser la piedra angular más importante en todos los procesos de la casa. Los quiero mucho y siempre tratare de ser mejor para que estén orgullosos del hijo que les ha tocado criar.

A mi hermano Gracias por ayudarme y escucharme en estos años, apoyarme a cuidar a nuestras chaparras Kohle, Kuma y Carlota. Por escucharme con mis problemas en el laboratorio y con el trabajo de bioinformática.

A los demás miembros de mi familia que siempre han confiado y creído en mi trabajo.

A Tormenta, Thor, Maggie y Kiba que por ellos estoy aquí

Agradecimientos

Al Departamento de Microbiología e Inmunología del que he recibido muchas enseñanzas, así como un segundo hogar de desarrollo personal.

A mi comité tutor por todas sus aportaciones al trabajo, así como su paciencia y tolerancia para hacerme ver mis errores y aciertos. Muchas gracias por todas sus enseñanzas al Dr. Jorge Ramírez Salcedo que siempre me brindo consejo y del que aprendí otras técnicas que no conocía; Dr. José Luis Puente García por todo su apoyo técnico y académico, además de todo lo que me ayudó a que la tesis se llevara a buen puerto; y al Dr. Antonio Verdugo Rodríguez que tiene años soportándome, guiándome y dándome diferentes oportunidades de crecimiento personal en el laboratorio.

Al Dr. Simón del Departamento de Microarreglos del Instituto de Fisiología Celular por todos sus consejos para poder trabajar con el de RNA.

Gracias Cloé, mi hermanita de maestría por todos los momentos que tuvimos que pasar en este camino gracias por la ayuda y la compañía.

Gracias Yolotzin por escucharme y por apoyarme con el trabajo, aunque fuera solo con pequeños detalles.

A todos los miembros del Laboratorio Microbiología Molecular que siempre me acompañaron a reír, jugar, mejorar y encontrar una nueva familia en la que se puede confiar gracias a Dalia, Liliana, Andrea, Isaac, Salvador, Carlos, Krez, Leonardo, Adolfo, Jocelyn, Saúl. Gracias a Pao aunque nos hemos conocimos poco tiempo has sido una agradable persona con quien contar. Gracias Dani por ser el espejo más nítido y la máscara más transparente, gracias por dejarme salir del espejo. Así como a Mariyol por el tiempo que nos ayudó en el laboratorio y a Minerva que también convivió con nosotros, aunque fuese por poco tiempo. Todos hicieron de la estancia en el laboratorio un lugar agradable para convivir.

Resumen

La brucelosis es una enfermedad zoonótica causada por el género *Brucella sp.*, que se transmite al humano mediante el contacto directo de secreciones y productos del aborto, o por el consumo de productos de origen animal contaminados. *Brucella melitensis* es la responsable de los cuadros clínicos más severos en humanos. Se ha reportado que entre la cepa 16M y las cepas vacunales M5 y REV-1, hay menos del 1% de diferencia a nivel genómico, lo cual parece estar asociado a diferencias fenotípicas que presentan estas cepas en experimentos de infección *in vitro* e *in vivo*. La cepa de referencia mexicana Bm133, que ha sido usada en diferentes estudios y es la cepa parental de la mutante LVM31, la cual ha mostrado una atenuación similar a la REV-1, no ha sido secuenciada a la fecha. Con la finalidad de conocer mejor a la cepa Bm133, en este trabajo se realizó un estudio de genómica comparativa entre las cepas de *B. melitensis* 16M y Bm133. El análisis y comparación de la secuencia de la cepa Bm133 con la cepa 16M arrojó la siguiente información. El cromosoma I tiene una completitud del 99.1 % y consiste de 1993 CDS completas y 66 CDS rotas, presenta 209 SNP's y 80 gaps. Por su parte, el cromosoma II tiene una completitud de 99.6 % y consiste en 1092 CDS completas y 47 CDS rotas, y posee 160 SNP's y 48 gaps. Entre los genes interrumpidos se encontraron algunos cuyos productos estarían asociados con la envoltura celular, transportadores tipo ABC, factores de virulencia, metabolismo de carbohidratos, replicación y crecimiento celular. Resaltan los cambios encontrados en los genes *parB* y *groL* que previamente habían sido reportados para la cepa vacunal REV-1 e involucrados en la atenuación de la cepa REV-1. A su vez, se comparó las cinéticas de crecimiento de las cepas 16M, Bm133 y LVM31 en caldo Brucella, lo cual reveló que las tres alcanzan la fase logarítmica tardía de crecimiento a diferentes horas: a las 16 h la cepa 16M, a las 18 h la cepa Bm133 y hasta las 28 h la cepa LVM31. Como parte de estos estudios se estandarizó el método de fenol caliente para la óptima extracción de RNA de *B. melitensis* para estudios futuros de transcriptómica. Por otra parte, se estandarizó una RT-PCR para el gen *omp31* que permite detectar el transcrito a una concentración mínima de 207.7 pg/μl de RNA total, dicha concentración de RNA es equivalente a unas 2077 bacterias aproximadamente.

Palabras clave: Brucelosis, *melitensis*, bioinformática, genómico, ensamble

Abstract

Brucellosis is a zoonotic disease caused by the genus *Brucella* sp. transmitted to humans through direct contact with secretions and abortion products or by the consumption of contaminated animal products. *Brucella melitensis* is the species that causes the most severe clinical presentations in humans. Less than 1% difference has been reported at the genomic level between the 16M strain and the M5 and REV-1 vaccine strains. Nevertheless, these are sufficient to modify expression patterns in infection experiments *in vitro* and *in vivo*. The genome of the Mexican reference strain Bm133 has not been characterized, despite being used in previous reports and being the parental strain of the LVM31 mutant, which has shown similar attenuation as REV-1. Therefore, sequencing and comparative genomics analyses are important to determine the differences between the Mexican reference strain Bm133 and the world reference 16M. Sequence analysis of strain Bm133 and comparison with that of strain 16M yielded the following information. Chromosome I is 99.1% complete and consists of 1993 CDS and 66 interrupted CDS; it has 290 SNPs and 80 gaps. Chromosome II has a completeness of 99.6% and consists of 1092 CDSs, 47 interrupted CDSs, 160 SNPs and 48 gaps. Among the interrupted genes, some whose products are associated with the cell envelope, ABC-type transporters, virulence factors, carbohydrate metabolism, cell replication and growth were found. The changes seen in the *parB* and *groL* genes stand out, as they have previously been reported for the REV-1 vaccine strain and are involved in its attenuation.

Furthermore, the growth kinetics of strains 16M, Bm133 and LVM3 were compared at 24 and 36 hours in Brucella broth, which revealed that the three reached the late logarithmic phase at different times: at 16 hours, the strain 16M, at 18 h the strain Bm133 and up to 28 h the strain LVM31. This difference in growth could be related to some of the genomic changes observed. As part of these studies, the hot phenol method for optimal RNA extraction from *B. melitensis* was standardized for future transcriptomics studies. Likewise, it was possible to standardize an RT-PCR for the *omp31* gene, allowing us to detect in a minimum of total RNA concentration of 207.7 pg/ μ l, equivalent to approximately 2077 bacteria.

Índice

1	Introducción.....	1
1.1	Epidemiología.....	1
1.2	Brucelosis.....	2
1.3	Brucella sp.....	3
1.4	<i>B. melitensis</i>	3
1.5	Patogenia.....	4
1.6	Transcriptómica de <i>B. melitensis</i> 16M.....	5
1.7	Análisis genómico de cepas vacunales.....	7
1.7.1	M5.....	7
1.7.2	REV-1.....	9
1.8	Cepas de <i>B. melitensis</i> mexicanas.....	10
1.8.1	La cepa de <i>B. melitensis</i> BM133.....	10
1.8.2	La cepa de <i>B. melitensis</i> LVM31.....	11
2	Justificación.....	12
3	Hipótesis.....	13
4	Objetivo General.....	13
4.1	Objetivos Específicos.....	13
5	Materiales y Métodos.....	14
5.1	Cepas y cultivos.....	14
5.2	Cinética de crecimiento de <i>B. melitensis</i> 16M, Bm133 y LVM31.....	14
5.3	Extracción de RNA bacteriano.....	15
5.4	Ensayos con DNAsa I.....	16
5.5	RT-PCR del RNA total de <i>B. melitensis</i>	17
5.6	Extracción de DNA.....	19
5.7	Secuenciación de las cepas Bm133 y LVM31.....	21
5.8	Ensamblaje de los genomas Bm133 y LVM31.....	22
5.9	Análisis de calidad de las lecturas de la cepa Bm133 y la mutante LVM31.....	23
5.10	Edición de lecturas “reads” de la cepa Bm133.....	23
5.11	Ensamble de novo del genoma de la cepa Bm133 y la mutante LVM31.....	24
5.12	Análisis de los resultados del ensamblaje de novo.....	25

5.13	Mejoramiento del ensamblaje mediante el uso de SSPACE	26
5.14	Ordenamiento y alineamiento de “contigs” de la cepa Bm133	26
5.15	Anotación de los genes de la cepa de referencia mexicana de <i>B. melitensis</i> Bm133.	27
6	Resultados.....	28
6.1	Cinéticas de crecimiento de las cepas de <i>B. melitensis</i> 16M, Bm133 y LVM31.....	28
6.2	Extracción de RNA total de las cepas 16M, Bm133 y LVM31	29
6.3	Ensamblaje Bm133 y LVM31.....	34
6.4	Análisis de calidad	35
6.5	Edición de lecturas de la cepa Bm133	39
6.6	Ensamble <i>de novo</i> del genoma de la cepa Bm133.....	41
6.7	Análisis de los resultados del Ensamblaje <i>de novo</i>	42
6.8	Mejoramiento del ensamblaje mediante el uso de SSPACE.....	47
6.9	Ordenamiento y alineamiento de “contigs” de la cepa Bm133	47
6.10	Anotación de los genes de la cepa de referencia mexicana de <i>B. melitensis</i> Bm133.	58
7	Discusión	62
8	Conclusiones	68
9	Prospectiva.....	69
10	Bibliografía	70
11	Apendices.....	75
11.1	Diagrama de flujo del ensamble	75
11.2	Lista de Programas usados.....	77

Lista de Cuadros

Cuadro 1.	Protocolo de reacción con DNAsa I para RNA Total	17
Cuadro 2.	PCR omp22.....	18
Cuadro 3.	PCR omp31.....	19
Cuadro 4.	Archivos que contienen las lecturas obtenidas en la secuenciación	22
Cuadro 5.	Archivos de “reads” editados con el programa Cutadapt 3.2.....	24
Cuadro 6.	Resumen de los valores obtenidos por el programa QAST para los diferentes ensamblajes que se realizaron.....	44
Cuadro 7.	CDS con sustituciones no sinonimas o mutaciones sin sentido del ChI de la cepa Bm133.	55
Cuadro 8.	CDS con sustituciones no sinónimas o mutaciones sin sentido del ChII de la cepa Bm133.	57

Lista de Figuras

Figura 1.	Cinética de crecimiento de las cepas 16M y Bm133.....	28
Figura 2.	Cinética de crecimiento de la cepa LVM31.....	29
Figura 3.	Extracción de RNA total de <i>B. melitensis</i> 16M, Bm133 y LVM31; y monocitos THP-1.	30
Figura 4.	RNA total tratado con DNAsa I.....	31
Figura 5.	RNA obtenido a partir de diluciones décuples seriadas de un cultivo de <i>B. melitensis</i> 16M.....	32
Figura 6.	RT-PCR semi-cuantitativo del gen <i>omp31</i> de la cepa 16M.	33
Figura 7.	Diagrama Kroona de la cepa <i>B. melitensis</i> Bm133 secuenciada.....	35
Figura 8.	Reporte de calidad combinado de las lecturas de la cepa Bm133 y de la mutante LVM31.	36
Figura 9.	Distribución de la longitud de las secuencias de las lecturas de las cepa Bm133 y LVM31.	37
Figura 10.	Contenido de N's en las secuencias de las lecturas para los 4 archivos.	37
Figura 11.	Resumen del análisis de calidad de las cepa Bm133 y LVM31.	38
Figura 12.	Reporte de calidad combinado de las lecturas de la cepa Bm133.	40
Figura 13.	Figura 13 Contenido de bases N en las lecturas.	40
Figura 14.	Mapa de frecuencias de los valores reportados en todos los analisis de FastQC de a cepa Bm133.....	41
Figura 15.	Gráfica de identidad del Cromosoma I (ChI) de Bm133 vs 16M.....	45
Figura 16.	Gráfica de identidad del Cromosoma II (ChII) de Bm133 vs 16M.....	46
Figura 17.	Representación del ordenamiento final de “contigs” del ChI con Mauve.....	48
Figura 18.	Representación del ordenamiento final de “contigs” del ChII con Mauve.....	49
Figura 19.	Gráfica de similitud del ensamble ordenado del ChI de Bm133 vs el ChI 16M. .	59
Figura 20.	Gráfica de similitud del ensamble ordenado del ChI de Bm133 vs el ChI 16M. .	60

Abreviaturas

µm	Micrómetros
kb	Kilobases
pH	Potencial de hidrogeniones
°C	Grados Celsius
h	Horas
µl	Microlitros
ml	Mililitros
rpm	Revoluciones por minuto
mg	Miligramos
v	Volumen

1 Introducción

1.1 Epidemiología

La brucelosis es una enfermedad infecto-contagiosa, causada por las bacterias pertenecientes al género *Brucella* sp. Es una zoonosis que se transmite por contacto directo con las secreciones y productos del aborto de un animal infectado o de manera indirecta por el consumo de productos de origen animal contaminados con la bacteria (Corbel, 2006). La brucelosis humana es una de las zoonosis más comunes en el mundo con más de 500,000 casos al año, que puede provocar discapacidad en las personas y estar asociada como una enfermedad en personas que se trasladan alrededor del mundo (Pappas *et al.*, 2006). La Organización Mundial de la Salud (WHO por sus siglas en inglés) reporta en promedio 393,239 casos de brucelosis de origen alimentario (WHO, 2014). Las regiones del mundo que más reportan casos de brucelosis humana son África, Asia y América del sur; en Europa, particularmente en países como España, Portugal y los países balcánicos, su incidencia es baja (Pappas *et al.*, 2006). De acuerdo con el boletín Epidemiológico emitido por el Sistema Nacional de Vigilancia Epidemiológica (SINAVE) de la Secretaría de Salud, México es el país de América del norte que reporta más casos de brucelosis en humanos con 2,447 casos en 2017 (Dirección de Campañas Zoonositarias and SENASICA, 2017) y 1,527 casos en 2018. Por esta razón, México mantiene una Campaña Nacional contra la Brucelosis en animales, la cual se encuentra bajo la legislación de la NOM-041-ZOO-1995, en la que se plantean las medidas de control, delimitación y segregación de zonas, vacunación obligatoria de animales y diagnóstico de animales, así como las pruebas diagnóstico de tipo inmunológicas (tarjeta, anillo en leche, rivanol, fijación de complemento e inmunodifusión doble) para la campaña, etc, (SAGARPA, 1995). El Servicio Nacional de Sanidad, Inocuidad y Calidad Agroalimentaria (SENASICA) reporta, dentro del periodo de enero-septiembre del año 2017, un total de 762,255 cabezas de ganado

vacunadas incluyendo bovinos, caprinos y ovinos (Dirección de Campañas Zoonositarias and SENASICA, 2017), solo de los bovinos se reportan 401,408 cabezas vacunadas, mientras que en el inventario ganadero del 2014 se reportan 28,415,337 cabezas totales en el país, lo que representa menos del 1.5% de animales vacunados. Dentro de la campaña se reporta que para el año 2019, los estados de Sonora y Baja California Sur se consideran libres de brucelosis, en específico de *Brucellas lisas*; el 30.75% del territorio nacional se encuentra en etapa de erradicación y el resto del país se encuentra en control (Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria, 2016).

1.2 Brucelosis

La brucelosis, fue descrita por David Bruce en el año de 1887. *Brucella melitensis* agente causal de la fiebre de Malta afecta a humanos y rumiantes principalmente. Es una enfermedad debilitante que puede cursar por tres fases: la primera es la fase de incubación, que dura de 2-3 semanas, en la cual los síntomas no son evidentes en el hospedador; la segunda es una fase aguda en la cual el patógeno invade y coloniza diferentes tejidos; por último, la tercera fase o la fase crónica puede terminar con muerte del hospedador. Sus signos son mialgia, artralgia, pirexia, anorexia, fatiga, etc. La bacteria también se puede establecer en el aparato reproductor provocando abortos, distocia, retención placentaria y en ocasiones el nacimiento de animales débiles. Los problemas crónicos que se presentan en la brucelosis pueden ser hepáticos, respiratorios, cardíacos, neurálgicos, etc. También se sabe que la bacteria puede transmitirse de madre a hijo por la vía láctea, la cual es comúnmente reportada en animales, pero muy rara vez en humanos (Corbel, 2006; De Figueiredo *et al.*, 2015).

1.3 *Brucella* sp.

El género *Brucella* se encuentra dentro de la familia *Brucellaceae* en el orden de los Rhizobiales, dentro de la clase de las α -proteobacterias. Es un bacilo corto de 0.5-0.7 x 0.6-1.5 μm , Gram negativo, intracelular facultativo, inmóvil (a pesar de tener los genes del flagelo en su cromosoma) y no esporula. Positiva a las pruebas de catalasa, de oxidasa (aunque algunas biovariedades pueden ser negativas), negativas en la producción de indol y con capacidad desnitrificante. La bacteria puede crecer de los 20 °C hasta los 40°C (Garrity, 2005). El origen de las especies de *Brucella* se ha enfocado, principalmente, en la aparente adaptación hacia un hospedero específico, por lo que se ha sugerido que la mayoría de las especies evolucionaron de un ancestro común muy parecido a *Brucella ovis*. La evolución de cada especie no está ligada únicamente al hospedero preferencial, sino también a los procesos de adaptación (Ficht, 2011). Existen diez especies principales de *Brucella* que son: *melitensis*, *abortus*, *suis*, *ovis*, *canis*, *neotomae*, *pinnipedialis*, *ceti*, *microti* e *inopinata* (Golshani and Buozari, 2017). De estas especies, cuatro son las que representan un mayor riesgo a la salud pública *B. melitensis*, *B. abortus*, *B. suis* y *B. canis* (Corbel, 2006; Pappas *et al.*, 2006).

1.4 *B. melitensis*

Es la especie que origina los cuadros clínicos más severos en humanos. Sus hospederos preferenciales son las cabras y los ovinos. En México el 93% de los casos en humanos son causados por *B. melitensis* provenientes de cabras mientras que sólo el 5% son de origen bovino (González-Rodríguez, 2012).

1.5 Patogenia

La vía de transmisión de *B. melitensis* es principalmente de manera horizontal tanto en humanos como en cabras y borregos. Esto ocurre por el contacto del hospedero con secreciones orales, vaginales, productos del aborto y consumo de leche contaminada; además, el consumo de productos lácteos contaminados con la bacteria también es una manera de transmisión. En el caso de las cabras y borregos se ha demostrado que existe la transmisión genital. Se sabe que las bacterias del género *Brucella* tiene una marcada preferencia por el eritritol, un polialcohol producto intermedio en la vía de las pentosas que se utiliza como fuente de energía, al igual que la fructosa, producida y utilizada por algunos tejidos del aparato reproductor (Lopez-Goñi and Moriyón, 2005). Por lo anterior, la bacteria puede producir los siguientes signos y lesiones asociadas al aparato reproductor como son: aborto, retención placentaria, metritis e infertilidad. Aunque estos signos y lesiones suceden principalmente en los caprinos y ovinos, existen reportes en humano que presentan dicha signología (González-Rodríguez, 2012; Verdiguél-Fernández, 2014).

La bacteria entra en contacto con las mucosas del hospedero (oral, conjuntival y respiratorio), lo cual le permite su interacción con las células epiteliales y macrófagos del área. Aunque por principio se considera que la vía de entrada de *Brucella* es la interacción con fagocitos profesionales en las zonas de contacto inicial, varios reportes presentan datos interesantes sobre la entrada de la bacteria a través del aparato digestivo, principalmente por las placas de Peyer que se localizan en la parte terminal del íleon (Rossetti *et al.*, 2013; De Figueiredo *et al.*, 2015). La infección de macrófagos por parte de la bacteria permite su diseminación inicial por vías linfáticas hacia linfonodos, bazo, hígado, articulaciones, etc. La invasión de macrófagos por *B. melitensis* se suscita por diferentes vías, de éstas la más importante sucede mediante la interacción con los microdominios conocidos como balsas lipídicas (*lipids rafts* denominación en

inglés), en los cuales el LPS de la bacteria interactúa con receptores celulares que permiten su internación con los macrófagos y que dan origen a la vacuola contenedora de *Brucella* (BCV por sus siglas en inglés).

La interacción de esta BCV con diferentes proteínas encargadas del tránsito intracelular como Rab5, EEA1 y TfR se presenta antes de los 30 minutos, esto sucede mediante la interacción de la vacuola con los endosomas tempranos. La modificación de las rutas de tránsito por medio de sus factores de virulencia, en específico el sistema de secreción tipo 4 (SST4) y sus diferentes efectores, evita que la vacuola se fusione con lisosomas y por tanto evita la lisis bacteriana. Después de 30 minutos se encuentran proteínas de endosoma intermedio y tardío como LAMP1, CD63 y Rab7, que demuestran la maduración de la BCV. Posteriormente, se pueden encontrar los marcadores de endosoma maduro IRE1 α , Atg9 y Sar1. La BCV se localiza en los sitios ERES que son los sitios que permiten el tráfico de vesículas del retículo endoplásmico al aparato de Golgi, donde *Brucella* comienza su replicación intracelular. Por último, la BCV adquiere proteínas como Beclina 1, ULK1 y ATG14L, lo cual facilita la salida de la bacteria de la célula (Gorvel and Moreno, 2002; Celli, 2015).

1.6 Transcriptómica de *B. melitensis* 16M

Una bacteria intracelular facultativa como *Brucella* y que sobrevive a cambios tan adversos como los que representa el ambiente extracelular, así como los diferentes microambientes intracelulares, debe tener una intrincada red de efectores y factores que le ayuden a la expresión o represión de sus genes (Viadas *et al.*, 2010). La activación o represión diferentes factores de virulencia como son: los sistemas de dos componentes, el sistema de secreción tipo 4, los transportadores ABC, la expresión de glucanos cíclicos, etc.; ayuda a que la bacteria evite el sistema inmune del hospedador.

Rossetti *et al.* (2009) reportaron que *B. melitensis* 16M invade mejor en células HeLa a partir de cultivos en la fase logarítmica tardía con respecto a fases tempranas o estacionaria. Demostraron que en la fase logarítmica tardía se recuperaron 130×10^3 CFU/ml. En este mismo reporte, mediante el uso de microarreglos. Se encontraron 454 genes expresados diferencialmente (14% del genoma de *Brucella*) entre las fases logarítmica tardía y la estacionaria. De éstos 414 fueron sobre-expresados y 40 subexpresados en la fase logarítmica, siendo la mayoría genes asociados a crecimiento y metabolismo. De los genes asociados a virulencia sólo 3 genes del operón *virB* (*virB1*, *virB3* y *virB10*) fueron regulados positivamente, así como 5 genes que participan en la biogénesis del flagelo o su regulación. Ya que la regulación genética de las bacterias se encuentra ligada a los cambios ambientales, nutricionales o estreses a los que se enfrenta el microorganismo, 22 de los genes que codifican para reguladores transcripcionales pertenecientes a diferentes familias, tales como AraC, AsnC, BetI, DeoR, GntR, IclR, LysR, LuxR, MarR, MerR y RpiR, fueron expresados en forma diferencial en la fase logarítmica tardía con respecto a la fase estacionaria, al igual que los genes que codifican para las histidin cinasas VsrB y HprK (regulador central del metabolismo de carbohidratos). Además, se encontró que el gen *rpoH1* que codifica para el factor sigma 32 que regula la transcripción de genes de choque térmico, presentó un incremento en su expresión (Rossetti *et al.*, 2009).

Las propiedades de la membrana externa de *B. melitensis* se encuentran ligadas a la capacidad de la bacteria de invadir células por sitios específicos ricos en colesterol, por lo que cambios importantes en la estructura de la membrana externa de la bacteria pueden mitigar su capacidad para adherirse a las balsas lipídicas, que son el sitio diana celular que prefiere *Brucella* para iniciar su ingreso en las células.

Rossetti *et al.*, (2017) reportaron que la cepa 16M de *B. melitensis* comienza a transcribir genes después de los 30 minutos post infección en las placas de Peyer

en el íleon ligado de ternero. La cantidad de genes que se modificaron a las 2h y a las 4h post-infección fue de 1,909 genes y 1,912 genes, respectivamente. Los genes que la bacteria comienza a expresar después de 30 minutos pertenecen a los sistemas de dos componentes BvrS/R, FeuQP, NtrBC, NtrXY, entre otros; además, mantiene ciertos genes de metabolismo encendidos. Por el contrario, los transportadores tipos ABC y el SST 4 se mantienen apagados aún después de 4 horas, lo que sugiere una estrategia que permite el evadir la estimulación de la respuesta inmune innata.

1.7 Análisis genómico de cepas vacunales

1.7.1 M5

La cepa vacunal M5, derivada de la cepa virulenta de *B. melitensis* M28, es utilizada principalmente para vacunar ovejas y cabras gestantes y no gestantes en China. Presenta una estructura similar a la cepa de referencia 16M con sólo un 1% de diferencias genéticas con respecto a la cepa de referencia 16M de *B. melitensis* (Jiang *et al.*, 2013). Estas diferencias genómicas, representadas por la presencia de deleciones, inserciones, regiones del genoma únicas y los polimorfismos de nucleótidos, son suficientes para que la cepa M5 sea menos virulenta con respecto a la 16M. El cromosoma 2 de la cepa M5 presenta una inversión genética de 56 kb, asociada a transposasas. En esta inversión se encontraron genes pertenecientes a transportadores de tipo ABC para péptidos que pueden jugar un rol importante en la capacidad de sobrevivencia intracelular de la bacteria. También se identificaron 5 regiones de eliminaciones en el genoma de la M5 que en la 16M corresponden a genes con funciones de chaperonas (ClpB), helicasas dependientes de ATP, proteínas hipotéticas pertenecientes a la familia de las resolvasas y genes relacionados a amidasas y aminotransferasas. A su vez, se encontraron 2,819 polimorfismos de un solo nucleótido de los cuales 585 fueron en regiones intergénicas y 2,234 ocurrieron

en regiones codificantes. De los 2,234 SNP's (polimorfismos simples de nucleótido por sus siglas en inglés), 798 fueron mutaciones sinónimas mientras que 1,436 fueron mutaciones no sinónimas. De tal forma que estos cambios en el genoma de la M5 podrían modificar la expresión de genes necesarios para la virulencia de la cepa en comparación con la cepa virulenta 16M. Se podría afectar la envoltura bacteriana por la mutación de genes involucrados en la síntesis del LPS como son *pgm*, *pmm*, *wbdA*, *lpsA*, *mgtA* y *amiC* lo cual resultaría en la reducción en virulencia de la cepa M5. Tres de los genes pertenecientes al operón VirB (*virB4*, *virB6* y *virB9*) con mutaciones no sinónimas, podrían ser importantes para explicar la incapacidad de la cepa M5 de llegar al nicho de replicación bacteriana. Así como los SNP's que se localizaron en los genes encargados del metabolismo de metales como son el hierro y aquellos SNP's que se encontraron en el metabolismo del nitrógeno y de los aminoácidos. Sin embargo, se requiere de ensayos experimentales para corroborar estos supuestos. Wang *et al.*,(2013) reportaron que la cepa M5-90 comparada con su cepa parental M28 tuvo 23 CDS candidatas a estar relacionadas a la atenuación de la bacteria, enfocando su trabajo en el gen *tuf2* involucrado en la traducción de proteínas. Este gen presenta el 30 % de SNP's, así como inserciones y deleciones, encontradas en la cepa M5-90 al compararla con la cepa M28. Fue descrito para la cepa vacunal un menor crecimiento en comparación con la cepa M28, cuando el gen *tuf2* fue mutado en la cepa parental, se observó una disminución en su crecimiento *in vitro* y en su presencia en el bazo de ratones infectados. La restitución de la función de la proteína EF-Tu en la cepa M5-90 restauró el crecimiento y la persistencia de la bacteria en un nivel intermedio, tanto en medios *in vitro* como en infecciones de ratón. Concluyendo que las mutaciones en el gen *tuf2* son en parte responsables por la atenuación de la cepa M5-90.

1.7.2 REV-1

Recientemente, la cepa vacunal Rev-1 que es utilizada en México en la campaña nacional contra la brucelosis, fue secuenciada (Salmon-Divon *et al.*, 2017; Kornspan *et al.*, 2020). Esta cepa REV-1 tiene un genoma 99.9 % idéntico a la cepa virulenta de referencia 16M y presenta 5 regiones de inserción y 3 regiones eliminadas. Una de las inserciones contiene el gen para el regulador MerR relacionado a la resistencia de mercurio y la otra el que codifica para la hidroxipiruvato reductasa, la cual presenta una modificación en su dominio de quinasa. En su análisis de SNPs reportaron mutaciones no sinónimas y cambios de marco de lectura en genes relacionados a virulencia involucrados en el metabolismo de lípidos, aminoácidos y síntesis de pared celular.

El análisis del transcriptoma de la cepa REV-1, en comparación con la cepa 16M crecidas a pH de 7.3 (a lo que denominaron pH normal) y pH 4.4 (a lo que denominaron pH bajo), reveló que 242 genes se expresaron diferencialmente entre las cepas REV-1 y 16M a pH 7.3. 172 genes mostraron un incremento y 70 una reducción (Salmon-Divon, Zahavi and Kornspan, 2019). Entre estos genes se encontraron aquellos que codifican proteínas con funciones de transportadores membranales de cationes, oxidoreductasas, hidrolasas y ATPasas. Además de la sobreexpresión del gen BMEII0704 que corresponde a la bacterioferritina y seis transportadores de tipo ABC (Salmon-Divon, Zahavi and Kornspan, 2019). Por su parte, a pH 4.4 los procesos biológicos más enriquecidos en la cepa REV-1 fueron traducción, procesos metabólicos y transporte (de aminoácidos, de carbohidratos, de proteínas y transportes transmembranales), para la cepa virulenta 16M múltiples procesos biológicos fueron enriquecidos incluyendo patogénesis, división celular y organización de pared celular; estos procesos son esenciales para la sobrevivencia de la bacteria durante su tráfico intracelular hasta la llegada a su nicho de replicación (Salmon-Divon, Zahavi and Kornspan, 2019).

1.8 Cepas de *B. melitensis* mexicanas

1.8.1 La cepa de *B. melitensis* BM133

La cepa de *B. melitensis* 133 (Bm133) es una cepa aislada de campo, con la cual se han realizado diferentes estudios en el Departamento de Microbiología e Inmunología de la Facultad de Medicina Veterinaria y Zootecnia de la UNAM. Mediante la generación de una biblioteca genómica construida en fagémidos, se obtuvo un inserto aproximado de 4000 pb que contiene los genes *atpBEXF* que codifican para las subunidades a, c, b y b' de la ATP sintasa, además de los genes *rnhB* (AF054610) y AF059568 que codifica para una proteína hipotética (Hernández-Castro *et al.*, 2000). En otro trabajo, se utilizó a la cepa Bm133 para generar una mutante en el gen BMEI0216, cuyo producto predicho fue, en su momento, clasificado como una proteína hipotética con cierta similitud a una transglicosilasa, la cual presenta un 100% de similitud con las transglicosilasas de *B. abortus* 2308 y *B. ovis* 25840 (Hernández-Castro *et al.*, 2008). La interrupción del gen mediante un casete de resistencia a kanamicina, generó una mutante que invadía 2 veces menos a 1 hora de estar en contacto con células en cultivo y hasta 19 veces menos a las 2 horas (Hernández-Castro *et al.*, 2008). A pesar de que la cepa Bm133 es virulenta, se ha reportado que comparada con la cepa 16M es menos capaz de invadir y sobrevivir intracelularmente. Por otro lado, la cepa 16 M, crece más rápido en medio líquido, alcanzando la fase logarítmica tardía por lo menos 2 horas antes que la cepa Bm133. Esta diferencia de crecimiento también se ve reflejada durante la invasión de células HeLA a una MOI de 500:1, ya que la cepa 16M invade 2.2 veces más y sobreviviendo a los 30 minutos post-infección (P.I.) 1.6 veces más que la cepa Bm133 sin embargo, esta diferencia no se presentó a los 60 minutos P.I. (Alva-Pérez *et al.*, 2014). Al ser la cepa Bm133 modelo de estudio en nuestro grupo y ser considerada una cepa de referencia mexicana, es de interés secuenciarla para identificar las

diferencias genéticas que a nivel genómico tiene con la cepa de referencia mundial 16M y la cepa de campo mexicana Bm133, diferencias que podrían ayudar a explicar las diferencias fenotípicas que se han observado entre ellas.

1.8.2 La cepa de *B. melitensis* LVM31

Ha sido también del interés de nuestro grupo generar una cepa atenuada a partir de la cepa Bm133. Uno de los esfuerzos se dirigió a la interrupción del gen *omp31*, que pertenece a las proteínas de membrana externa (Omp's) mayoritarias involucradas en el intercambio de solutos; interacción con la pared celular y la estabilidad de la membrana externa, lo cual dio lugar a la cepa LVM31. Esta cepa presenta una caída prematura de viabilidad durante la fase estacionaria comparada con la cepa Bm133 y la cepa vacunal REV-1 (Verdiguél-Fernández, 2014). Además, en el mismo estudio se demostró que la cepa LVM31 es susceptible a la acción de polimixina B (con 42.32% de sobrevivencia) y deoxicolato de sodio (con 50.71% de sobrevivencia) comparada con la cepa Bm133 que presentó 77.15% y 86.52 % de sobrevivencia, respectivamente. La mutante LVM31 también presentó una menor capacidad de sobrevivir intracelularmente en células HeLa y macrófagos murinos, comparada con la cepa silvestre (Verdiguél-Fernández *et al.*, 2017). Más aún, coloniza de manera deficiente el bazo en ratones hembra BALB/c de acuerdo a lo observado a las 3 y hasta las 15 semanas post-inoculación, comparada con la cepa Bm133 (Verdiguél-Fernández *et al.*, 2020). A su vez, se demostró que la protección conferida por la cepa LVM31 a las semanas 3, 6 y 9 post-desafío con la cepa Bm133, provocó una disminución en la colonización esplénica de una manera similar a la de la vacuna REV-1, con lo que se concluyó que la cepa mutante confiere una protección similar a la cepa vacunal en hembras de ratón BALB/c (Verdiguél-Fernández, 2019; Verdiguél-Fernández *et al.*, 2020).

2 Justificación

La coordinación especializada de la red de genes presentes en *B. melitensis* ha demostrado ser determinante en sus procesos de adherencia, invasividad y persistencia intracelular. Las diferencias genómicas existentes entre las cepas vacunales de *B. melitensis* M5 y REV-1 comparadas con la cepa 16M no pasan del 1%, siendo suficientes para generar una atenuación marcada de las dos primeras cepas, lo que permite su uso como vacunas. Sin embargo, los cambios genéticos en las cepas vacunales son variados y modifican de forma muy marcada la virulencia de la bacteria (Jiang *et al.*, 2013; Salmon-Divon, Yeheskel and Kornspan, 2018). La cepa de referencia mexicana de *B. melitensis* Bm133 presenta diferencias reportadas en su velocidad de crecimiento y capacidad de invadir células HeLa y macrófagos murinos.

De aquí el interés en conocer cuáles son las diferencias genotípicas entre las cepas 16M y BM133 que puedan servir para explicar las diferencias fenotípicas en su crecimiento

3 Hipótesis

El crecimiento de las cepas 16M y Bm133 en cultivo líquido muestra cinéticas distintas que influyen su capacidad de invadir y replicarse en macrófagos THP-1, las cuales pueden explicarse en parte por las diferencias identificadas a nivel genómico.

4 Objetivo General

Analizar las diferencias fenotípicas entre las cepas de *B. melitensis* 16M y Bm133 mediante el uso de genómica comparativa, para entender mejor las diferencias existentes entre las cepas, que se relacionen con su fenotipo y que puedan aportar información de la carga genética que posee la mutante LVM31.

4.1 Objetivos Específicos

- 1 Comparar la cinética de crecimiento de las cepas de *B. melitensis* 16M, Bm133 y LVM31 en medio líquido.
- 2 Optimizar un método para la extracción de RNA de *Brucella* para futuros estudios de transcriptómica.
- 3 Secuenciar el genoma de la cepa Bm133 mediante estudios de genómica comparativa con la cepa de referencia 16M.

5 Materiales y Métodos

5.1 Cepas y cultivos

Se usaron las cepas de *B. melitensis* 16M, Bm133 y LVM31 que se obtuvieron del cepario del Laboratorio de Microbiología Molecular. Se cultivaron en agar *Brucella* (AB) a 37 °C bajo una atmósfera de 5% de CO₂ por 72 h como cultivos primarios. También se utilizó caldo *Brucella* (CB) el cual se incubó en agitación a 150 o 200 rpm a 37 °C por 18 h y por 14 h para los ensayos de extracción de RNA total de las cepas de *B. melitensis*.

Se utilizaron monocitos de origen humano de la línea THP-1 ATCC 9184[®], que fueron incubados a 37 °C en una atmósfera con 5% de CO₂, en medio RPMI-1640 suplementado con 0.45% de glucosa, 25 mM de HEPES, 10% de suero fetal bovino, 2 mM de L-glutamina, 1 mM de piruvato de sodio y 0.05 mM de β-mercaptoetanol.

5.2 Cinética de crecimiento de *B. melitensis* 16M, Bm133 y LVM31

Para conocer la fase de crecimiento se realizaron conteos de las Unidades Formadoras de Colonia (UFC's). Se inoculó en 5 ml de CB, 100 µl de un vial congelado de cada cepa y se incubaron a 150 rpm, 37 °C por 24 h. Posteriormente a las 24 h se realizó un cultivo secundario con una dilución 1/50 del cultivo inicial. En 4.9 ml de CB se añadieron 100 µl del cultivo primario, homogenizando perfectamente la muestra. Antes de incubar los cultivos, se tomaron 20 µl del cultivo y se sembraron en 180 µl de CB, con lo que se realizaron diluciones decuples seriadas hasta 10⁻⁹, se sembraron placas de AB 3 gotas de 20 µl por dilución. Los cultivos secundarios se incubaron a 150 rpm, 37 °C para

tomar una muestra del cultivo cada 2 h hasta las 24 h o 36 h según fuese la cepa. Las placas de agar se incubaron por 72 h a 37 °C con 5% CO₂. Pasado el tiempo de incubación de las placas se realizó el conteo de las UFC's de cada dilución. Con esta cinética se determinó la hora del cultivo secundario en la que la bacteria se encuentra en fase logarítmica tardía y el número de UFC's presentes en el cultivo.

5.3 Extracción de RNA bacteriano

La extracción del RNA total de las cepas 16M, Bm133 y LVM31 se realizó utilizando la metodología de Fenol ácido proporcionada por el doctor José Luis Puente, del Laboratorio de Microbiología Molecular, del Instituto de Biotecnología (IBT-UNAM), modificada de acuerdo con las indicaciones del Doctor Simón Guzmán del Departamento de Microarreglos de DNA, del Instituto de Fisiología Celular-UNAM. A continuación se describe la metodología, se dividió en dos partes, la primera se realizó en la Unidad de Bioseguridad del departamento de Microbiología e Inmunología de la FMVZ en donde se manejaron los cultivos de *B. melitensis*, así como la lisis y extracción con fenol:cloroformo:alcohol isoamílico 125:24:1 (v/v/v) pH 4.5 de la marca Ambion. Mientras que la segunda parte de la extracción se llevó a cabo en el Laboratorio de Microbiología Molecular ya con las muestras en fenol.

A partir de placas sembradas con las cepas 16M, Bm133 y LVM31, se inoculó un asada del cultivo en 20 ml de CB en tubos de 50 ml con cierre hermético de la marca Corning Star. Se incubaron a 37 °C a 150 rpm por 14 h. Se centrifugó el cultivo a 3000 rpm por 20 min y se decantó el sobrenadante al terminar. Se añadió 500 µl de la solución A fresca (10mM de EDTA pH 8 0.5M, 20mM acetato de sodio pH 4.8 3M y SDS 10%). Se homogenizó en vortex hasta que la solución se volvió viscosa. Se agregaron 5 a 7 µl de DEPC de Sigma y se homogenizó en vortex. El líquido se transfirió a tubos de 1.5 ml con O-Ring. Se agregaron 500 µl

de fenol ácido, precalentado a 62 °C en baño seco (tiempo de precalentado de 10-20 minutos), por cada tubo. Mezclar en vortex vigorosamente por 15 segundos. Con esto finaliza la primera parte del protocolo. Se incubaron los tubos en baño maría a 62 °C, agitando el tubo por inversión durante 10 a 15 segundos cada 2 minutos. Terminado el tiempo de incubación se centrifugaron los tubos a 13,500 rpm por 30 minutos a 4 °C, en microcentrifuga Eppendorf. Se transfirió la fase acuosa a un tubo nuevo de 1.5 ml, solo se tomaron 400 µl. Se añadió 0.1 volúmenes o 40 µl de Acetato de sodio pH 5.3 3M (ajustado con ácido acético glacial) y se homogenizó por pipeteo. Se añadió etanol al 100%, 2.5 volúmenes y se mezcló 10 veces por inversión de tubo. Se reservaron los tubos a -80 °C durante mínimo 1 h. Se centrifugaron los tubos a 13,500 rpm por 20 a 30 minutos a 4 °C. Se decantó el sobrenadante por inversión rápida del tubo y se secó el exceso con un paño limpio. Se agregó 1 ml de etanol al 70% a -20 °C. Los tubos fueron mezclados por inversión del tubo y centrifugados a 13,500 rpm por 15 minutos a 4 °C. Se decantó el sobrenadante por inversión del tubo y sin invertirlos de nuevo se dejaron secar, abiertos y sobre una sanita limpia durante 20 a 30 minutos. Se suspendieron con 30 a 40 µl agua libre de nucleasas sin DEPC de Ambion, se dejaron reposar con el agua por 20 minutos a 4 °C. Después se homogenizó cada muestra con pipeta.

Para comprobar la integridad y la cantidad de RNA obtenido se realizó una electroforesis horizontal de agarosa al 1% en TAE 1X, el gel se corrió de 90 V por 40 minutos. Cada muestra fue teñida con 5-8 µl de amortiguador de carga, mezclado con 10 µl de bromuro de etidio (10 mg/ml) por mililitro de amortiguador. Los geles se digitalizaron mediante el Mini Lumi®.

5.4 Ensayos con DNAsa I

Para eliminar el DNA contaminante en las extracciones de RNA bacteriano, se decidió utilizar la DNAsa I de la marca Thermo Scientific® como se especifica en el cuadro 1.

Protocolo de reacción con DNAsa I para RNA Total	
RNA	10 µg cbp en µl
Amortiguador de DNAsa	2 µl
DNAsa	2 µl
H ₂ O libre de nucleasas	Cbp µl
Total de reacción	20 µl

Cuadro 1.

Se descongelaron todos los reactivos necesarios y las muestras en hielo. El amortiguador de la enzima se homogenizó en vortex durante 15 a 20 segundos. Se agregaron todos los reactivos a un tubo de 600 µl, estos se mezclaron por vortex y se les centrifugó por 5 segundos. Se incubaron las reacciones a 2 diferentes temperaturas 25 °C y 37 °C por 15 minutos en termociclador. Se detuvo la reacción agregando 2 µl de EDTA 25 mM e incubando a 65 °C por 10 minutos en termociclador. Los ensayos se verificaron por medio de electroforesis como se describió previamente.

5.5 RT-PCR del RNA total de *B. melitensis*

Se realizó una RT-PCR para la amplificación de los genes *omp22* (Verdiguell-Fernández, 2019) y *omp31* (iniciador delantero F'FCASC 5'-CATCCATGGTCCATCGCCGCTATGTTTCG-3' e iniciador reverso R'FCASC CTATACGTACTTCTTCATTCCGTGTTTCGTGG-3' , diseñados por el alumno de licenciatura Vilchis Malvaez Carlos Alberto) de Bm con iniciadores diseñados por tesisistas del laboratorio de Microbiología Molecular, utilizando la enzima SuperScript III de Invitrogen se modificaron las indicaciones del fabricante para

la retro transcripción: Primer paso se añadió a un microtubo de 200 µl, 2 µl de primer reversa a 10 µM/µl de los pares mencionados, 1 µl de RNA bacteriano a una concentración de 1 µg/µl, deoxinucleótidos tri-fosfatados o DNTP's 1µl a una concentración de 10 mM/µl por nucleótido y H2O libre de nucleasas cbp 14.5 µl de volumen final de la reacción. Las reacciones se colocaron en termociclador a 65 °C por 5 minutos y luego a 4 °C por 1 minuto. Al terminar el ciclo se añadió a cada reacción 4 µl de Amortiguador 5X, 0.5 µl de inhibidor de RNAsas y 1 µl de la Enzima SuperScript III a una concentración de 200 u/µl. Los tubos fueron homogenizados en vortex y centrifugados inmediatamente, para después ser colocados en el termociclador a 50°C por 30 minutos, 85 °C por 5 minutos y 4°C. Como siguiente paso a cada reacción de retro-transcripción se le añadió 2 unidades de RNAsa H de la marca Thermo Scientific. Los tubos fueron incubados a 37 °C por 20 minutos en termociclador. Para la PCR punto final se escribirán los protocolos que se siguieron para cada reacción en los siguientes cuadros:

PCR <i>omp22</i>	
Amortiguador [] 5X	5 µl
MgCl ₂ [] 25 mM	5 µl
DNTP's [] 10 mM	2.5 µl
Iniciador delantero [] 10 µM	2 µl
Iniciador reverso [] 10 µM	2 µl
cDNA	3 µl
Taq recombinante 1 u/µl	2.5 µl
H2O cbp 50 µl volumen final	28 µl

Cuadro 2.

Las condiciones de temperatura y tiempo para la PCR de *omp22* son las siguientes: 94 °C por un minuto, 55.5 °C por un minuto; y 72 °C un minuto y medio por 30 ciclos, con una extensión final a 72 °C por 5 minutos.

PCR <i>omp31</i>	
Amortiguador [] 5X	2.5 µl
MgCl ₂ [] 25 mM	4.5 µl
DNTP's [] 10 mM	1 µl
Iniciador delantero [] 10 µM	1 µl
Iniciador reverso [] 10 µM	1 µl
cDNA	1 µl
Taq recombinante 5 u/µl	0.5 µl
H ₂ O cbp 25 µl volumen final	13.5 µl

Cuadro 3.

La PCR de *omp31* se realizó bajo las siguientes condiciones de tiempo y temperatura: 94 °C por 30 segundos, 55 °C por 30 segundos y 72 °C por 45 segundos; todo lo anterior por 30 ciclos. Con una síntesis final a 72 °C por 5 minutos.

5.6 Extracción de DNA

La extracción de DNA de las cepas Bm133 y LVM31 se llevó a cabo utilizando el reactivo DNAzol de la marca Invitrogene con las siguientes modificaciones al protocolo. Se sembraron las cepas Bm133 y LVM31 en una placa de AB se incubaron a 37° con 5% CO₂ por 72 h. Se sembró una asada de cada cepa en 20 ml de CB en tubos cónicos de 50 ml de tapa hermética y se incubaron a 37 °C en agitación (200 rpm) por 24 h. Al día siguiente se realizó la extracción de DNA en la unidad de bioseguridad mediante el uso del reactivo DNAzol modificando el procedimiento de la siguiente manera: Se centrifugaron los cultivos a 3000 rpm por 20 minutos, se descartó el sobrenadante y se suspendió el precipitado bacteriano en 1 ml de la solución de DNAzol. Se homogenizaron

los tubos por agitación vigorosa en vortex durante 20 segundos y se transfirieron 500 µl de la suspensión a tubos de 1.5 ml con tapa O-Ring y se añadieron 500 µl de Fenol:Cloroformo 1:1 (v/v). Se cerraron perfectamente los tubos y se homogenizaron mediante vortex por 20 segundos. Con las bacterias neutralizadas por el fenol fue posible continuar con la extracción del DNA en el Laboratorio de Microbiología Molecular. Se centrifugaron los 4 tubos de 1.5 ml a 12,000 rpm por 10 minutos. Se transfirió la fase acuosa a un nuevo microtubo de 1.5 ml y se añadieron 500 µl de Fenol:Cloroformo:Alcohol Isoamílico (25:24:1) homogenizando la muestra por 20 segundos y se centrifugó a 12,000 rpm por 10 minutos. El sobrenadante fue transferido a un microtubo de 1.5 ml nuevo, después se añadieron 0.7 volúmenes de isopropanol y se incubó a -80 °C por 1 h. Los tubos fueron centrifugados a 12,000 rpm por 15 minutos. Se decantó el sobrenadante y se le agregó 1 ml de etanol al 70 % homogenizando mediante inversión del tubo 10 veces. Las muestras se centrifugaron a 12,000 rpm por 10 minutos y al terminar el sobrenadante fue decantado. El precipitado de DNA fue secado en Savant por 20 minutos. La muestra se suspendió en 100µl de H₂O libre de nucleasas.

Las muestras de DNA de la cepa Bm133 y LVM31 se observaron mediante electroforesis en un gel de agarosa al 1% en TAE 1X y se cuantificaron en nanodrop. Posteriormente, los 100 µl del DNA fueron tratados con 4 µl de RNAsa A (5 mg/ml) a 37 °C por 20 minutos. Con la finalidad de eliminar los restos de fenol y RNAsa A presentes en la muestra debido a los procedimientos anteriores, el material genético de ambas cepas fue tratado con el kit comercial EZ-10 Spin Column DNA Gel Extraction Miniprep Kit de BIO BASIC INC. para purificación de DNA como lo recomienda el fabricante. 30 µl del DNA puro correspondiente a las cepas Bm133 y LVM31 fueron enviados para su secuenciación.

5.7 Secuenciación de las cepas Bm133 y LVM31

Los DNAs de las cepas de *B. melitensis* Bm133 y LVM31 fueron enviados y procesados por el Centro Nacional de Servicios de Diagnostico en Salud Animal (CENASA) de la SENASICA, que es parte de la SADER. Mediante el uso de un secuenciador de segunda generación MiSeq de Illumina, se realizó una secuenciación pareada de 150 pb mediante el uso del kit Nextera DNA Flex Library siguiendo las recomendaciones del fabricante. A continuación, se describe el proceso de secuenciación por síntesis de Illumina de forma general: Como primer paso el DNA se fragmentó de manera aleatoria mediante nebulización o sonicación, para después seleccionar el tamaño de 150 pb de los fragmentos que se secuenciaron, mediante una electroforesis horizontal. Los fragmentos son entonces purificados de la matriz del gel y se les prepara para añadir los adaptadores o etiquetas siguiendo las recomendaciones del fabricante del kit Nextera. Los fragmentos ya con las etiquetas o adaptadores se agregaron a la laminilla en donde se hibridaron con los oligos que se encuentran fijados en la matriz. El siguiente paso fue la amplificación de los fragmentos para la formación de los “clusters” o agrupaciones mediante una PCR de puente, lo que permitió la generación de cadenas complementarias fijas a la laminilla y así como aumentar la cantidad hasta 1000 copias por “cluster”. Con las agrupaciones ya formadas se eliminan las cadenas complementarias y solo se conservaron las cadenas sencillas. El proceso de secuenciación por síntesis, comienza con la adición de los nucleótidos marcados con fluoróforos por la polimerasa uno por vez, gracias a la modificación de los dNTPs. Después de que se sintetizó una base, el fluoróforo es excitado mediante un LASER que barre por completo la laminilla, la fluorescencia emitida es capturada mediante una fotografía que se guarda para cada ciclo. Terminando la toma de la imagen se elimina el bloqueo en el dNTP y se continúa con el ciclo de síntesis, el proceso se repite las veces necesarias de acuerdo al tamaño a amplificar y si la secuenciación fue pareada

o sencilla. La computadora divide la laminilla en cuadrantes, por lo que la posición de los “clusters” será específica para cada lectura que genere el equipo para cada cuadrante y para cada ciclo, lo que genera miles de fotografías de alta resolución con las que la computadora después genera las secuencias por lecturas y las asocia a la calidad de fluorescencia que obtuvo. Se recibieron los archivos del cuadro 4.

Archivos que contienen las lecturas obtenidas en la secuenciación	
Nombre de las lecturas pareadas	Tamaño del archivo en Megabytes (Mb)
BM133-BRUCELLA-MELITENSIS_S11_L001_R1_001.fastq.gz	30.9
BM133-BRUCELLA-MELITENSIS_S11_L001_R2_001.fastq.gz	33.1
LVM-31-PURA_S12_L001_R1_001.fastq.gz	33.9
LVM-31-PURA_S12_L001_R2_001.fastq.gz	36.2

Cuadro 4.

Se verificó el taxón de pertenencia de cada muestra con el programa KRAKEN2, el cual también nos permite ver el nivel de contaminación que hubo en la muestra mediante un esquema circular generado con el programa Kroona. Este proceso se realizó por parte del personal de CENASA y los reportes fueron recibidos junto con las lecturas de la secuenciación.

5.8 Ensamblaje de los genomas Bm133 y LVM31

Para el análisis de la calidad de secuencias, ensamblaje, alineamientos genómicos y anotación del genoma, se utilizaron diferentes programas bioinformáticos que son de libre acceso en el sistema operativo Linux, así como

la plataforma en línea de libre acceso Galaxy.org (Afgan *et al.*, 2018). Los programas y los procedimientos serán mencionados en los siguientes párrafos.

5.9 Análisis de calidad de las lecturas de la cepa Bm133 y la mutante LVM31

La revisión de calidad de las lecturas presentadas en el cuadro (4), se realizó mediante el programa FastQC versión 0.11.8 con el cual se analiza la calidad de las lecturas o “reads” obtenidas del procedimiento de secuenciación masiva. El programa también nos permite hacer una revisión de la cantidad de lecturas y su tamaño promedio con lo cual se calculó la cobertura que se tenía del material genético de referencia de la cepa 16M, dividiendo el número total de nucleótidos obtenidos en los archivos pareados para cada cepa entre el tamaño total del genoma de *B. melitensis* 16M (que se obtiene de la suma del tamaño en nucleótidos del cromosoma I [ChI] y el cromosoma II [ChII]). Esto se obtuvo mediante el uso de `zgrep` y `awk` con el siguiente comando:

```
zgrep . Nombre del archivo.fastq.gz |
  awk 'NR%4==2{c++; l+=length($0)}
  END{
    print "Number of reads: "c;
    print "Number of bases in reads: "l
  }'
```

5.10 Edición de lecturas “reads” de la cepa Bm133

La edición de “reads” cuadro (5) se realizó con el programa Cutadapt 3.2 en Python 3.7 (Martin, 2021), que nos permitió editarlas mediante el retiro de los adaptadores utilizados en la secuenciación; así como una selección con base en su nivel de calidad. Estos niveles de calidad hacen referencia a la probabilidad

de que se presenten errores de secuenciación a lo largo de las lecturas. Para los valores de 40 se espera 1 error cada 10,000 bases o un 99.99% de certeza en que la base que se obtuvo en el proceso es la correcta y para los niveles menores a 10, se espera 1 error cada 10 bases o una certeza del 90 %. Se eligieron los niveles de calidad desde 22 hasta 30, usando solo niveles pares de acuerdo con el reporte que se obtuvo en el FastQC.

Archivos de “reads” editados con el programa Cutadapt 3.2

Nombre de los archivos pareados	Tamaño de los archivos en Mb
BM133_R1_cuttred_q22_min18.fastq	87.9
BM133_R2_cuttred_q22_min18.fastq	87.6
BM133_R1_cuttred_q24_min18.fastq	87.6
BM133_R2_cuttred_q24_min18.fastq	87.0
BM133_R1_cuttred_q26_min18.fastq	87.0
BM133_R2_cuttred_q26_min18.fastq	86.2
BM133_R1_cuttred_q28_min18.fastq	86.3
BM133_R2_cuttred_q28_min18.fastq	85.2
BM133_R1_cuttred_q30_min18.fastq	85.1
BM133_R2_cuttred_q30_min18.fastq	84.0

Cuadro 5.

5.11 Ensamble de novo del genoma de la cepa Bm133 y la mutante LVM31

Los archivos pertenecientes a la cepa Bm133 se usaron para realizar un ensamblaje de *novo* de las lecturas cortas obtenidas en la edición, con el programa SPAdes genome assembler v3.13.1 (Bankevich *et al.*, 2012). De la misma manera, se utilizó como comparativa del ensamblaje el programa

Unicycler v0.4.8 (Wick *et al.*, 2017) utilizado por Salmon-Divon *et al.*, (2017) para el ensamblaje de la cepa vacunal REV-1 de *B. melitensis*. Se utilizaron los valores programados por defecto en cada uno de los programas y solo para el programa de SPAdes se decidió utilizar la modalidad “single cell”, ya que se ha reportado en la literatura una mejor capacidad de ensamblaje para genomas y detección de genes. Además, el programa fue diseñado principalmente para este propósito.

5.12 Análisis de los resultados del ensamblaje de *novovo*

Para evaluar la calidad de ensamble producido por los programas antes mencionados se utilizó el programa QUAST: Quality Assessment Tool for Genome Assemblies, Versión: 5.0.2 (Gurevich *et al.*, 2013), que nos permite comparar la calidad de los ensambles de secuencias contiguas o “contigs” por parte de los ensambladores y compararlas con las secuencias de los cromosomas de referencia de *B. melitensis* 16M que se encuentran anotadas con número de referencia: **AE008917.1** *B. melitensis* bv. 1 str. 16M chromosome I y **AE008918.1** *B. melitensis* bv. 1 str. 16M chromosome II en el NCBI. Es necesario recalcar que se probaron todos los ensambles producidos con SPAdes y Unicycler con cada cromosoma de *B. melitensis* 16M por separado para verificar los valores de N50, el NG50, el número total de “contigs”, el porcentaje de genoma ensamblado, así como un aproximado de genes que se lograron ensamblar con respecto a la referencia (cepa 16M) y por último, el orden y dirección de los “contigs” ensamblados. Se escogió el ensamblado que tuvo el menor número de “contigs”, el N50 más alto y el mayor porcentaje de completitud. El ensamble con nombre BM133_R1_cuttad_q22_min18.fastq y BM133_R2_cuttad_q22_min18.fastq fue elegido, ya que SPAdes solo le da el nombre al archivo como “contigs.fasta”. Para evaluar dirección y porcentaje de identidad de las secuencias contiguas, con respecto al Genoma de referencia de *B. melitensis* 16M, se utilizó NUCmer (NUCleotide MUMmer) versión 3.1,

recordando que esto se realizó de manera independiente para cada cromosoma antes de tener los “contigs” de cada cromosoma separados.

5.13 Mejoramiento del ensamblaje mediante el uso de SSPACE

Después de evaluar la identidad con Nucmer, el archivo “contigs.fasta” obtenido con el programa SPAdes fue procesado por el programa SSPACE para tratar de realizar un andamiaje o “scaffolding”. SSPACE requiere de ciertos valores que se obtienen de los archivos de las lecturas mediante el programa Bbmap. Se generó el archivo “BM133_scaff_ext.final.scaffolds.fasta”, que se analizó nuevamente con el programa QUAST, para determinar si hubo una mejora en los “contigs” del ensamblaje.

5.14 Ordenamiento y alineamiento de “contigs” de la cepa Bm133

El archivo “BM133_scaff_ext.final.scaffolds.fasta” fue ordenado con el programa Mauve Snapshot_2015-02-13-Genome Alignment Visualization (Rissman *et al.*, 2009) y la opción “Move contigs” permitió ordenar las secuencias contiguas con respecto al genoma de referencia de la cepa 16M. Para este proceso se utilizó el archivo gbk de cada uno de los cromosomas de la cepa 16M lo que permite una anotación de las secuencias codificantes o “coding sequences” (CDS por sus siglas en inglés). Mauve entrega una lista con los “contigs” que se deben usar por cada cromosoma, pero no los separa del archivo “fasta” que se obtiene al final del reordenamiento de “contigs”. Para separar las secuencias que conforman cada cromosoma de la cepa Bm133, se utilizó el programa faSomeRecords.py con el cual se pudo separar los “contigs” que pertenecían al ChI y al ChII de *B. melitensis* 133.

Gráficas de similitud o “dotplots” fueron realizadas con el archivo de contigs ordenado por Mauve por medio de mummerplot, un programa que se usó desde la plataforma en línea Galaxy, para verificar el ordenamiento del archivo BM133_scaff_ext.final.scaffolds.fasta para cada cromosoma, así como la concordancia del genoma ensamblado comparada con la referencia (Bm133 contra 16M).

5.15 Anotación de los genes de la cepa de referencia mexicana de *B. melitensis* Bm133

Ya con los genes ordenados por cromosoma se procedió a realizar un previo de anotación del genoma con el programa Prokka que está especializado en la anotación de genomas procarióticos. El programa Prokka se usó desde la plataforma en línea Galaxy (Afgan *et al.*, 2018) y los archivos se reservaron para el proceso de anotación del genoma en la plataforma de Genbank.

6 Resultados

6.1 Cinéticas de crecimiento de las cepas de *B. melitensis* 16M, Bm133 y LVM31

Considerando que *B. melitensis* 16M presenta una mayor internación cuando se encuentra en la fase de crecimiento logarítmica tardía (Rossetti *et al.*, 2009), evaluamos la cinética de crecimiento de las cepas de *B. melitensis* 16M, Bm133 y LVM31 en cultivos de CB líquidos. Esto con el propósito de establecer el tiempo en el que cada cepa alcanza dicha fase.

Como se puede observar en la Figura 1, a las 18 horas la cepa Bm133 alcanza la fase logarítmica tardía con un desfase de 2 h con respecto a la cepa 16M (16 h).

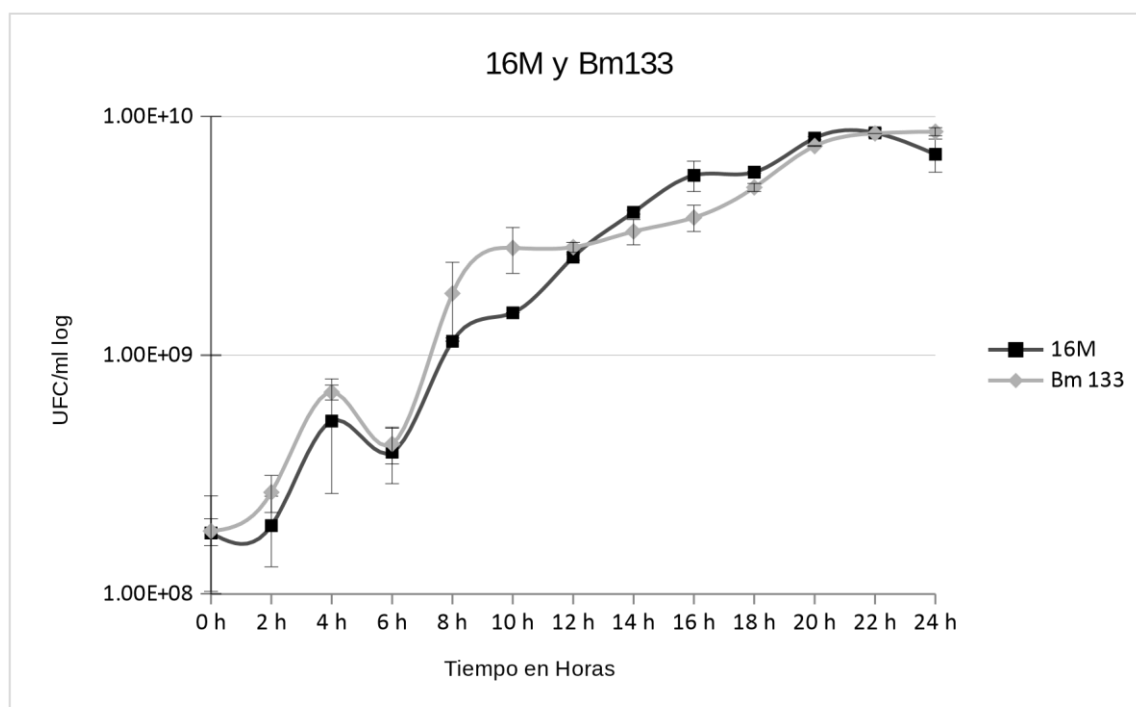


Figura 1. Cinética de crecimiento de las cepas 16M y Bm133. *B. melitensis* 16M (■) y Bm133 (◆) En crecidas en CB líquido. El crecimiento se siguió cada dos horas durante 24 h como se describe en los métodos. Los puntos representan la media de 3 repeticiones con sus respectivas desviaciones estándar.

Por su parte, la cepa LVM31 presenta un crecimiento atípico con respecto a la cepa parental Bm133 ya que empieza a crecer después de las 22 h en cultivo, pasando por una larga fase de adaptación al medio. Entre las 22 y 24 h crece rápidamente para alcanzar la fase logarítmica tardía entre las 24 y las 32 h. Este crecimiento se atribuye a la interrupción del gen *omp31*, localizado en el cromosoma II, por un casete de Kanamicina; sin embargo, la cepa no ha sido complementada por lo que su secuenciación será importante para descartar la existencia de mutaciones secundarias.

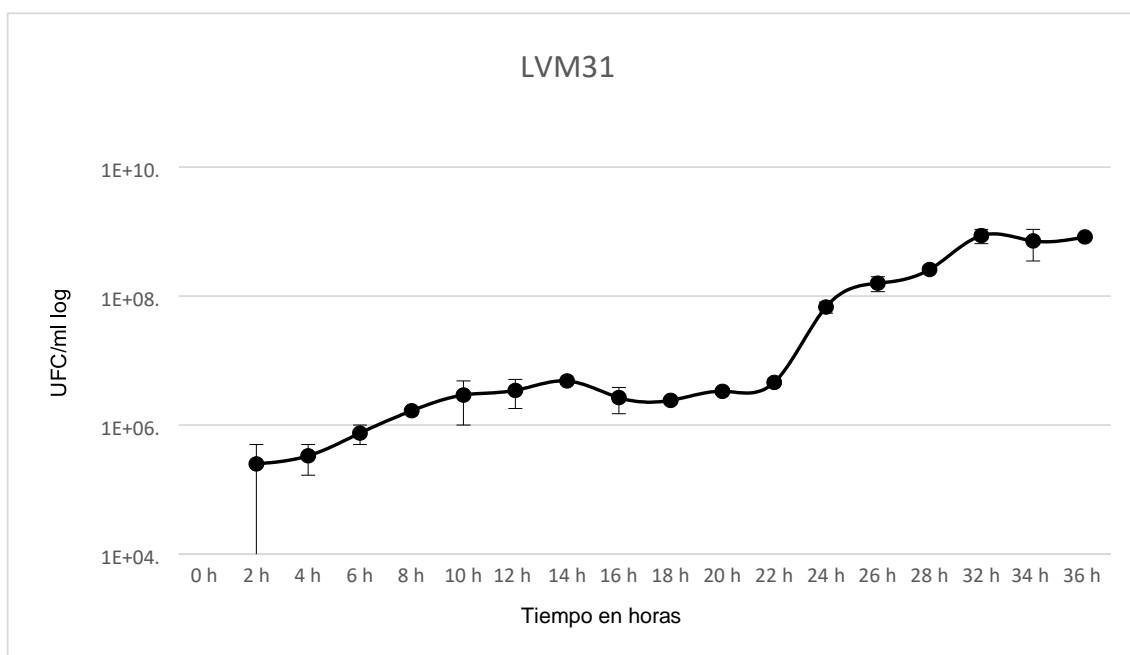


Figura 2. Cinética de crecimiento de la cepa LVM31. El crecimiento se siguió cada dos horas durante 36 h como se describe en los métodos. Cada punto representa la media de 3 repeticiones con sus respectivas desviaciones estándar.

6.2 Extracción de RNA total de las cepas 16M, Bm133 y LVM31

Con el propósito de adecuar un protocolo que permita purificar RNA total de *Brucella* reduciendo los tiempos de trabajo sin comprometer la calidad del RNA para que se pueda usar en estudios de transcripción se realizaron diferentes pruebas con la metodología de fenol ácido como se describe en materiales y métodos. Se pudo comprobar que mientras la muestra se encuentre con el fenol

ácido de Ambion (pH 4.5), el RNA se mantiene estable incluso con un tiempo de espera de 40 minutos a temperatura ambiente. Esto permite realizar la extracción del material genético con un mínimo de pérdida de calidad y disminuyendo el tiempo de estancia en la Unidad de Bioseguridad 3 (Figura 3A).

El material que se obtuvo es de calidad, se optó por la utilización del protocolo TriReagent de Sigma, con modificaciones al protocolo con células THP-1 como punto de evaluación de los resultados de la extracción del RNA. Los resultados de dicha extracción se muestran en la Figura (3, Panel B).

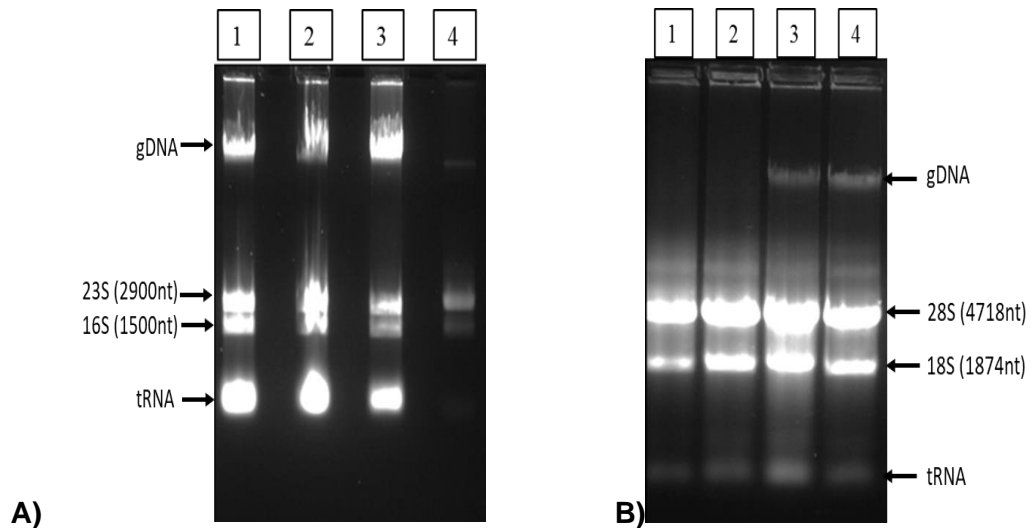


Figura 3. Extracción de RNA total de *B. melitensis* 16M, Bm133 y LVM31; y monocitos THP-1. **Panel A)** RNA total cepa 16M (carril 1), RNA total Bm133 (carril 2), RNA total LVM31 (carril 3) y RNA total de *E. coli* DH5 α (carril 4). Todas las muestras se obtuvieron siguiendo el protocolo de fenol ácido. **Panel B)** Carril 1 al 4 RNA total de monocitos THP-1.

Un problema que presentó la metodología utilizada fue la obtención del RNA contaminado con DNA genómico, por lo que se realizó un tratamiento con DNAsa I libre de RNAsas a 37 °C; sin embargo, se observó que se modificaba la proporción de las bandas del RNA bacteriano de manera drástica (dato no mostrado). El tratamiento se realizó entonces a 25 °C por 15 y 20 minutos (Figura 4, Panel A).

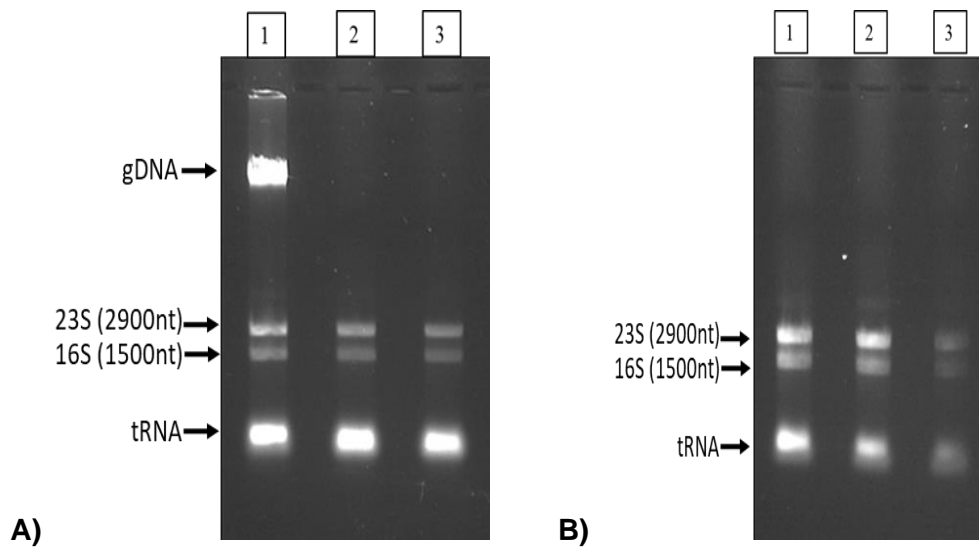


Figura 4. RNA total tratado con DNasa I.

A) RNA total de la cepa 16M sin tratar (carril 1), tratamiento por 15 minutos (carril 2) y tratamiento por 20 minutos (carril 3). **B)** RNA total de las tres cepas de *Brucella* después de tratamiento por 15 minutos. Cepas 16M (carril 1), Bm133 (carril 2) y LVM31 (Carril 3).

Con este resultado se determinó que, aunque no existe una diferencia aparente entre los tiempos de 15 y 20 minutos, es preferible manejar lo menos posible el RNA tanto en procesos como en tiempo, así que en los ensayos subsiguientes se utilizó el tiempo de 15 minutos de incubación a 25 °C (Figura 4B).

La cantidad de brucelas que logran internarse en macrófagos infectados es alrededor de 1×10^4 hasta 1×10^5 bacterias por los 3.5 millones de células, variando el número según la proporción de células y la fase de crecimiento de la bacteria. Con la finalidad de estimar la cantidad posible de RNA bacteriano que se podría extraer a partir de células infectadas para futuros experimentos de transcriptómica, se realizaron diluciones seriadas a partir de cultivos de 14 h de *B. melitensis* 16M. Un primer ensayo utilizando para la extracción de RNA diluciones décuples seriadas desde 1/10 hasta 1/10,000,000, a partir de un cultivo de *B. melitensis* 16M, mostró que el RNA comienza a ser menos aparente en el gel a partir de la dilución 1/100 (Figura 5A). La misma serie de diluciones se realizó con RNA ya purificado a partir de la cepa 16M (Figura 5 B).

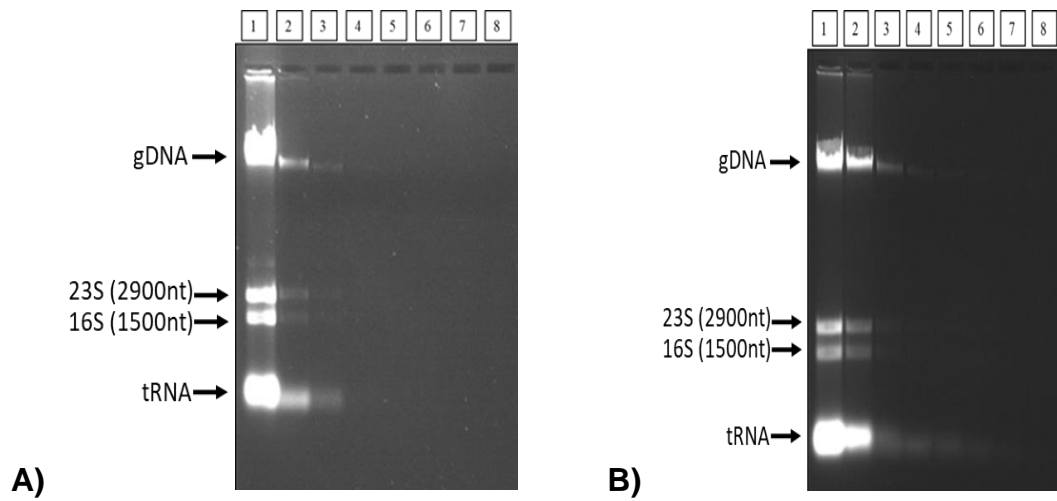


Figura 5. RNA obtenido a partir de diluciones decuples seriadas de un cultivo de *B. melitensis* 16M. **A)** RNA extraído de *B. melitensis* 16M a partir de diluciones decuples seriadas de un cultivo crecido en CB por 14 h. Cepa sin diluir (carril 1). Diluciones 10 a la -1, -2, -3, -4, -5, -6 y -7 respectivamente (carriles 2-8). **B)** Diluciones decuples seriadas a partir de 2 $\mu\text{g}/\mu\text{l}$ de RNA total de *B. melitensis* 16M purificados partiendo de 6×10^9 bacterias. RNA sin diluir (carril 1), diluciones 10 a la -1, -2, -3, -4, -5, -6 y -7 respectivamente.

Posteriormente, se hizo un ensayo semi-cuantitativo por RT-PCR con diluciones decuples seriadas iniciadas en 1/10 de los RNA's purificados usando ribonucleótidos que amplifican el gen *omp31* (Figura 6)

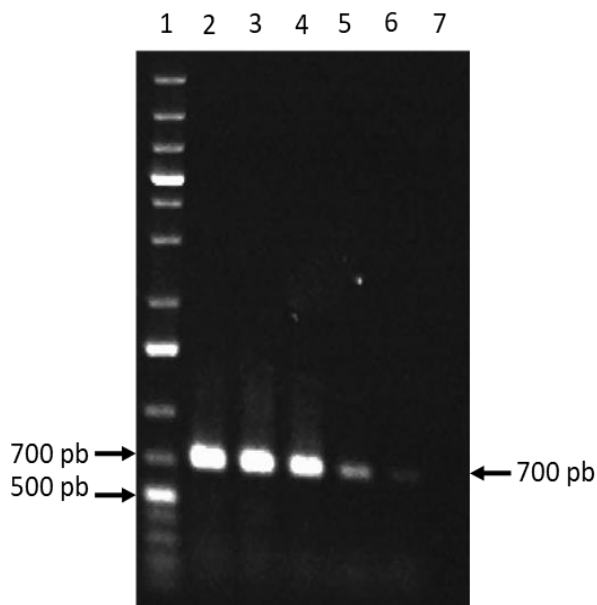


Figura 6. RT-PCR semi-cuantitativo del gen *omp31* de la cepa 16M. Marcador de peso molecular 1Kb plus (carril 1), RNA de la cepa 16M sin diluir (carril 2), diluciones del RNA desde 1 a -4 (carriles 3-6) y Control negativo (carril 7).

La cantidad aproximada de RNA por bacteria según la literatura es en promedio de 100 fg por bacteria (Westermann, Gorski and Vogel, 2012; Hör, Gorski and Vogel, 2018) valor que se toma aproximado ya que puede variar dependiendo de la especie y las condiciones en las que se encuentra la bacteria. La cantidad de RNA de la cepa 16M sin diluir con la que se realizaron estos experimentos fue de 2077 ng/ μ l cuantificados en el Nanodrop. Esto permitió inferir que en la dilución -4 se tenía una concentración de 0.2077 ng/ μ l o 207.7 pg/ μ l, que en proporción con los 0.1 pg que tiene una bacteria de RNA representan 2077 bacterias. Con base en esta información, se concluyó que mediante la PCR descrita se podría evaluar la presencia de RNA de *Brucella* a partir de RNA total obtenido de células THP-1 infectadas.

6.3 Ensamblaje Bm133 y LVM31

El genoma de las cepas Bm133 y LVM31 fue secuenciado en CENASA, de donde se recibieron tres carpetas por cada cepa conteniendo las lecturas en crudo de cada una de las secuenciaciones, un ensamblado previo hecho en el software CLC bio de Quiagen que se utiliza para realizar el proceso de ensamblaje de las lecturas y además el estudio de taxonomía hecho con el programa Kraken2, que permite asignar etiquetas taxonómicas a secuencias cortas de DNA y estimar la abundancia de secuencias. Kraken2 determinó que la ruta taxonómica más baja de las secuencias, pertenecían a un solo organismo con un 99.13 % de las secuencias, de estas secuencias el 98.37% se etiquetaban al género *Brucella* y en específico para la especie *melitensis* se determinaba una ruta a la especie de *Brucella* con el 14.16 % de las secuencias, para la cepa Bm133 (Figura 7).

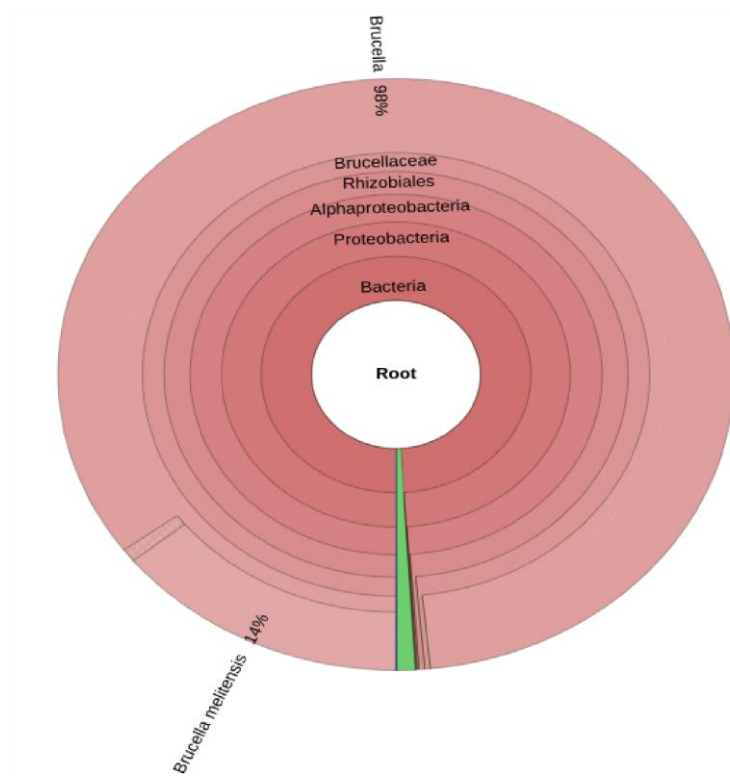


Figura 7. Diagrama Krona de la cepa *B. melitensis* Bm133 secuenciada.

Se observa de forma jerárquica siguiendo el diagrama del centro a la periferia, la abundancia taxonómica a la que pertenecen las lecturas (*reads*) obtenidos en la secuenciación. En color rosado se presenta las secuencias pertenecientes a *B. melitensis*, mientras que de color verde se puede observar las secuencias contaminantes o que no tuvieron una secuencia que alineara con algún organismo

6.4 Análisis de calidad

La calidad de las lecturas se analizó con la herramienta FastQC y se realizó el cálculo de la cobertura del genoma por medio del uso de comandos de zreg y awk. Se obtuvieron 295,964 y 316,996 lecturas para las cepas Bm133 y LVM31, respectivamente, de entre 35-151 bases, cero secuencias de calidad pobre con un porcentaje de G-C de 56 % en ambos casos que es muy cercano al porcentaje de la cepa referencia *B. melitensis* 16M. Ya que los resultados gráficos de dicho proceso se obtienen por cada archivo, para poderlos usar de manera más

sintetizada se usó la herramienta MultiQC v1.12.dev0 que permite combinar los resultados de FastQC y presentar juntos los resultados de los 4 archivos de calidad que se obtuvieron (Figuras 8-11):

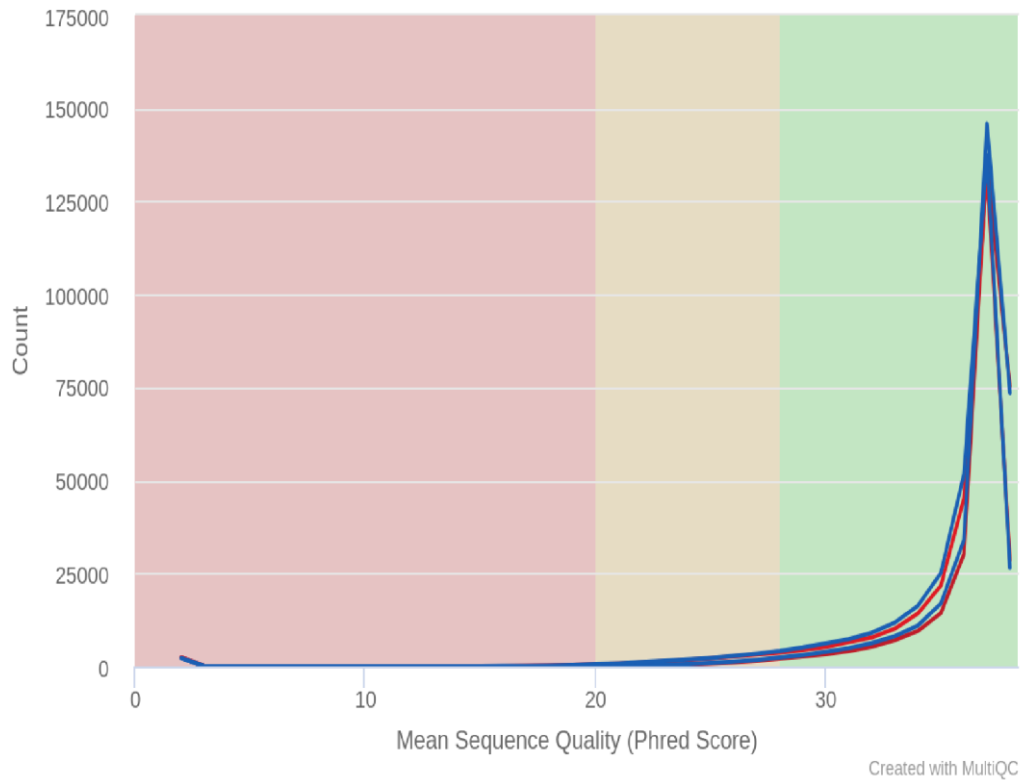


Figura 8. Reporte de calidad combinado de las lecturas de la cepa Bm133 y de la mutante LVM31. Distribución de la cantidad de lecturas de acuerdo a su calidad para los 4 archivos. Se observa en la gráfica que la distribución tiende a calidades arriba de 30, lo cual nos da una probabilidad de error por base de 0.0001. Cepa Bm133 (líneas rojas ■) y mutante LVM31 (líneas azules ■). Gráfico creado en MultiQC.

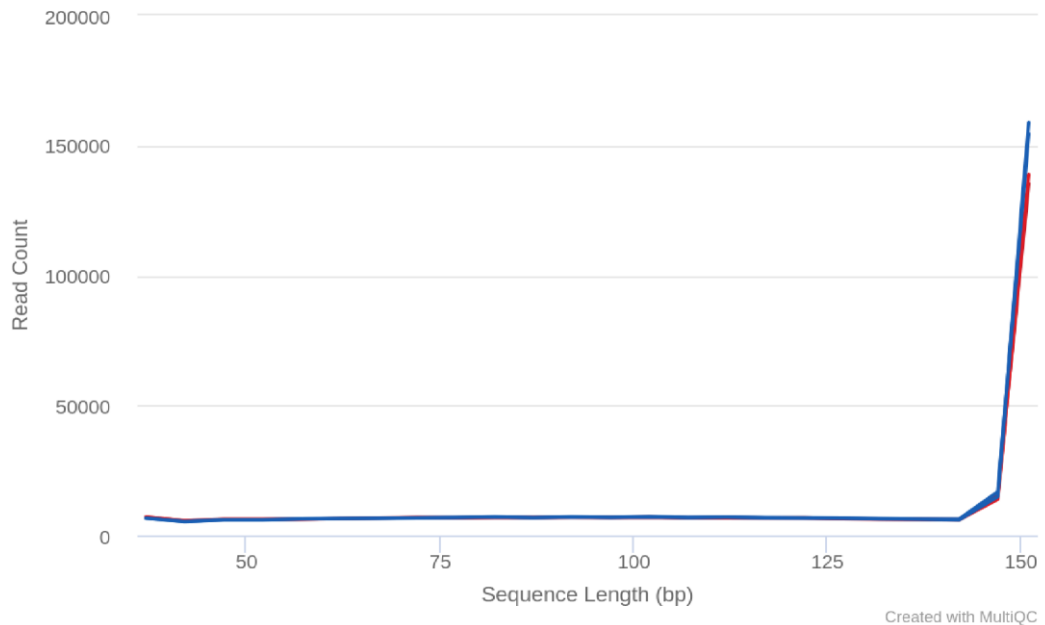


Figura 9. Distribución de la longitud de las secuencias de las lecturas de las cepas Bm133 y LVM31. Se observa en la gráfica que la distribución del largo de secuencias que se encuentran en el rango de 35-151 nt, así como la cantidad de lecturas que se encuentran dentro de este rango. La cepa Bm133 (líneas rojas ■) y la mutante LVM31 (líneas azules ■). Gráfico creado en MultiQC.

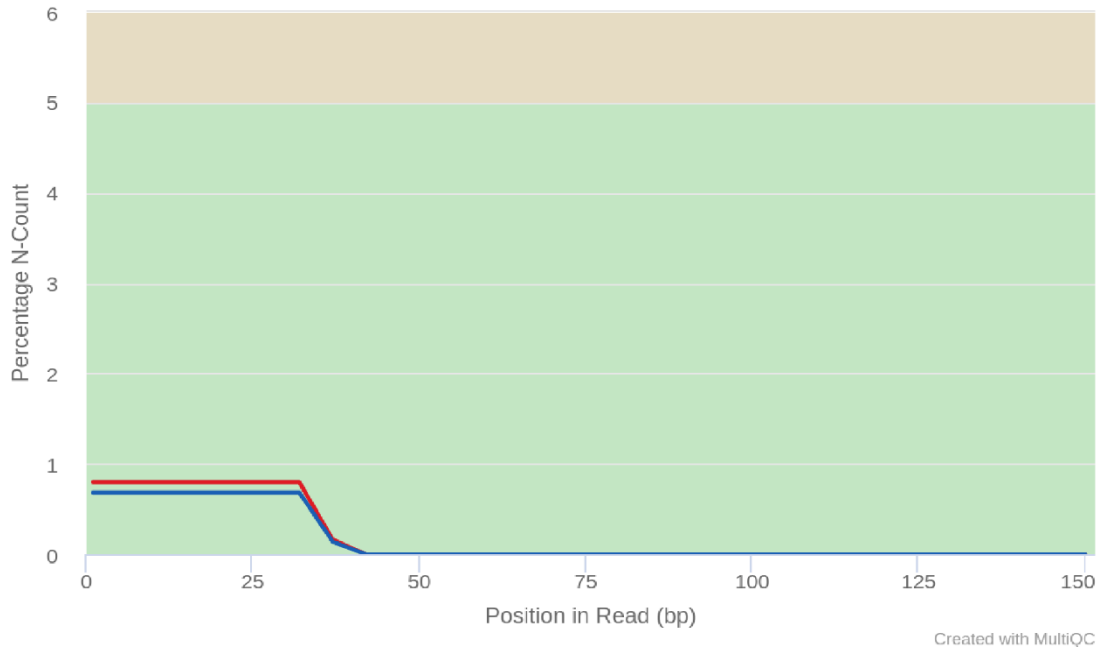


Figura 10. Contenido de N's en las secuencias de las lecturas para los 4 archivos. Se observa en la gráfica que algunas bases dentro de las lecturas no fueron correctamente identificadas en la secuenciación por lo que se les dio una clasificación N. Cepa Bm133 (■) y mutante LVM31 (■). Gráfico creado en MultiQC.

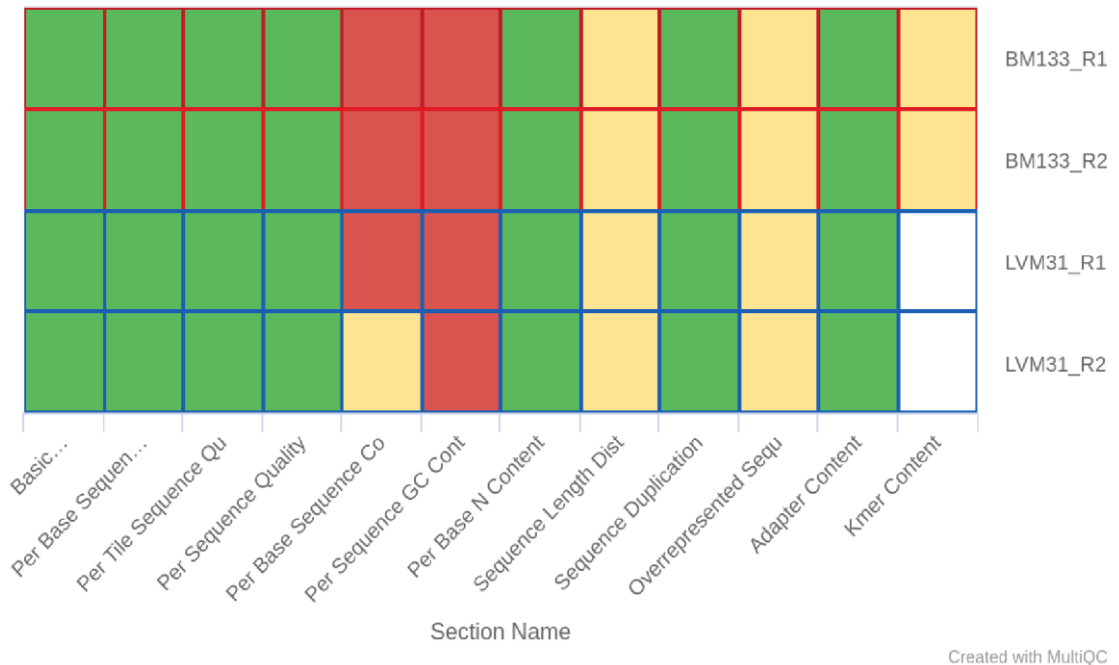


Figura 11. Resumen del análisis de calidad de las cepas Bm133 y LVM31. Mapa de calor de cada uno de los valores reportados por FastQC. Parámetros en rango adecuado (■), presencia de advertencias (■), parámetros con problemas (■) y parámetros sin datos (□). El cuadrículado (■) enmarca los datos de la cepa Bm133 y en cuadrículado azul (■) los datos de la LVM31. Gráfico creado en MultiQC.

Estos análisis permitieron determinar qué parámetros tener en cuenta para realizar un corte o refinamiento de las lecturas mediante el uso del programa Cutadapt. Por ejemplo, se pudo determinar que las lecturas debían tratarse para eliminar nucleótidos indeterminados (N) y que, además, debían eliminarse los adaptadores que se usaron para la secuenciación.

El cálculo de cobertura de genoma se realizó a partir de las lecturas crudas de la cepa Bm133. Entre los dos archivos se obtuvieron 591,928 lecturas un total de nucleótidos de 110,554,264, lo que dividido entre 3,294,931 nucleótidos, que es el total del genoma de *B. melitensis* bv. 1 str. 16M (código de ensamblaje en Assembly del NCBI ASM712v1), se estableció una cobertura de 33.55 veces, lo

que es suficiente para poder ensamblar un borrador del genoma de la cepa Bm133 dentro de los parámetros aceptados.

6.5 Edición de lecturas de la cepa Bm133

La edición de las lecturas fue mediante el uso del programa Cutadapt utilizando 5 diferentes cortes de calidad (22, 24, 26, 28 y 30). Todas aquellas lecturas por debajo de estos cortes de calidad fueron retiradas del archivo fastq final y se usaron para evitar que al momento de la edición se perdiera más de un 10 % de las lecturas, lo cual no es deseable para el proceso de ensamble; con esos niveles de calidad no se redujo más del 5 % del total de las lecturas. Así mismo, se procedió a eliminar las lecturas que estuviesen por debajo de una longitud de 18 nt. En las Figuras 12-14 se muestran las gráficas de calidad obtenidas para la cepa Bm133 comparadas con las lecturas sin tratar; mediante FastQC y combinadas con MultiQC después de que las secuencias fueron tratadas en el programa Cutadapt.

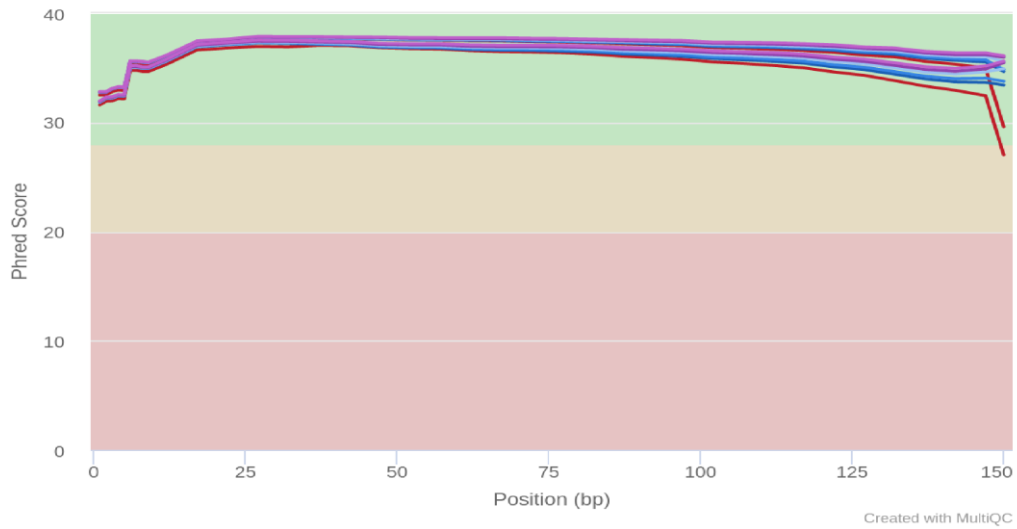


Figura 12. Reporte de calidad combinado de las lecturas de la cepa Bm133. Calidad de las lecturas para los archivos. Cepa Bm133 sin tratar (■), BM133_R1_cuttred_q22_minl18.fastq y el archivo R2 editado (■), BM133_R1_cuttred_q24_minl18.fastq y el R2 editado (■), BM133_R1_cuttred_q26_minl18.fastq y el archivo R2 editado (■), Líneas de color (■) corresponden a BM133_R1_cuttred_q28_minl18.fastq y el archivo R2 editado y las líneas de color (■) BM133_R1_cuttred_q30_minl18.fastq y el archivo R2 editado. Gráfico creado en MultiQC.

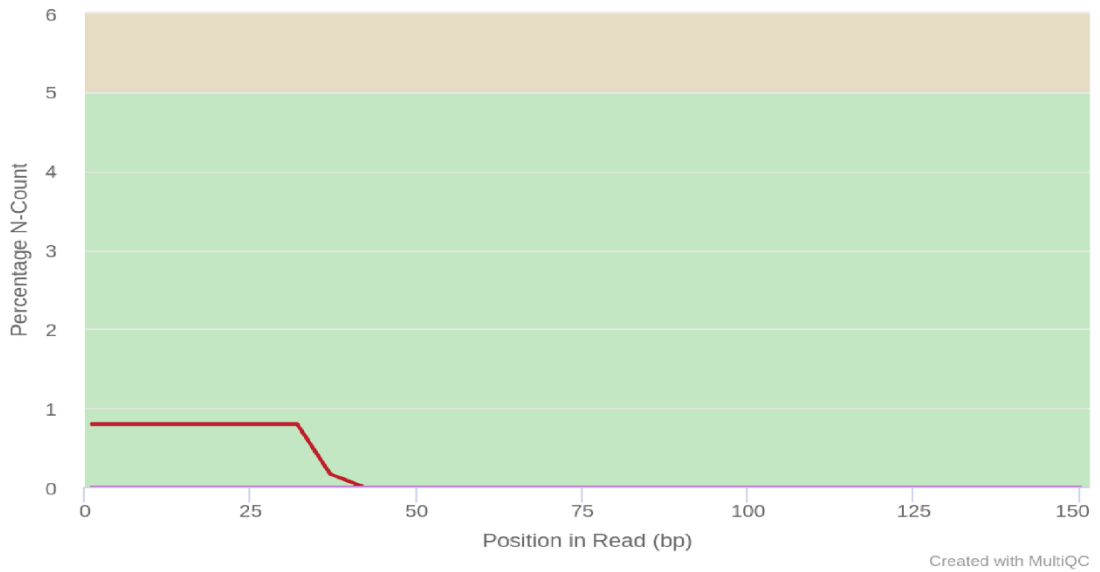


Figura 13. Figura 13 Contenido de bases N en las lecturas. En líneas (■) la cepa Bm133 si tratar. Líneas de color (■) BM133_R1_cuttred_q22_minl18.fastq y el archivo R2 editado, en líneas de color (■) BM133_R1_cuttred_q24_minl18.fastq y el archivo R2 editado, BM133_R1_cuttred_q26_minl18.fastq y el archivo R2 editado (■). Líneas de color (■) corresponden a BM133_R1_cuttred_q28_minl18.fastq y el archivo R2 editado. Líneas de color (■) corresponden a BM133_R1_cuttred_q30_minl18.fastq y el archivo R2 editado. Gráfico creado en MultiQC

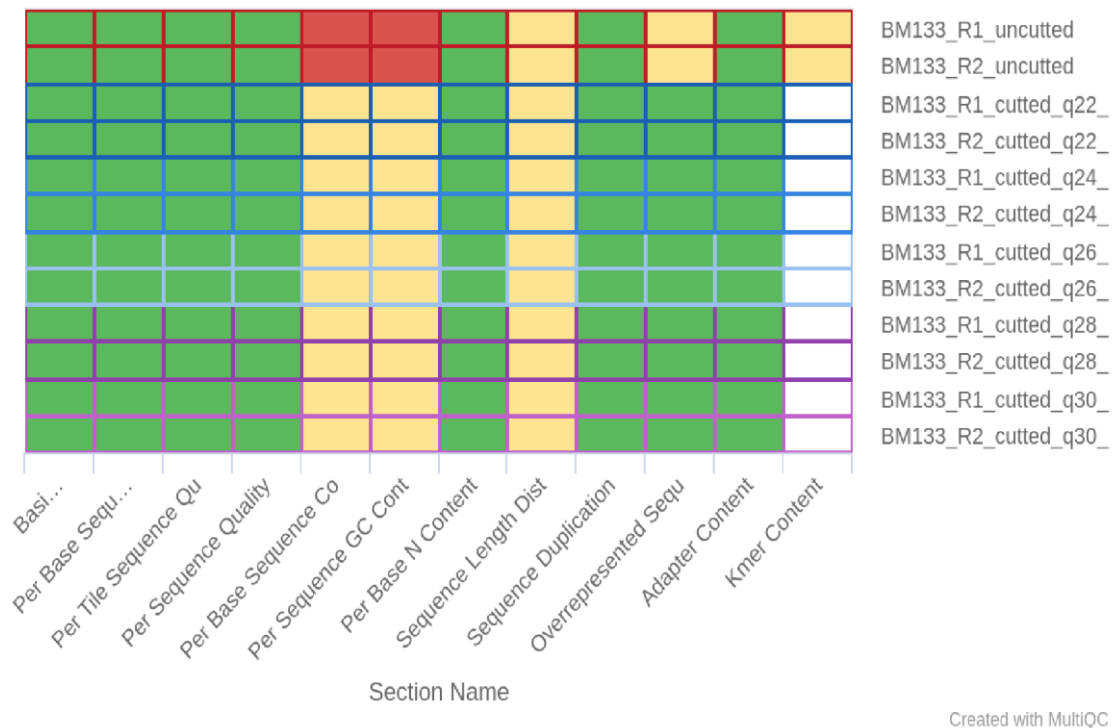


Figura 14. Mapa de frecuencias de los valores reportados en todos los análisis de FastQC de una cepa Bm133.

Parámetros en rango adecuado (■), presencia de advertencias (■), parámetros con problemas (■) y parámetros sin datos (□). En (■) líneas rojas representan los datos de la cepa Bm133 si tratar, las lecturas BM133_R1_cutted_q22_min18.fastq y el archivo R2 editado (■), BM133_R1_cutted_q24_min18.fastq y el archivo R2 editado (■), en líneas de color (■) BM133_R1_cutted_q26_min18.fastq y el archivo R2 editado, la cuadrícula de color (■) corresponde a BM133_R1_cutted_q28_min18.fastq y el archivo R2 editado. BM133_R1_cutted_q30_min18.fastq y el archivo R2 editado en (■). Gráfico creado en MultiQC

6.6 Ensamble *de novo* del genoma de la cepa Bm133

Con las secuencias ya refinadas con el programa de corte, se procedió a su ensamblaje mediante el uso de SPAdes que es un ensamblador de secuencias cortas. El protocolo de ensamblaje se basó en las opciones que tiene el programa por defecto, solo modificando la modalidad de ensamblaje a la opción “single cell”. Siendo el caso para la cepa Bm133 el uso de todos los archivos que se obtuvieron en la edición.

6.7 Análisis de los resultados del Ensamblaje *de novo*

Ya con los ensamblajes terminados, se utilizó el programa Quast (Quality Assessment Tool for Genome Assemblies o Herramienta para la Evaluación de la Calidad de Ensamblajes Genómicos) para evaluar la calidad de los ensamblajes usando como referencia los cromosomas de *B. melitensis* 16M (AE008917.1 *B. melitensis* bv. 1 str. 16M chromosome I y AE008918.1 *B. melitensis* bv. 1 str. 16M chromosome II) que se encuentran reportados en el NCBI. Es necesario recalcar que, para este procedimiento, sólo se puede usar uno de los cromosomas de referencia por evaluación, por lo que todos los archivos se trabajaron dos veces, una para evaluar el ensamblaje con respecto al cromosoma I (ChI) y otro para evaluar el ensamblaje con respecto al cromosoma II (ChII). Como resultados destacables del proceso de evaluación, el ensamblaje que obtuvo mejores evaluaciones por parte del programa Quast fue el del archivo SPADES_def_q22_min18_bm133_contigs, el cual presentó un menor número de secuencias contiguas (*contigs*) con un total de 64, un valor N50 de 92,000 pb, siendo el *contig* más larga de 179.1 kbp, con un 99.291 % total de nucleótidos que alinean entre los *contigs* del ensamblaje y los nucleótidos del genoma con respecto al ChI de la cepa referencia 16M. Además de las 4263 características genómicas (lista de genes presentes en el archivo gbk de la referencia) presentes en el ChI de *B. melitensis* 16M, la cepa Bm133 tuvo 4156 + 55 partes de genes en el ChI.

Para el caso de la evaluación con respecto al ChII, los valores son los mismos para la N50 y el *contig* más largo; pero para la totalidad de nucleótidos del cromosoma es mayor con un porcentaje de 99.8 %, de la misma manera para las características genómicas del ChII, que en la cepa de referencia son de 2333 y en el ensamblado son de 2305 + 17 partes. Estos valores (cuadro 6) indicaron

que se podía continuar con el proceso de andamiaje, alineamiento y ordenamiento de *contigs* con base en una cepa de referencia.

Resultados de Quast modificados Bm133 vs 16M ChI								
Sample Name	N50 (Kbp)	# contigs	Largest contig (Kbp)	Length (Mbp)	Mismatches/100kbp	Indels/100kbp	Genes	Genome Fraction
SPADES_def_bm133_withoutcut_contigs_ChI	92	65	179.1Kbp	3.3Mbp	10.51	3.62	3004	99.30%
SPADES_def_q22_min18_bm133_contigs_ChI	92	64	179.1Kbp	3.3Mbp	10.61	3.62	3004	99.30%
SPADES_def_q24_min18_bm133_contigs_ChI	71.3	74	172.2Kbp	3.3Mbp	10.66	3.66	3007	99.30%
SPADES_def_q28_min18_bm133_contigs_ChI	71.5	76	172.1Kbp	3.3Mbp	10.99	3.62	3010	99.20%
SPADES_def_q30_min18_bm133_contigs_ChI	71.5	80	172.1Kbp	3.3Mbp	10.95	3.62	3010	99.20%

Resultados de Quast modificados Bm133 vs 16M ChII								
Sample Name	N50 (Kbp)	#contigs	Largest contig (Kbp)	Length (Mbp)	Mismatches/100kbp	Indels/100kbp	Genes	Genome Fraction
SPADES_def_bm133_withoutcut_contigs_ChII	92	65	179.1Kbp	3.3Mbp	20.84	5.44	3004	99.80%
SPADES_def_q22_min18_bm133_contigs_ChII	92	64	179.1Kbp	3.3Mbp	20.5	5.44	3004	99.80%
SPADES_def_q24_min18_bm133_contigs_ChII	71.3	74	172.2Kbp	3.3Mbp	21.1	5.53	3007	99.80%
SPADES_def_q28_min18_bm133_contigs_ChII	71.5	76	172.1Kbp	3.3Mbp	21.1	5.44	3010	99.80%
SPADES_def_q30_min18_bm133_contigs_ChII	71.5	80	172.1Kbp	3.3Mbp	21.01	5.36	3010	99.80%

Cuadro 6. Resumen de los valores obtenidos por el programa QUASt para los diferentes ensamblajes que se realizaron. La evaluación se realizó comparado los ensamblajes con los cromosomas ChI y ChII de *B. melitensis* 16M. Barras de color (■) para la tabla Bm133vsChI. Barras de color (■) para la tabla Bm133vsChII. Tabla modificada de MultiQC

Para evaluar cómo interactúan los *contigs* con el genoma de referencia se realizó una corrida de alineamiento en Nucmer del archivo de ensamblaje vs los dos cromosomas ChI y ChII de la cepa de referencia 16M. Se pudo apreciar que para ambos cromosomas el porcentaje de identidad de los *contigs*, con respecto a las áreas en las que se alineaban, estuvo dentro del rango de 99-100%, para los *contigs* del ChI (Figura 15) y de 98-100% para el ChII (Figura 16). Otra cosa que se pudo observar para ambas alineaciones fue que los *contigs* que estaban al inicio también aparecían al final del alineamiento lo que parece indicar circularidad de cada uno de los cromosomas, la cual se podría lograr con un ensayo de ensamblaje híbrido con una secuenciación de lecturas largas.

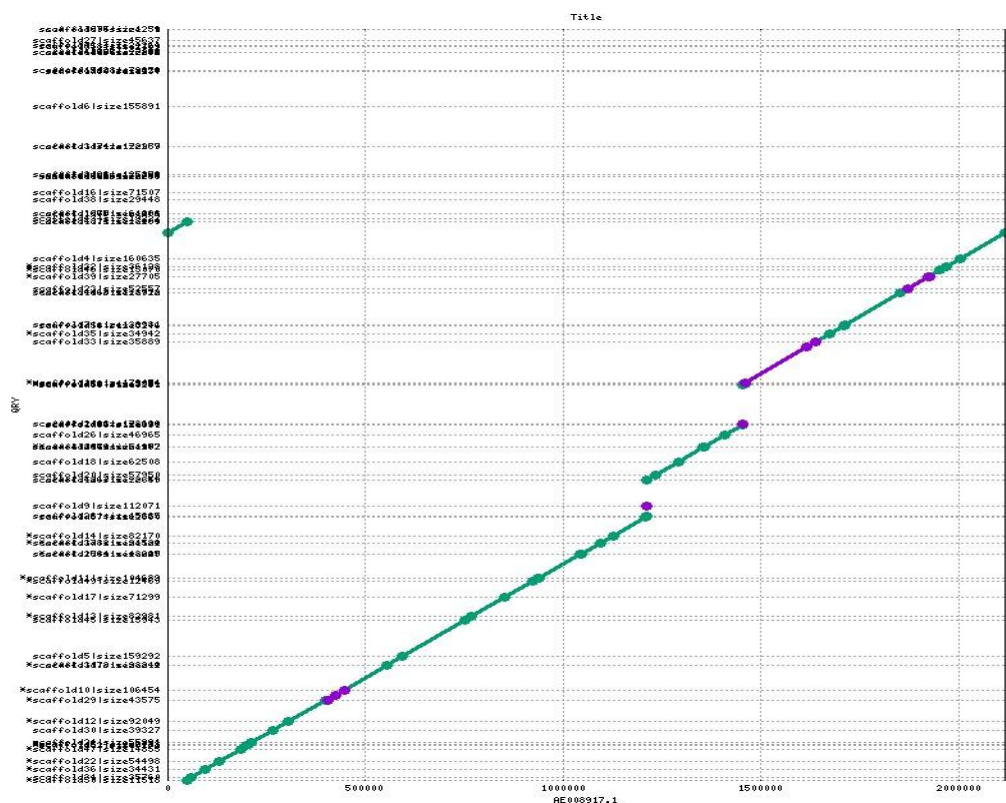


Figura 15. Gráfica de identidad del Cromosoma I (ChI) de Bm133 vs 16M. El eje de las Y representa todos los “contigs” presentes en el ensamblado. El eje de las X es la secuencia completa del ChI de la cepa 16M. Obtenida con el programa MUMmerplot por Nucmer en la plataforma Galaxy.

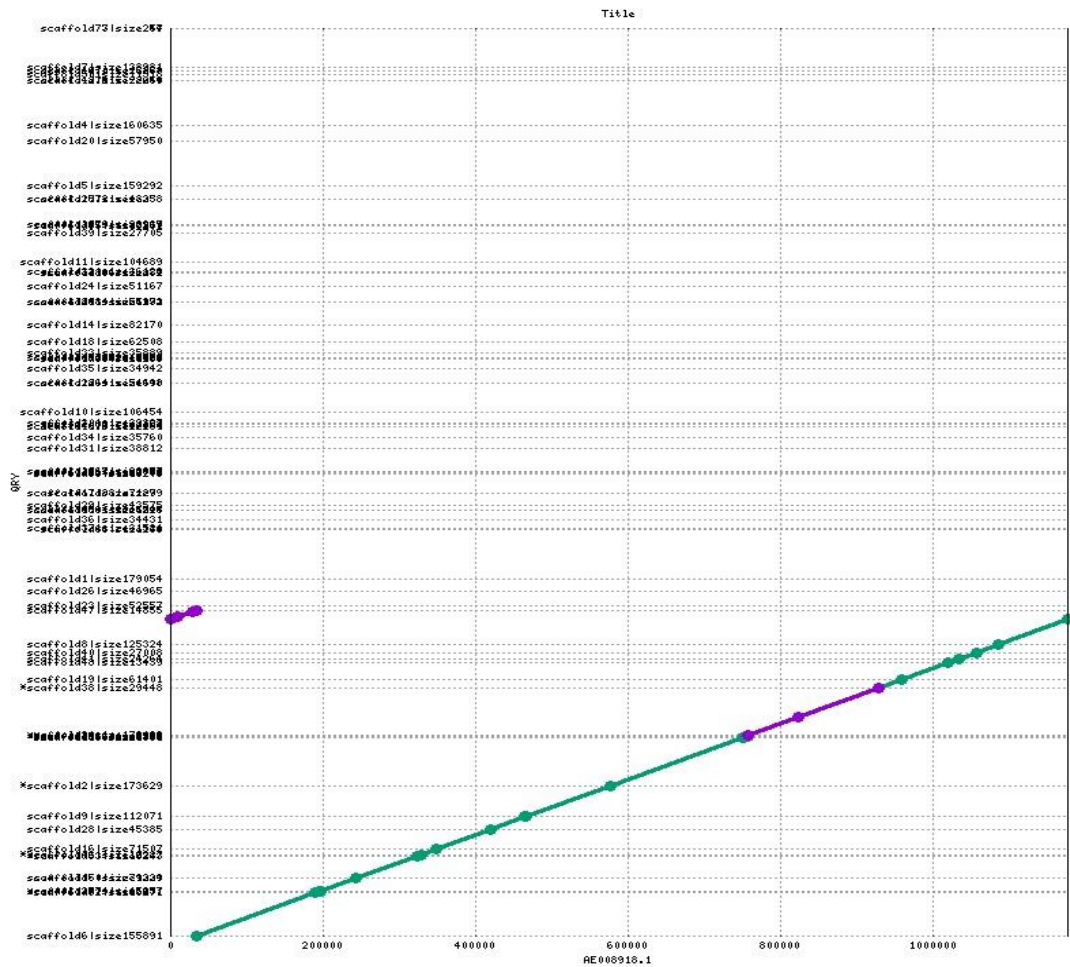


Figura 16. Gráfica de identidad del Cromosoma II (ChII) de Bm133 vs 16M. El eje de las Y representa todos los “contigs” presentes en el ensamblado. El eje de las X es la secuencia completa del ChII de la cepa 16M. Obtenida con el programa MUMmerplot por Nucmer en la plataforma Galaxy.

En las figuras anteriores se puede apreciar que la relación de continuidad para el alineamiento del ChI de las cepas Bm133 y 16M presenta cambios más abruptos incluso dentro de algunos *contigs* (Figura 15), mientras que la relación que para el ChII es mucho más lineal (Figura 16). En las gráficas se puede apreciar que las secciones iniciales y finales de los dos cromosomas comparten un *contig* para cada uno de los alineamientos.

6.8 Mejoramiento del ensamblaje mediante el uso de SSPACE

Para mejorar el ensamblado realizado en el programa SPAdes y hacer que los *contigs* que se generaron aumentaran un poco de tamaño produciendo mejores sobrelapamientos entre los *contigs* dando lugar a la formación de andamios (*scaffolds*) más largos, se utilizó la herramienta SSPACE. Se logró una mejora mínima en el llenado de los espacios existentes entre los *contigs*, sin lograr generar *scaffolds* más grandes. La evaluación del ChI y del ChII que se obtuvo solo mejoró en una milésima parte del porcentaje de alineamiento para cada cromosoma, por lo que se decidió pasar directamente al procedimiento de alineamiento.

6.9 Ordenamiento y alineamiento de “contigs” de la cepa Bm133

El ordenamiento de los *contigs* con respecto de un genoma de referencia se realizó utilizando la herramienta MAUVE que permite tanto el alineamiento de los genomas, como el reordenamiento de secuencias contiguas con el uso de una secuencia de referencia, aunque solo permite el uso de una única referencia por corrida por lo que se hicieron para cada cromosoma por separado.

Con el ensamble SPADES_def_q22_min18_bm133_contigs se procedió al reordenamiento, usando como referencia los archivos .gbk de los dos cromosomas de la cepa 16M AE008917.1 *B. melitensis* bv. 1 str. 16M chromosome I y AE008918.1 *B. melitensis* bv. 1 str. 16M chromosome II. La razón para usar los archivos .gbk y no los .fasta, es que al usar los archivos .gbk Mauve produce un apartado de cuáles son las secuencias codificantes (CDS por sus siglas en inglés), que presentan SNP's o polimorfismos simples de nucleótidos, presencia de mutaciones sin sentido, “*gaps*” o espacios entre las secuencias codificantes, y sustituciones de aminoácidos. Además de *gaps* o

espacios y SNP's en el total del alineamiento. En las figuras (17 y 18), se observa el resultado final del reordenamiento de los dos cromosomas de la cepa Bm133.

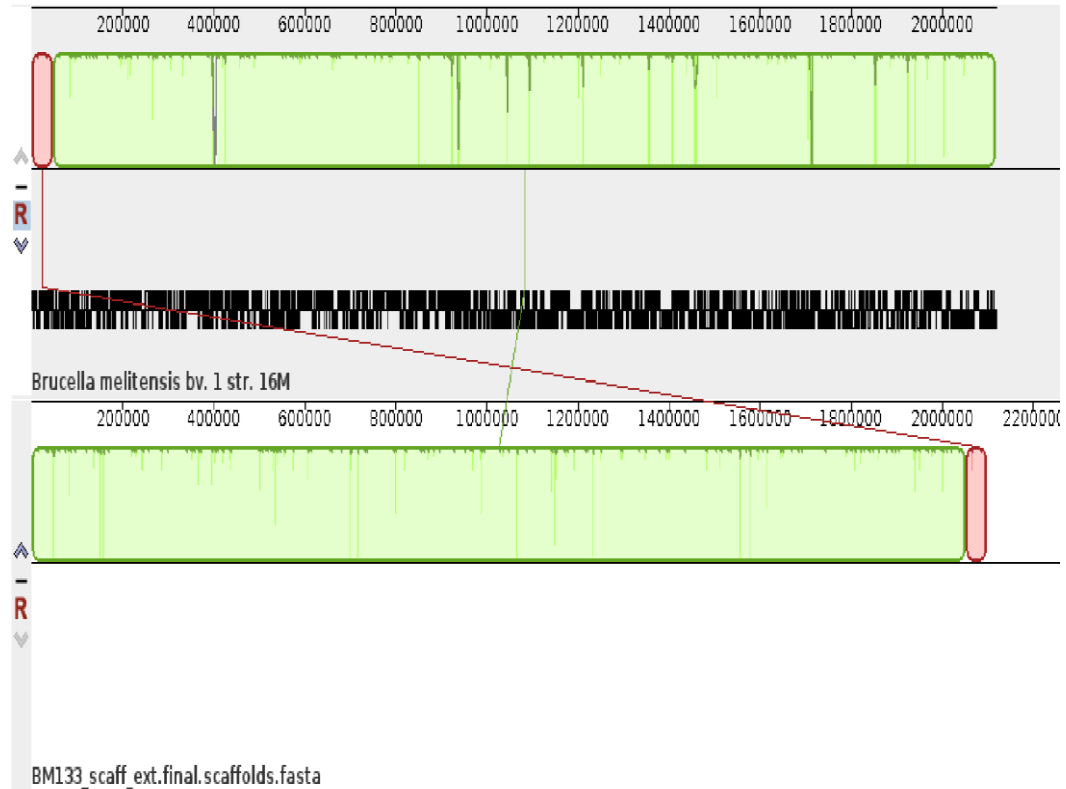


Figura 17. Representación del ordenamiento final de "contigs" del ChI con Mauve. Se utilizó el ensamble SPADES_def_q22_min18_bm133_contigs, usando como referencia la secuencia del ChI de la cepa 16M en .gbk completo. Barra superior roja (inicio de la secuencia) y barra superior verde (el total de la secuencia) es la representación del ChI de la cepa 16M. Barra inferior roja (final de la secuencia) y barra inferior verde (el total de la secuencia) es la representación del ordenamiento ensamble de la cepa bm133 con respecto al ChI de la cepa de referencia. Barra negra segmentada representa los CDS anotados para la cepa de referencia y para el ChI.

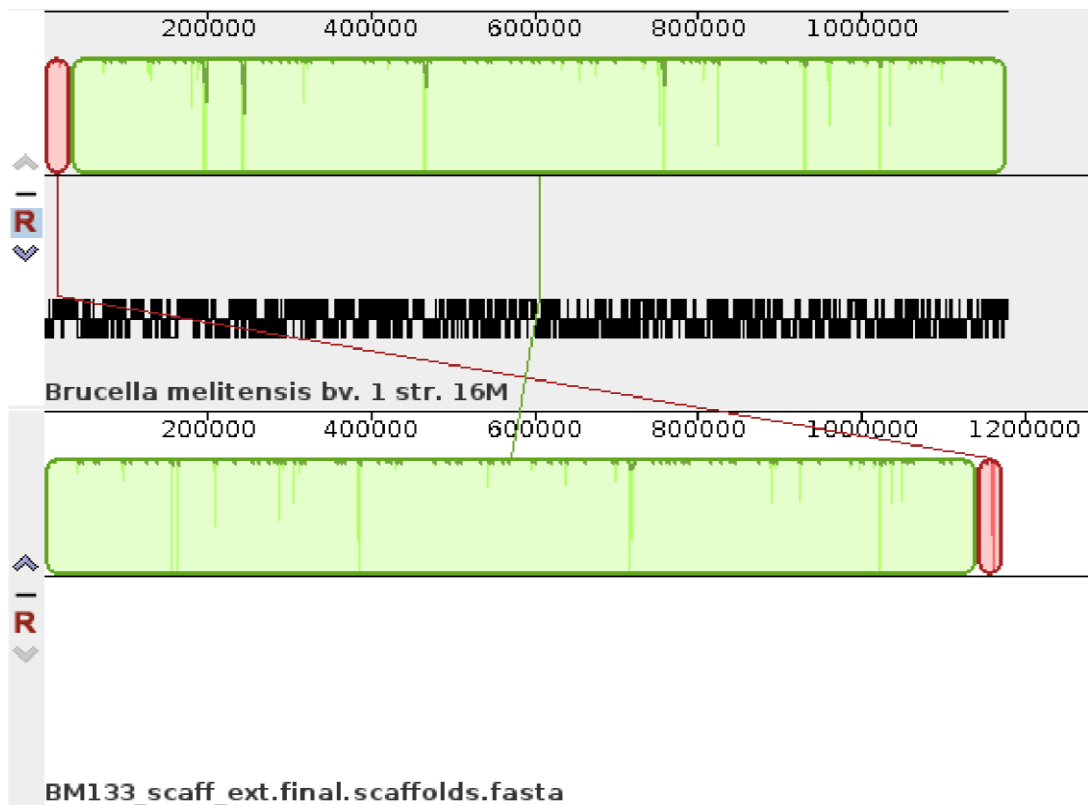


Figura 18. Representación del ordenamiento final de “contigs” del ChII con Mauve. Se utilizó el ensamblado denominado SPADES_def_q22_min18_bm133_contigs, utilizando como referencia la secuencia del ChII de *B. melitensis* 16M en tipo de archivo .gbk completo. Barra superior roja (inicio de la secuencia) y barra superior verde (el total de la secuencia) es la representación del ChII de la cepa 16M. Barra inferior roja (final de la secuencia) y barra inferior verde (el total de la secuencia) es la representación del ordenamiento ensamblado de la cepa Bm133 con respecto al ChII de la cepa de referencia.

Los datos que el resumen del proceso entrega son los siguientes:

Para el ChI el reporte señala que el ensamblado posee 3,276,423 bases en el total del ensamblado, mientras que la referencia tiene 2,117,144 bases solo para el ChI de la cepa 16M. El reporte contiene una lista de 1993 CDS completas, 66 CDS rotas (secuencias codificantes que presentan mutaciones puntuales con cambio de sentido, presencia de codones de paro y espacios o gaps dentro de las CDS), 209 SNP's y 80 gaps en el ensamblado. Por su parte, en el ensamblado del ChII contiene 3,276,423 bases por 1,177,787 de la referencia, así como 1092 CDS completas, 47 CDS rotos, 160 SNP's y 48 gaps.

La suma total de los 2 cromosomas de la cepa de referencia 16M es de 3,294,931 nt mientras que el ensamble de la Bm133 consiste de 3,276,423 nt presentando una diferencia de 18,508 nt. Mediante los reordenamientos y alineamientos genómicos que se realizaron con los archivos gbk que se obtuvieron del GenBank, Mauve genera una calificación del ensamble (Score assembly) en el que se listan las secuencias codificantes rotas (Broken CDS) con los genes anotados en donde se presentaron cambios en la secuencia que llevan a la sustitución de aminoácidos o la aparición de mutaciones sin sentido (presencia de codones de paro). El ChI de la cepa Bm133 de los 66 CDS rotos presentes, en 34 secuencias codificantes se encontraron sustituciones de aminoácidos y una mutación sin sentido en la secuencia CDS BMEI0393 que codifica para una proteína de unión a sustratos (principalmente involucrada en el metabolismo de carbohidratos). Para el caso de las 32 CDS restantes, presentan *gaps* en las CDS, no se incluyeron en la evaluación ya que se tiene que comprobar que estos nos son artefactos de la secuenciación. También se observan cambios en genes involucrados en la síntesis de proteínas ribosomales, *rplB* (50S) y *rpsG* (30S) con un cambio y el gen de la proteína S2 (30S) con dos sustituciones no sinónimas. La CDS BMEI0437, que codifica para un transportador de tipo ABC que pueden tener una relación estrecha con los procesos de virulencia de *Brucella*, es uno de los genes que más cambios no sinónimos presenta, con un total de 7 codones que corresponden a los aa 43, 44, 46, 49, 55, 57 y 58. Dentro de las secuencias que presentan tres sustituciones de codones se encuentran BMEI1606, BMEI0468 y BMEI0892, los cuales corresponden a genes anotados en el GeneBank que codifican para la proteína VsrB, una proteína hipotética y un transportador periplásmico de flujo RND, respectivamente. De estos genes, el transportador de flujo está involucrado en el metabolismo de formación de la membrana externa, mientras que *vsrB* está asociado a mecanismos de transducción de señales. Los genes BMEI0275 y BMEI0574, que codifican para una enzima helicasa que es parte importante del proceso de replicación de la bacteria y la proteína MurE (ligasa de peptidoglicano) relacionada a el metabolismo de formación de la pared celular,

presentan dos sustituciones no sinónimas. El resto de CDS (26 genes) que presentan solo un cambio no sinónimo en un codón incluyen genes para transportadores de diferentes sustratos, genes reguladores, genes involucrados en el proceso de replicación, en el metabolismo del LPS, así como para proteínas hipotéticas de las cuales se desconoce su función y otros son pseudo genes.

Para el ChII de la cepa Bm133, se identificaron 25 secuencias con sustituciones o codones de paro. BMEII0840 presenta un codón de paro en el aa 190 de la proteína (apenas por delante de la mitad de la secuencia de aa [59%]), por lo que este gen, que codifica para una proteína glicosil-transferasa de la familia 2 que está involucrada en el metabolismo de la membrana externa, de forma más concreta en el metabolismo del LPS de la bacteria, ahora se considera inactivo. Las CDS que presentaron sustituciones de tres aa fueron BMEII1024 que codifica para un pseudo gen de 412 pb del que no se posee más información, y BMEII0489, en el cual se substituyen los aa 60, 61 y 69, codifica para la proteína NikC, una permeasa parte del transportador de tipo ABC de níquel. Varios de estos transportadores se han visto asociados al proceso de adaptación de *Brucella* al ambiente celular en macrófagos y células epiteliales. Con cambios de dos aminoácidos tenemos a los genes BMEII0284 y BMEII0437. El primero codifica para la proteína de unión a sustrato de un transportador de tipo ABC; y el segundo para la prenil-transferasa UbiA involucrada en el proceso de formación del Ubiquinona 8 o Q-8 que participa en procesos de descarboxilación del piruvato y transporte de electrones en la membrana citoplasmática. Uno de los cambios en BMEII0284 es el de la Alanina 190 por Prolina, el cual podría provocar una variación en el plegamiento de la proteína y por tanto de su funcionalidad.

Los 20 genes restantes presentan una sustitución simple de aminoácido. Las secuencias BMEII0713 y BMEII0710 se consideran pseudogenes. Por su parte, las secuencias BMEII0033, BMEII0432, BMEII1048 y BMEII0839 son de relevancia para el proceso de patogénesis de la bacteria, BMEII0033 que corresponde al gen *virB9* que codifica para un componente del Sistema de

Secreción Tipo IV de *Brucella*, presenta un cambio de Arginina por Glicina en el extremo carboxilo terminal. Virb9 es considerado parte de uno de los principales factores de virulencia por estar involucrado en el proceso de sobrevivencia intracelular en células eucariontes (O'Cellaghan *et al.*, 1999; Ke *et al.*, 2015; Rossetti *et al.*, 2017). BMEII0432 codifica para un transportador de azúcares de tipo ABC dependiente de ATP; este tipo de transportadores se encuentran ligados estrechamente a los procesos iniciales de internación de la bacteria y las etapas tempranas de maduración de la VCB (Supriya, Umapathy and Ravikumar, 2010; Rossetti *et al.*, 2011, 2017; He, 2012; Yan *et al.*, 2021). BMEII1048 corresponde al gen *groL* parte de la chaperonina GroEL cuya función principal es evitar el plegamiento incorrecto de proteínas bajo condiciones de estrés: también se ha visto involucrada en los procesos de internación de la bacteria (Godefroid *et al.*, 2010; Wenxiao *et al.*, 2016; Salmon-Divon, Zahavi and Kornspan, 2019). Por último, BMEII0839, que codifica para una alfa-N-acetilglucosaminil-fosfotransferasa, está involucrada en el proceso de cebado del bactoprenol y comparte homología con el gen *wecA* de *B. abortus*. Al estar involucrada en la ruta de ensamblaje del LPS, el cual se considera uno de los factores de virulencia de *Brucella*, representa una mutación relevante la cual se deberá estudiar más a fondo. Las demás sustituciones incluyen genes asociados a metabolismo de carbohidratos como galactosa y fructosa, reparación del material genético, metabolismo de nucleótidos metabólicamente activos y metabolismo de vitaminas, entre otros. Las tablas siguientes resumen la información descrita anteriormente (Cuadros 7 y 8).

Es importante mencionar que los datos de secuenciación de la mutante LVM31 se trabajaron siguiendo los mismos procedimientos que se usaron para el ensamblaje del genoma de la cepa Bm133, sin embargo, ya que se obtuvo el ensamble y los contigs ordenados del ChII de la mutante, se observó que el gen *omp31* no estaba interrumpido y que las secuencias pertenecen probablemente

a la cepa Bm133. Por esta razón, se descartó el trabajo con este genoma que se secuenció por un posible error al enviar el material a secuenciar.

CDS con sustituciones de aminoacidos del ChI de la cepa Bm133				
CDS_ID	Pep_Len	Subs_Pos	Subs	Genebank name
BMEI1012	337	275	Y->H	Site-specific integrase
BMEI1059	472	389	A->T	M48 family metalloprotease
BMEI0123	332	11	N->K	Peptidyl-prolyl cis-trans isomerase.
BMEI0288	610	208	L->F	ABC-F family ATP-binding cassette domain-containing protein
BMEI0760	278	9	I->T	<i>rplB</i> 50S ribosomal protein L2
BMEI0275	1029	561,607	L->F,D->A	Helicase
BMEI0274	505	365	C->S	Pseudo-gen
BMEI0393	326	-	-	Substrate-binding domain-containing protein (CHO's)
BMEI0192	296	161	K->T	<i>prmC</i> peptide chain release factor N(5)-glutamine methyltransferase
BMEI0437	286	43,44,46,49,55,57, 58	E->K,S->P,S->G,S->P,V->F,G->V,L->F	ddpD ABC transporter ATP
BMEI0876	824	130	D->H	lon endopeptidase La
BMEI0753	157	53	T->K	<i>rpsG</i> 30S ribosomal protein S7
BMEI1606	1306	1108,1128,1129	T->S,P->Q,X->R	<i>vsrB</i> Transduction signal
BMEI0143	213	84	N->Y	LysE family translocator
BMEI0061	120	61	F->S	hypothetical protein
BMEI0468	87	11,14,18	Q->E,S->R,D->E	hypothetical protein
BMEI0545	354	336	M->L	DUF475 domain-containing protein
BMEI1117	434	140	I->S	adenylosuccinate lyase purB
BMEI0103	428	120	R->H	cystathionine gamma-synthase family protein
BMEI2047	185	90	T->A	hslV – ATP-dependent protease subunit HslV

BMEI1519	428	307	V->I	<i>purD</i> ligasa de fosforibosilamina—glicina
BMEI1957	408	26	V->L	<i>fabB</i> beta-ketoacyl-ACP synthase I
BMEI1912	190	124	V->L	sugar transporter
BMEI0823	267	81,96	Y->F,G->A	30S ribosomal protein S2
BMEI0451	584	293	G->V	<i>leuA</i> -2-isopropylmalate synthase
BMEI0010	294	277	H->R	chromosome partitioning protein parB >>
BMEI1383	180	95	G->D	AraC family transcriptional regulator
BMEI1823	814	493	E->K	<i>gyrB</i> – DNA topoisomerase (ATP-hydrolyzing) subunit B
BMEI0337	151	74	M->I	<i>tolR</i> protein TolR
BMEI1225	140	19	V->A	hypothetical protein BMEI1225
BMEI1587	72	69	V->G	hypothetical protein BMEI1587
BMEI0574	500	7,11	V->G,A->T	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase MurE
BMEI2037	492	260	G->D	phosphoenolpyruvate carboxykinase
BMEI0892	365	127,128,137	M->S,P->A,K->E	efflux RND transporter periplasmic

Cuadro 7. CDS con sustituciones no sinonimas o mutaciones sin sentido del ChI de la cepa Bm133.

El gen BMEI0393 presenta una mutación sin sentido en el nucleótido 193. BMEI0010 es una mutación que se encuentra reportada en la cepa REV-1

CDS con sustituciones de aminoacidos del ChII de la cepa Bm133				
CDS_ID	Pep_Len	Subs_Pos	Subs	Genebank_name
BMEII0033	248	234	R->G	<i>virB9</i> P-type conjugative transfer protein VirB9
BMEII0078	588	210	H->P	2,3-dihydroxybenzoyl adenylate synthase
BMEII1024	413	22,28,43	A->T,P->A,L->F	Pseudogene
BMEII0097	674	27	K->R	Cation-translocating P-type ATPase
BMEII0037	324	72	G->V	DUF459 domain-containing protein
BMEII0157	191	138	N->Y	Lytic transglycosylase domain-containing protein
BMEII0432	515	505	T->A	Sugar ABC transporter ATP-binding protein
BMEII1048	547	284	C->R	<i>groL</i> chaperonin GroEL > >
BMEII0688	1171	612	G->D	<i>mfd</i> transcription-repair coupling factor
BMEII0128	419	197	H->S	Tm-1-like ATP-binding domain-containing protein
BMEII0946	339	211	A->D	Substrate-binding domain-containing protein
BMEII0284	507	189,190	N->K,A->P	ABC transporter substrate-binding protein
BMEII0543	258	244	A->S	SDR family protein
BMEII0840	319	-	-	Glycosyltransferase family 2 protein
BMEII0423	355	159	A->G	<i>fb</i> a fructose-bisphosphate aldolase class II
BMEII0489	285	60,61,69	R->W,V->L,N->D	Nickel ABC transporter permease subunit NikC
BMEII1096	452	284	H->Y	MFS transporter
BMEII1053	413	35	S->W	Sugar MFS transporter
BMEII0380	391	190	E->G	Efflux RND transporter periplasmic adaptor subunit
BMEII0839	355	122	M->L	Undecaprenyl/decaprenyl-phosphate alpha-N-acetylglucosaminyl 1-phosphate transferase
BMEII0437	480	206,254	V->L,H->Q	UbiA family prenyltransferase

BMEII0713	373	220	A->S	Pseudogene
BMEII0710	122	54	S->P	Pseudogene
BMEII0776	379	219	H->N	8-amino-7-oxononanoate synthase bioF
BMEII0617	282	144	P->R	Xanthine/uracil/vitamin C permease, AzgA family

Cuadro 8. CDS con sustituciones no sinónimas o mutaciones sin sentido del ChII de la cepa Bm133.

En el gen BMEII0840 tiene un codón de paro en el nucleótido 190. El gene BMEII1048 mutación ya reportada para la cepa REV-1

6.10 Anotación de los genes de la cepa de referencia mexicana de *B. melitensis* Bm133

Con los “*contigs*” ordenados para cada cromosoma de la cepa Bm133 se procedió a separarlos por cromosomas con la herramienta `faSomeRecords.py`, la cual permite retirar de un archivo de texto plano una lista previamente especificada de elementos del archivo. Gracias a que los archivos de salida de Mauve incluyen un archivo `tab` con los *contigs* ordenados, aunque no separados, fue más sencillo separarlos ya que mediante el despliegue gráfico del programa, se anotó cuál era el último “contig” del reordenamiento y, con esto, se preparó la lista de contigs necesarios para que el programa separase las secuencias enlistadas. Aunque se eligieron los contigs que el programa Mauve indicó que formaban parte de cada cromosoma para la cepa Bm133, cabe mencionar que se tomó un contig extra de la secuencia para cada cromosoma, ya que en el nodo final una pequeña sección sí alineaba con las secuencias de cada cromosoma. Por lo que para cada secuencia de los cromosomas se tiene secuencias extras. Con los “contigs” separados en los cromosomas ChI y ChII se procedió a realizar un `mummerplot` en la plataforma Galaxy entre la cepa Bm133 y la cepa 16M (Figuras 21 y 22).

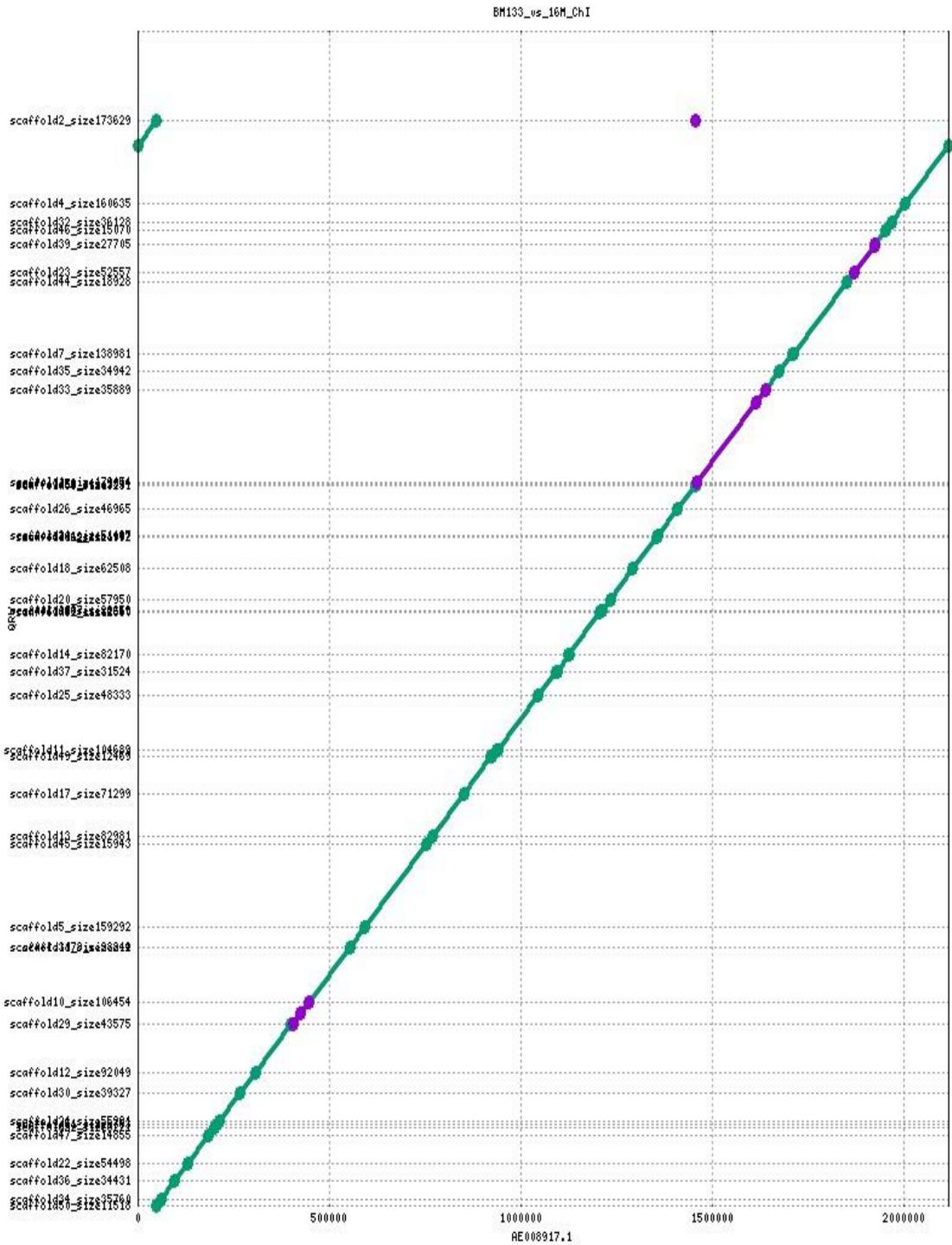


Figura 19. Gráfica de similitud del ensamblado ordenado del ChI de Bm133 vs el ChI 16M.

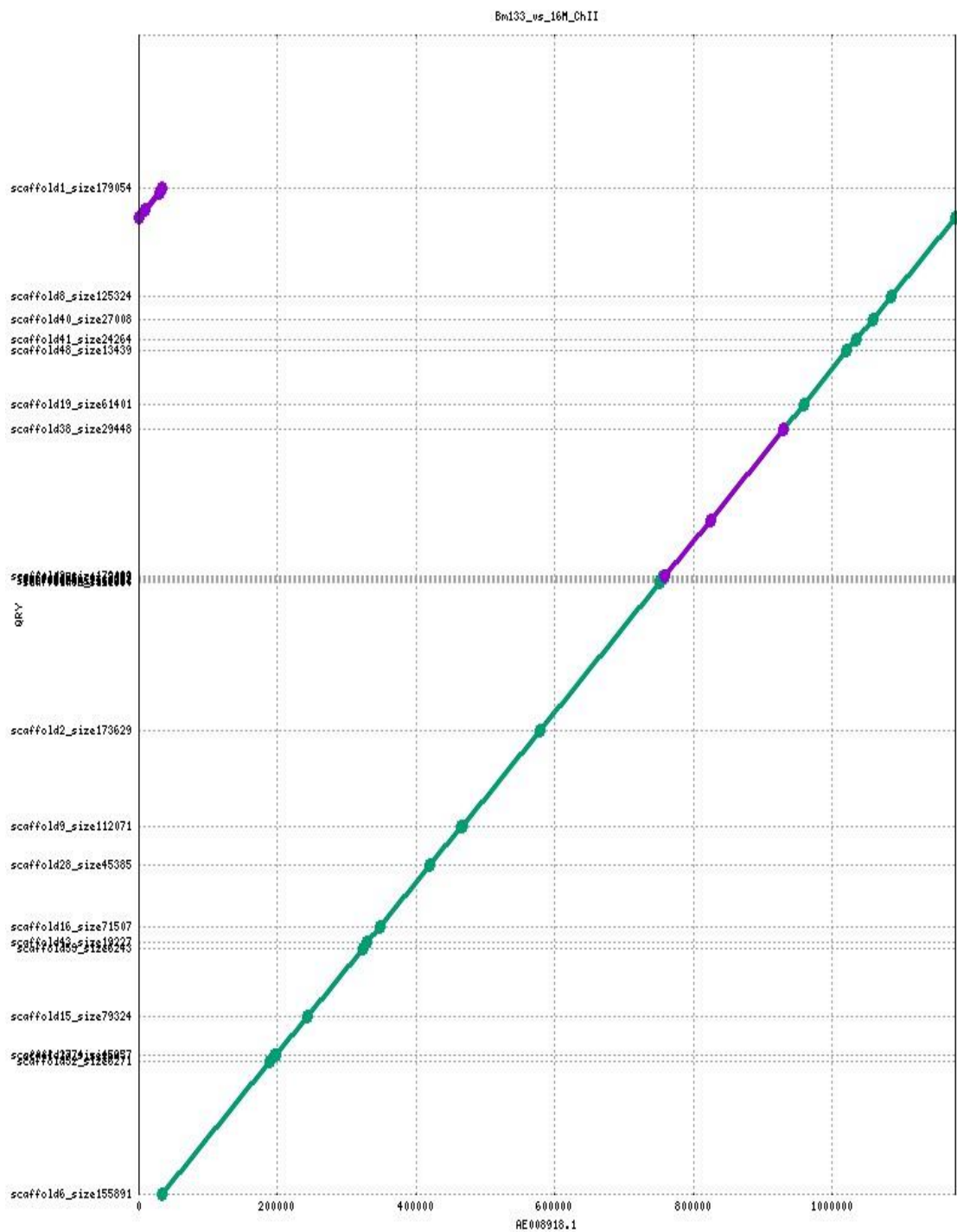


Figura 20. Gráfica de similitud del ensamble ordenado del ChI de Bm133 vs el ChI 16M.

Después de verificar la linealidad del ensamble, se procedió a la anotación mediante la herramienta Prokka que se encuentra asociada a la plataforma en línea Galaxy. Esto se realizó para cada cromosoma, obteniendo los siguientes datos preliminares. Para el ChI de la cepa Bm133 el número de contigs usados en la anotación fue de 45; número de bases presentes 2,272,809 lo que representa el 7.3% más de bases de las que tiene el cromosoma ChI de la cepa de referencia 16M. Esto indica que aún le falta refinamiento a la separación de contigs. El número de CDS fue de 2180; 3 genes ribosomales; 42 tRNA; y 1 tmRNA. En el caso del ChII 23 “contigs” fueron usados para ChII; el número de bases presentes en el ensamble en la anotación fue de 1,353,704, lo que representa un 14 % más que las bases con respecto a la referencia. Las CDS presentes en el ChII fue de 1275; 3 genes ribosomales; y 16 tRNA.

Con estos resultados es necesario eliminar las bases extras que se encuentran en el “contig” final para cada ordenamiento de los cromosomas para con ello eliminar las bases extra que pudieran estar introduciendo un mayor número de genes del que realmente se logró en el ensamble.

A los archivos de los “contigs” ordenados de la bacteria se les realizó una corrida con el programa QUAST con el cual se pudo verificar de mejor manera la cantidad de bases que realmente pertenecen a cada cromosoma en sus alineamientos. Obteniendo los siguientes resultados: el ChI de la cepa Bm133 tuvo un total de nt en el ensamble de 2,098,522 de la referencia, lo que corresponde a un 99.084 % generado a partir de un total de 43 “contigs”; el más largo fue de 179,054 nucleótidos. Para el ChII de la cepa Bm133 se tuvo un total de 1,173,629 nt del total 1,177,787 nt de la cepa de referencia, lo que representa un 99.6 % del total del cromosoma; esto se logró usando 20 contigs, el “contig” más largo fue de 173,629 nt.

7 Discusión

Para definir las condiciones experimentales que nos permitan realizar un análisis de transcriptómica comparativa entre las cepas de *B. melitensis* 16M, cepa de referencia para la cual se reportó un análisis de expresión diferencial en comparación con la cepa vacunal REV-1 mediante el uso de microarreglos de DNA (Rossetti *et al.*, 2009) y la Bm133 (aislado mexicano de referencia) y su mutante $\Delta omp31$ LVM31, en condiciones de cultivo líquido. En este trabajo se analizó la cinética de crecimiento de estas cepas. Como se describe en la sección de resultados, las tres cepas muestran cinéticas de crecimiento distintas, particularmente la cepa LVM31, un fenotipo asociado probablemente a la ausencia del gen *omp31*. Nuestras observaciones son consistentes con reportes previos que indican diferencias de crecimiento en medios de cultivo de laboratorio, así como en invasividad en células en cultivo (González-Rodríguez, 2012; Hernandez-Pineda, 2012; Alva-Pérez *et al.*, 2014; Verdiguél-Fernández *et al.*, 2017, 2020).

A su vez, se estandarizó un procedimiento reproducible para la extracción de RNA total para las cepas en estudio y un ensayo de RT-PCR punto final del gen *omp31* que permitiría la detección de hasta 2,070 bacterias en promedio. Debido a causas ajenas al grupo de investigación no se pudieron realizar los ensayos de infección de macrófagos humanos con las tres cepas de *B. melitensis* y por consiguiente tampoco el proceso de amplificación selectiva de mRNA y los estudios de microarreglos de DNA de *Brucella* que eran parte de los objetivos del trabajo, los resultados obtenidos representan las bases metodológicas necesarias para que el proyecto de evaluación del transcriptoma tenga continuidad, aportando conocimiento sobre la patogenia de la cepa de referencia mexicana Bm133 y su mutante LVM31.

Como alternativa para avanzar en el conocimiento de estas cepas, se realizó la secuenciación, ensamble y análisis del genoma (ambos cromosomas) de la cepa

de referencia mexicana de *B. melitensis* Bm133. Este objetivo permitió la introducción del grupo al análisis bioinformático a través del cual se aprendió a aplicar diferentes herramientas de análisis y a establecer un correcto flujo de trabajo que dio lugar a una primera anotación y análisis comparativo de la cepa Bm133 con la cepa de referencia 16M.

El análisis de las lecturas con FastQC (Leggett *et al.*, 2013) reportó calidades por encima de 30 con un promedio de 120 nt por lectura y 316,996 lecturas para ambos archivos. La cobertura que se obtuvo fue de 33.55 X dentro de los parámetros aceptados para un ensamble genómico bacteriano, sin embargo, no se obtendría al 100 % el genoma y mucho menos la posibilidad de la circularización. Una secuenciación de lecturas largas (Nanopore de Oxford o SMRT de Pacbio) podría permitir un ensamble híbrido para circularizar cada cromosoma (Salmon-Divon *et al.*, 2017; Wick *et al.*, 2017).

La edición del genoma nos permitió quitar las secuencias de adaptadores, secuencias repetidas, y mejorando la calidad de las lecturas sin perder >5 % de las lecturas (Martin, 2021). La herramienta QUAST (Gurevich *et al.*, 2013) permitió la evaluación del ensamblado para ambos ChI y ChII para la cepa Bm133, se logró completar el 99.291% del ChI de la Bm133; mientras para el ChII el porcentaje de completitud fue de 99.8 %. El uso de Nucmer y Mummerplot (Kurtz *et al.*, 2004) permitieron verificar los porcentajes de identidad de los contigs en porcentajes de identidad, con un mínimo de 98% y un máximo de 100%. El despliegue de los datos en matriz nos permitió observar la linearidad y similitud de las secuencias de las cepas Bm133 contra la cepa 16M. El programa Mauve se utilizó para la separación y re-ordenamiento de contigs, realizando corridas por separado para cromosoma. Los contigs fueron separados por cromosoma con faSomeRecord.py sin ser perfecto. En ambos archivos fasta se obtuvo un contig extra, que solo pertenecía a una parte de la secuencia del final. En la anotación con el programa Prokka (Seemann, 2014) los valores de CDS y

el total de nt utilizados en la anotación, aumentaron con respecto a los obtenidos en el programa Mauve.

Mediante la herramienta Mauve se encontraron polimorfismos de nucleótido simple (SNP's), Gaps o huecos en el genoma de la cepa Bm133 comparada contra la referencia 16M los cuales deberán ser revisados más a fondo para determinar cuáles de estos son artificios generados por el proceso de secuenciación o por el proceso de ensamblaje.

En el ChI de la cepa de referencia mexicana Bm133, la secuencia BMEI0393 que codifica para una proteína de unión a sustrato del transporte de ribosas para uso de la bacteria por parte de un transportador de tipo ABC asociado al metabolismo de monosacáridos; presenta un codón de paro en el aminoácido 193 el cual inactiva la función de esta proteína por completo, sin embargo *Brucella melitensis* presenta en su genoma tres diferentes transportadores de tipo ABC que se especializan en el transporte de ribosa, que podrían suplir la función perdida para este gen, siempre y cuando las condiciones de expresión de los otros dos transportadores sean las mismas o presenten homología con el transportador mutado (Jenner *et al.*, 2009). La secuencia BMEI0288 codifica para una proteína de la familia de transportadores del tipo ABC que presenta solo un cambio de aminoácido en la posición 208 (de Leucina a Fenilalanina), este gen es asociado con una función reguladora de resistencia a antibióticos. El gen BMEI0437 presenta 7 aminoácidos con sentido diferente en las posiciones 43, 44, 46, 49, 55, 57 y 58 y los cambios fueron los siguientes: E→K, S→P, S→G, S→P, V→F, G→V y L→F. Esta secuencia codifica para el gen *ddpD* parte de la maquinaria de importación del transportador y que está involucrada al transporte de di-péptidos y el metabolismo del níquel (Jenner *et al.*, 2009).

De la misma manera en el ChII de la cepa Bm133 se obtuvieron tres secuencias que presentaban cambios en la secuencia de aminoácidos la cuales fueron: BMEII0432, BMEII0284, BMEII0489. BMEII0432 codifica en su secuencia para un transportador de tipo ABC involucrado en el transporte transmembranal del

sustrato eritritol uno de los sustratos importantes dentro del metabolismo de la bacteria, este transportador solo presenta un cambio en el aminoácido 505, de una proteína de tamaño 515 aa en total; el cambio presente es de Triptofano a Alanina (Lopez-Goñi and Moriyón, 2005; Jenner *et al.*, 2009).

BMEI0284 codifica para una proteína de unión a sustrato, de un transportador ABC asociada al transporte de di-péptidos y el metabolismo del níquel. Presenta el cambio de dos aminoácidos en las posiciones 189 y 190 que cambian de N→K y A→P. El gen BMEI0489 presentó la sustitución de 3 aminoácidos en las posiciones 60, 61 y 69 cambiando los aa de la siguiente manera R→W, V→L y N→D. Esta secuencia codifica para el gen *nikC* parte del sistema de transporte de níquel, requerido en eventos de baja oxigenación y su represor transcripcional es el gen *nikR* que es el encargado de activar al gen *ureA* el cual se ha visto que se encuentra sobre-expresado a la 12 h postinfección de células HeLa (Lopez-Goñi and Moriyón, 2005; Jenner *et al.*, 2009; Rossetti *et al.*, 2010) .

Dentro del metabolismo de síntesis de pared celular y de envoltura se encontraron tres genes modificados en el ChI (BMEI0574) y 2 en el ChII (BMEI0840 y BMEI0839) (Ratushna *et al.*, 2006; Salmon-Divon, Yeheskel and Kornspan, 2018). BMEI0574 codifica para la ligasa de peptidoglicano denominada MurE. Salmon-Divon *et al.*, (2018) reportaron una mutación en el gen MurE de la cepa vacunal Rev-1. BMEI0840 codifica para una proteína de la familia 2 de las glicosiltransferasas que están asociadas al metabolismo de síntesis de pared celular (Ratushna *et al.*, 2006), este gen presenta una mutación sin sentido en el aa 190 de la secuencia que se encuentra a la mitad del gen. Por lo que se puede considerar que este gen queda inactivado. El gen BMEI0839 presentó una mutación de sentido erróneo en la posición 122 de Metionina por Lisina. Este gen se encuentra involucrado de forma principal en el metabolismo de síntesis de envoltura celular y que presenta homología con el gen *wecA* de *Escherichia coli* (Lopez-Goñi and Moriyón, 2005; Ratushna *et al.*, 2006; González *et al.*, 2008) . Estos cambios genéticos podrían afectar de alguna

forma el metabolismo de membrana externa de la bacteria el cual es considerado uno de los principales factores de virulencia.

La biogénesis de la vacuola contenedora de *Brucella* (BCV por sus siglas en inglés) depende en gran medida del Sistema de Secreción Tipo IV, codificado en el operón *VirB* (O'Cellaghan *et al.*, 1999; Myeni *et al.*, 2013; Rossetti *et al.*, 2017). Mutantes en el operón *vir* han demostrado ser incapaces de sobrevivir y replicarse en las células hospederas como macrófagos humanos THP-1 o periféricos (O'Cellaghan *et al.*, 1999). En el ChII de la cepa Bm133 se encontró que el gen *virB9* (BMEII0033) presentó una mutación de sentido incorrecto en el aa 234 de la secuencia sufriendo un cambio de Arginina a Glicina. Dicha mutación deberá investigarse a futuro para identificar si afecta la actividad de la proteína.

Las secuencias BMEI0010 y BMEII1048 son un hallazgo dentro de la secuencia de relevancia, estas mutaciones fueron descritas para la cepa vacunal Rev-1 en su reciente secuenciación y análisis del genoma que se realizó en los años 2017 y 2018 (Salmon-Divon *et al.*, 2017; Salmon-Divon, Yeheskel and Kornspan, 2018) . BMEI0010 codifica para la proteína ParB que se reporta asociada al sitio de origen de replicación *oriC* por lo que es una proteína que ayuda al proceso de crecimiento y división celular en *Brucella* (De Bolle *et al.*, 2015). En estudios comparativos de proteoma entre las cepas 16M y REV-1 se demostró la presencia de esta proteína para REV-1 no así para la 16M, Salmon-Divon *et al.*, (2018) reportaron una mutación de sentido erróneo para la proteína ParB, el aminoácido que cambió fue el 277 y la modificación que se produjo fue una Histidina por una Arginina; esta misma mutación es la que se reporta en este trabajo para la proteína ParB. Lo que permite especular si la mutación tendrá relevancia en la diferencia en la curva de crecimiento de la cepa Bm133 y para la concentración de UFC's por unidad de tiempo. Por último, la CDS BMEII1048 situada en el ChII codifica para la Chaperonina GroEL, proteína que presenta actividad de chaperona, que facilita el plegamiento proteico, previene la

desnaturalización de proteínas y puede estar involucrada en varios procesos celulares desde el crecimiento bacteriano hasta procesos de transcripción o reparación por mutagénesis por luz UV. Salmon-Divon *et al.*, (2019) reportaron para la cepa REV1 que la proteína GroEL mostró una regulación negativa en condiciones de crecimiento de pH bajo cuando se comparó con la cepa 16M, en el reporte los autores encuentran el argumento que una posible acumulación de proteínas mal plegadas, debido a la baja de transcripción de la chaperona de GroEL, podría llevar a la atenuación de la bacteria. En el reporte previo (Salmon-Divon, Yeheskel and Kornspan, 2018) donde se realiza el análisis genético de la cepa REV-1 se encontró una mutación de sentido erróneo en la secuencia de la proteína en el aminoácido 284 de la chaperona, reportando un cambio de una Cisteína por una Arginina (C→R). Esta misma mutación de sentido erróneo fue encontrada en la cepa Bm133 como se explicó en los resultados de este trabajo. Con esto tenemos dos marcadores genéticos similares, que ya han sido reportados en la cepa atenuada REV-1, estos dos marcadores son importantes para el estudio y comprensión de aquellas posibles diferencias genéticas que hacen que la cepa Bm133 se comporte de una manera diferente a la cepa de referencia 16M, aunque sin perder su virulencia por completo. Por todo lo anterior realizar una comparación entre el genoma reportado de la REV-1 y la cepa Bm133 se vuelve relevante para entender por qué la cepa mexicana mantiene su virulencia, aunque no de la misma manera que la cepa 16M.

8 Conclusiones

1. En este trabajo se establecieron las horas a las cuales las cepas de *Brucella melitensis* 16M, Bm133 y LVM31 inician la fase logarítmica tardía (16 h, 18 h y 28 h respectivamente).
2. Se estandarizó el procedimiento de Extracción de RNA total para las cepas de *Brucella melitensis* que nos permitirá continuar con trabajos de ensayos de transcriptómica a futuro.
3. Se logró la detección de material genético (RNA) mediante la estandarización de una RT-PCR punto final para el gen *omp31* de 712 pb, con una cantidad mínima detectable de RNA de 207.7 picogramos por microlitro. Lo que permitirá verificar la presencia de RNA bacteriano en muestras de RNA total obtenidas de ensayos de infección.
4. Se secuenció y se ensambló el genoma de la cepa de referencia mexicana de *B. melitensis* Bm133, consiguiendo un 99.084 % de total del ChI y un 99.6 % del ChII con respecto a la cepa *B. melitensis* 16M. Que no ha sido reportado en la literatura, esto permitirá estudiar el genoma y sus interacciones en diferentes experimentos de transcriptómica diferencial.
5. Se demostró la presencia de SNP's, Gaps y CDS rotos en la secuencia de la cepa Bm133. Algunas de estas mutaciones sucedieron en transportadores de tipo ABC, en genes involucrados en el metabolismo de envoltura de la bacteria y genes de virulencia.
6. Se encontraron dos mutaciones en los genes *parB* y *groL* que también fueron reportadas en la cepa vacunal REV-1, lo que podría ayudarnos a correlacionar las diferencias en la cinética de crecimiento entre las cepas 16M y la Bm133.
7. Se realizó un borrador de la anotación de los cromosomas I y II de la cepa Bm133, que deberá ser trabajado antes de su anotación en el GenBank.

9 Prospectiva

1. Completar las secuencias genómicas de la cepa de *B. melitensis* Bm133 mediante secuenciación de lecturas largas para que se pueda circularizar, mediante el uso de la herramienta Unicycler.
2. Refinar y completar las anotaciones genéticas realizadas para la cepa de *B. melitensis* Bm133, así como confirmar los cambios encontrados en CDS de la cepa Bm133 con respecto a la cepa 16M mediante su amplificación por PCR y secuenciación de los amplificadores.
3. Realizar la secuenciación y análisis del genoma de la cepa de *B. melitensis* LVM31, para realizar estudios de genómica comparativa con la cepa parental Bm133 para determinar si, además de la eliminación del gen *omp31*, presenta mutaciones adicionales.
4. Realizar ensayos de evaluación del transcriptoma diferencial de las cepas de *B. melitensis* 16M, Bm133 y LVM31 de su fase logarítmica tardía en caldo Brucella a pH neutro o creciendo a pH ácido mediante el uso de microarreglos de DNA, para conocer las diferencias genéticas iniciales con las que se realizan los ensayos de infección
5. Evaluar la dinámica de la infección de las cepas de *Brucella melitensis* 16M, Bm133 y LVM31 en macrófagos THP-1 a las dos y cuatro horas post infección, así como el perfil de expresión genética.

10 Bibliografía

Afgan, E. *et al.* (2018) 'The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update', *Nucleic Acids Research*, 46(W1), pp. W537–W544. Available at: <https://doi.org/10.1093/nar/gky379>.

Alva-Pérez, J. *et al.* (2014) 'The *invA* gene of *Brucella melitensis* is involved in intracellular invasion and is required to establish infection in a mouse model', *Virulence*, 5(5), pp. 563–574. Available at: <https://doi.org/10.4161/viru.28589>.

Bankevich, A. *et al.* (2012) 'SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing', *Journal of Computational Biology*, 19(5), pp. 455–477. Available at: <https://doi.org/10.1089/cmb.2012.0021>.

De Bolle, X. *et al.* (2015) 'Brucella abortus Cell Cycle and Infection Are Coordinated', *Trends in Microbiology*, 23(12), pp. 812–821. Available at: <https://doi.org/10.1016/j.tim.2015.09.007>.

Celli, J. (2015) 'The changing nature of the Brucella-containing vacuole', *Cellular Microbiology*, 17(7), pp. 951–958. Available at: <https://doi.org/10.1111/cmi.12452>.

Corbel, M.J. (2006) *Brucellosis in humans and animals*, WHO Library catalogue in publication Data. Edited by [WHO] World Health Organization, W.O. for A.H. OIE, and L. Geneva: WHO Press.

Dirección de Campañas Zoonositarias and SENASICA (2017) *Datos sobre la vacunación contra la brucelosis*.

Ficht, T. (2011) 'Brucella taxonomy and evolution', *Future microbiology*, 5(6), pp. 859–866. Available at: <https://doi.org/10.2217/fmb.10.52.Brucella>.

De Figueiredo, P. *et al.* (2015) 'Pathogenesis and immunobiology of brucellosis: Review of Brucella-host interactions', *American Journal of Pathology*, 185(6), pp. 1505–1517. Available at: <https://doi.org/10.1016/j.ajpath.2015.03.003>.

Garrity, G.M. (2005) *BERGEY'S MANUAL OF Systematic Bacteriology*. 2nd Editio. Edited by D.J. Brenner, N.R. Krieg, and J.T. Staley. East Lansing: Springer US.

Godefroid, M. *et al.* (2010) 'Brucella melitensis 16M produces a mannan and other extracellular matrix components typical of a biofilm', *FEMS Immunology and Medical Microbiology*, 59(3), pp. 364–377. Available at: <https://doi.org/10.1111/j.1574-695X.2010.00689.x>.

Golshani, M. and Buozari, S. (2017) 'A Review of Brucellosis in Iran: Epidemiology, Risk Factors, Diagnosis, Control, and Prevention', *Department of Molecular Biology, Pasteur Institute of Iran, Tehran, Iran*, 21(x), pp. 8–18. Available at: <https://doi.org/10.18869/acadpub.ibj.21.6.349>.

- González-Rodríguez, D. (2012) *Expresion de la proteina sintaxina 4, durante la invasion de macrofagos humanos de la linea thp-1 por brucella melitensis*. Universidad Nacional Autónoma de México.
- González, D. *et al.* (2008) 'Brucellosis vaccines: Assessment of *Brucella melitensis* lipopolysaccharide rough mutants defective in core and O-polysaccharide synthesis and export', *PLoS ONE*, 3(7). Available at: <https://doi.org/10.1371/journal.pone.0002760>.
- Gorvel, J.P. and Moreno, E. (2002) 'Brucella intracellular life: From invasion to intracellular replication', *Veterinary Microbiology*, 90(1–4), pp. 281–297. Available at: [https://doi.org/10.1016/S0378-1135\(02\)00214-6](https://doi.org/10.1016/S0378-1135(02)00214-6).
- Gurevich, A. *et al.* (2013) 'QUAST: Quality assessment tool for genome assemblies', *Bioinformatics*, 29(8), pp. 1072–1075. Available at: <https://doi.org/10.1093/bioinformatics/btt086>.
- Hernández-Castro, R. *et al.* (2008) 'The BMEI0216 gene of *Brucella melitensis* is required for internalization in HeLa cells', 44, pp. 28–33. Available at: <https://doi.org/10.1016/j.micpath.2007.08.008>.
- Hernandez-Pineda, J.A. (2012) *Efecto del bloqueo de la expresion de sintaxina 4 sobre el transito intracelular de brucella melitensis.pdf*. Universidad Nacional Autónoma de México.
- Hör, J., Gorski, S.A. and Vogel, J. (2018) 'Bacterial RNA Biology on a Genome Scale', *Molecular Cell*, 70(5), pp. 785–799. Available at: <https://doi.org/10.1016/j.molcel.2017.12.023>.
- Jenner, D.C. *et al.* (2009) 'ATP-binding cassette systems of *Brucella*', *Comparative and Functional Genomics*, 2009. Available at: <https://doi.org/10.1155/2009/354649>.
- Jiang, H. *et al.* (2013) 'Comparative Genomic Analysis of *Brucella melitensis* Vaccine Strain M5 Provides Insights into Virulence Attenuation', *PLoS ONE*, 8(8). Available at: <https://doi.org/10.1371/journal.pone.0070852>.
- Ke, Y. *et al.* (2015) 'Type IV secretion system of *Brucella* spp. and its effectors', *Frontiers in Cellular and Infection Microbiology*, 5(October), pp. 1–10. Available at: <https://doi.org/10.3389/fcimb.2015.00072>.
- Kornspan, D. *et al.* (2020) 'Genomic analysis of natural rough *brucella melitensis* rev.1 vaccine strains: Identification and characterization of mutations in key genes associated with bacterial lps biosynthesis and virulence', *International Journal of Molecular Sciences*, 21(24), pp. 1–15. Available at: <https://doi.org/10.3390/ijms21249341>.
- Kurtz, S. *et al.* (2004) 'Versatile and open software for comparing large genomes', *Genome biology*, 5(2), p. 12. Available at: <http://www.tigr.org/software/mummer>.
- Leggett, R.M. *et al.* (2013) 'Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics', *Frontiers in Genetics*, 4(DEC), pp. 1–5.

Available at: <https://doi.org/10.3389/fgene.2013.00288>.

Lopez-Goñi, I. and Moriyón, I. (2005) *Brucella: Molecular and Cellular Biology, Journal of Antimicrobial Chemotherapy*. Available at: <https://doi.org/10.1093/jac/dki211>.

Martin, M. (2021) 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal*, 17(1), pp. 10–12. Available at: <https://doi.org/doi:https://doi.org/10.14806/ej.17.1.200>.

Myeni, S. *et al.* (2013) 'Brucella Modulates Secretory Trafficking via Multiple Type IV Secretion Effector Proteins', *PLoS Pathogens*, 9(8). Available at: <https://doi.org/10.1371/journal.ppat.1003556>.

O'Cellaghan, D. *et al.* (1999) 'A homologue of the *Agrobacterium tumefaciens* VirB and *Bordetella pertussis* Ptl type IV secretion systems is essential for intracellular survival of *Brucella suis*', *Molecular Microbiology*, 33(6), pp. 1210–1220. Available at: <https://doi.org/10.1046/j.1365-2958.1999.01569.x>.

Pappas, G. *et al.* (2006) 'The new global map of human brucellosis.', *The Lancet infectious diseases*, 6(2), pp. 91–99. Available at: [https://doi.org/10.1016/S1473-3099\(06\)70382-6](https://doi.org/10.1016/S1473-3099(06)70382-6).

Ratushna, V.G. *et al.* (2006) 'Molecular targets for rapid identification of *Brucella* spp', *BMC Microbiology*, 6. Available at: <https://doi.org/10.1186/1471-2180-6-13>.

Rissman, A.I. *et al.* (2009) 'Reordering contigs of draft genomes using the Mauve Aligner', *Bioinformatics*, 25(16), pp. 2071–2073. Available at: <https://doi.org/10.1093/bioinformatics/btp356>.

Rossetti, C.A. *et al.* (2009) 'Brucella melitensis global gene expression study provides novel information on growth phase-specific gene regulation with potential insights for understanding Brucella:host initial interactions', *BMC Microbiology*, 9(1), p. 81. Available at: <https://doi.org/10.1186/1471-2180-9-81>.

Rossetti, C.A. *et al.* (2010) 'Selective amplification of *Brucella melitensis* mRNA from a mixed host-pathogen total RNA'.

Rossetti, C.A. *et al.* (2013) 'Systems biology analysis of *Brucella* infected Peyer's patch reveals rapid invasion with modest transient perturbations of the host transcriptome', *PLoS ONE*, 8(12). Available at: <https://doi.org/10.1371/journal.pone.0081719>.

Rossetti, C.A. *et al.* (2017) 'Systems Biology Analysis of Temporal In vivo *Brucella melitensis* and Bovine Transcriptomes Predicts host:Pathogen Protein-Protein Interactions.', *Frontiers in microbiology*, 8(July), p. 1275. Available at: <https://doi.org/10.3389/fmicb.2017.01275>.

SAGARPA (1995) 'Nom-041-Zoo-1995'.

Salmon-Divon, M. *et al.* (2017) 'Complete Genome Sequence of the Live Attenuated

Vaccine Strain *Brucella melitensis* Rev.1', *Genome Announcements*, (8), pp. 4–5.

Salmon-Divon, M., Yeheskel, A. and Kornspan, D. (2018) 'Genomic analysis of the original elberg *brucella melitensis* rev.1 vaccine strain reveals insights into virulence attenuation', *Virulence*, 9(1), pp. 1436–1448. Available at: <https://doi.org/10.1080/21505594.2018.1511677>.

Salmon-Divon, M., Zahavi, T. and Kornspan, D. (2019) 'Transcriptomic Analysis of the *Brucella melitensis* Rev.1 Vaccine Strain in an Acidic Environment: Insights Into Virulence Attenuation', *Frontiers in Microbiology*, 10(February), p. 250. Available at: <https://doi.org/10.3389/fmicb.2019.00250>.

Seemann, T. (2014) 'Prokka: Rapid prokaryotic genome annotation', *Bioinformatics*, 30(14), pp. 2068–2069. Available at: <https://doi.org/10.1093/bioinformatics/btu153>.

Servicio Nacional de Sanidad Inocuidad y Calidad Agroalimentaria (2016) *Situación actual del control de la brucelosis en México*. Available at: <https://www.gob.mx/senasica/documentos/situacion-actual-del-control-de-la-brucelosis-en-mexico>.

Verdiguél-Fernández, L. *et al.* (2017) 'Omp31 plays an important role on outer membrane properties and intracellular survival of *Brucella melitensis* in murine macrophages and HeLa cells', *Archives of Microbiology*, 199(7), pp. 971–978. Available at: <https://doi.org/10.1007/s00203-017-1360-7>.

Verdiguél-Fernández, L. *et al.* (2020) 'Brucella melitensis omp31 mutant is attenuated and confers protection against virulent brucella melitensis challenge in BALB/c mice', *Journal of Microbiology and Biotechnology*, 30(4), pp. 497–504. Available at: <https://doi.org/10.4014/JMB.1908.08056>.

Verdiguél-Fernández, L.F. (2014) *Construcción de dos mutantes en los genes rpsI y omp31 de brucella melitensis*. Universidad Nacional Autónoma de México.

Verdiguél-Fernández, L.F. (2019) *Construcción y caracterización de mutantes en los genes omp31, omp22 y rpsL de Brucella melitensis*. UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO.

Viadas, C. *et al.* (2010) 'Transcriptome analysis of the *Brucella abortus* BvrR/BvrS two-component regulatory system', *PLoS ONE*, 5(4), pp. 1–8. Available at: <https://doi.org/10.1371/journal.pone.0010216>.

Wang, F. *et al.* (2013) ' Comparison of Genomes of *Brucella melitensis* M28 and the B. melitensis M5-90 Derivative Vaccine Strain Highlights the Translation Elongation Factor Tu Gene *tuf2* as an Attenuation-Related Gene ', *Infection and Immunity*, 81(8), pp. 2812–2818. Available at: <https://doi.org/10.1128/iai.00224-13>.

Wenxiao, L. *et al.* (2016) 'RNA-seq reveals the critical role of CspA in regulating *Brucella melitensis* metabolism and virulence', *Science China Life Sciences*, 59(4), pp. 417–424. Available at: <https://doi.org/10.1007/s11427-015-4981-6>.

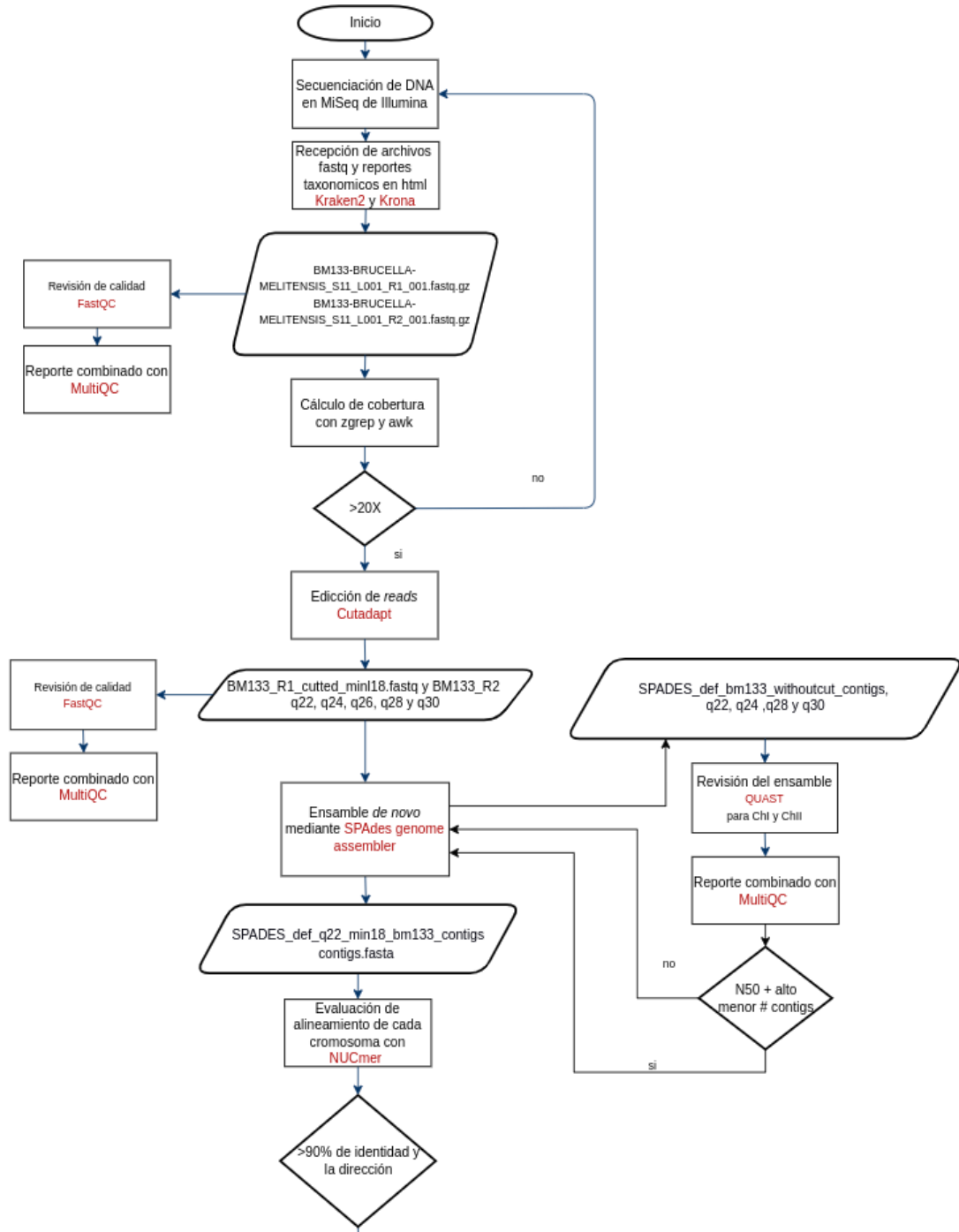
Westermann, A.J., Gorski, S.A. and Vogel, J. (2012) 'Dual RNA-seq of pathogen and host', *Nature Reviews Microbiology*, 10(9), pp. 618–630. Available at: <https://doi.org/10.1038/nrmicro2852>.

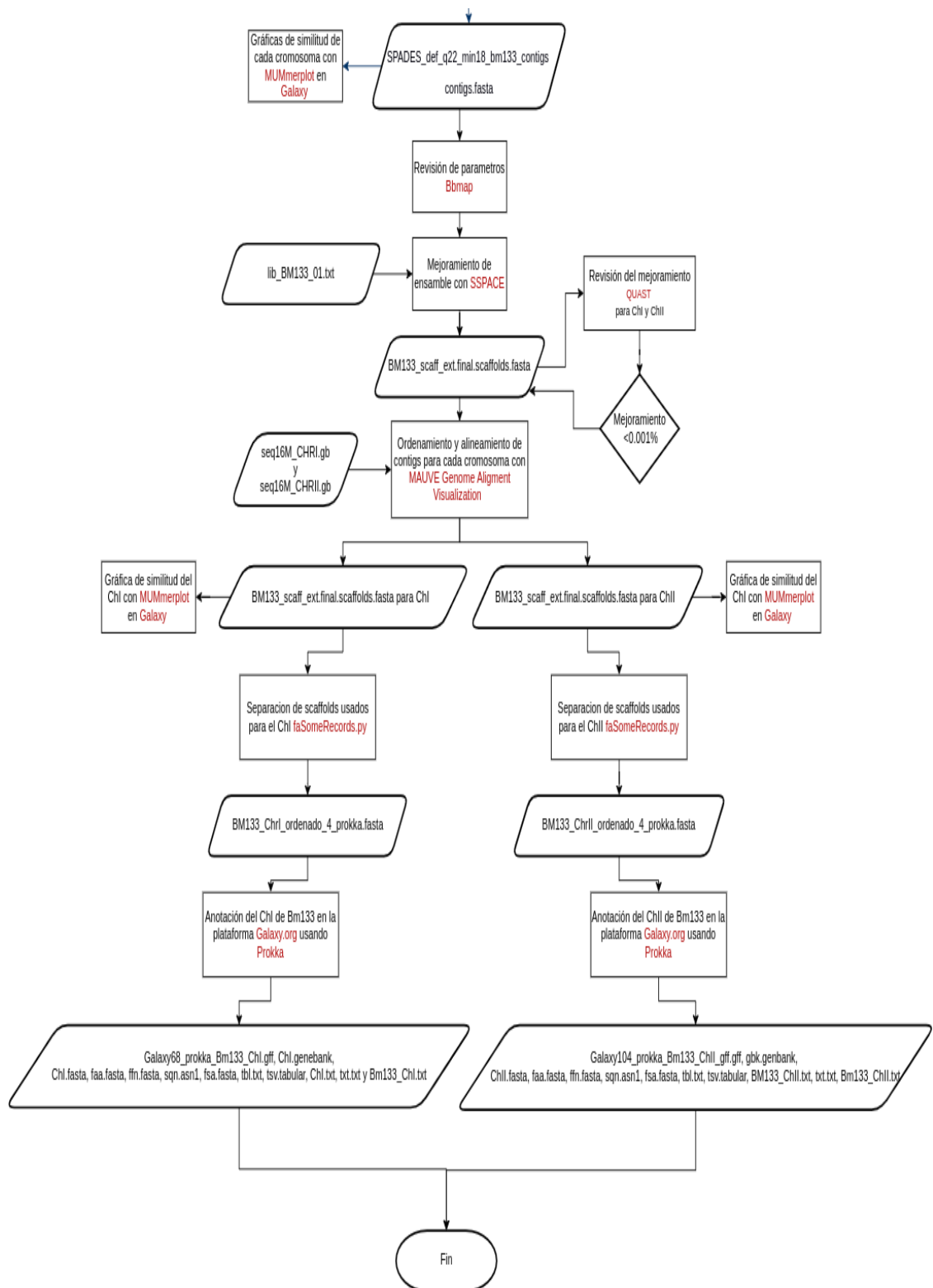
WHO (2014) 'WHO estimates of the global burden of diseases', *Who*, 46(3), pp. 1–15. Available at: <https://doi.org/10.1016/j.fm.2014.07.009>.

Wick, R.R. *et al.* (2017) 'Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads', *PLoS Computational Biology*, 13(6), pp. 1–22. Available at: <https://doi.org/10.1371/journal.pcbi.1005595>.

11 Apendices

11.1 Diagrama de flujo del ensamble





11.2 Lista de Programas usados

Nombre del programa	Versión	Descripción	Archivos de entrada	Archivos de salida
Kraken2	S/D	Asignar etiquetas taxonómicas a lecturas de secuenciación	.fastq o .fastq.gz	Kraken_report.html
Krona	S/D	Genera gráficas interactivas HTML de datos jerárquicos	Kraken_report.html	Krona.html
FastQC	0.11.9	Detecta potenciales problemas en un conjunto de datos obtenidos de secuenciaciones masivas	Fastq, SAM, BAM	Report.html
MultiQC	V1.12.dev0	Combina los resultados de análisis bioinformáticos de varias muestras en un reporte único	Fastq y quast_report.html	multiqc_report.html
Cutadapt	4.2	Remueve los adaptadores utilizados en el proceso de secuenciación y nos permite el mejoramiento de los parámetros promedio de las lecturas	input_R1.fastq e input_R2.fastq	output_R1.fastq e output_R2.fastq
SPAdes	3.13.1	Herramienta de ensamblaje que contiene varios canales para el ensamblaje de secuencias	input_R1.fastq e input_R2.fastq	output_R1.fastq e output_R2.fastq
QUAST	5.0.2	Evalúa ensamblajes genómicos	contigs.fasta o scaffolds.fasta	Report.html
NUCmer	3.1	Herramienta de MUMmer dedicada al alineamiento de contigs contra una referencia genómica	contigs.fasta, seq_16M_ChI.fasta y seq_16M_ChII.fasta	archivo.delta, archivo.delta.filter y archivo.delta.filter.coords
MUMmerplot	5.2.7	Comandos de generación de gráficos utilizando de base los archivos obtenidos en NUCmer	contigs.fasta, seq_16M_ChI.fasta y seq_16M_ChII.fasta	archivo.tabular y archivo.png
Galaxy.org	4.0.0beta2+galaxy0	Plataforma de recursos web abierta para uso intensivo de datos biomédicos	contigs.fasta, seq_16M_ChI.fasta y seq_16M_ChII.fasta	NA
Bbmap	VN 38.86	Alineador global para el empalme de lecturas obtenidas en la secuenciación de DNA y de RNA	archivos.fastq y archivos.fasta	archivo.txt y archivos.sam
SSPACE	Standard_v3.0_linux	Comando que permite extender y realizar scaffolds a partir de contigs pre-ensamblados usando bibliotecas de lecturas	archivo.txt, archivos.fastq y archivos.fasta	archivo_extendido.fasta
MAUVE	Snapshot_2015-02-13	Software que alinea regiones ortologas y xenologas entre dos o más secuencias genómicas que hayan o no sufrido cambios locales o a gran escala	archivo.fasta y archivo.gb	archivo.fasta, archivo.tab y aarchivo.fasta

faSomeRecord s.py	S/D	Extrae archivos FASTA a partir de un archivo multiFASTA basada en una lista de titulos de FASTA	archivo.txt y archivo.fasta	archivo.fasta
				archivo.gff, .genebank, .fasta, faa.fasta, ffn.fasta, sqn.asn1, fsa.fasta, tbl.txt, tsv.tabular, .txt, txt.txt y
Prokka	1.14.5	Prokka es una herramienta de software para una anotacion rápida de genomas procarióticos.	archivo.fasta	archivo.txt