

# **República de Cuba**



**Tesis de Doctorado**

## **Nuevos antígenos de superficie Mycroplasma gallisepticum identificados mediante un enfoque de vacunología reversa**

**Jose Antonio Agüero Fernández**

***Nuevos antígenos de superficie Mycroplasma gallisepticum identificados mediante un enfoque de vacunología reversa*** (Tesis de Doctorado) / Jose Antonio Agüero Fernández. – La Habana : Editorial Universitaria, 2016.

© **Autor:** Jose Antonio Agüero Fernández.

**Digitalización:** Editorial Universitaria del Ministerio de Educación Superior de la República de Cuba, 2016.

Calle 23 entre F y G, # 564.

El Vedado, Ciudad de La Habana, CP 10400,

Cuba.

Página web: <http://eduniv.mes.edu.cu>



Universidad Agraria de La Habana  
Centro Nacional de Sanidad Agropecuaria

**NUEVOS ANTÍGENOS DE SUPERFICIE DE**  
*Mycoplasma gallisepticum*  
**IDENTIFICADOS MEDIANTE UN ENFOQUE DE**  
**VACUNOLOGÍA REVERSA**

TESIS EN OPCIÓN AL GRADO DE DOCTOR EN CIENCIAS VETERINARIAS

JOSÉ ANTONIO AGÜERO FERNÁNDEZ

Mayabeque  
2016

Universidad Agraria de La Habana  
Centro Nacional de Sanidad Agropecuaria

**NUEVOS ANTÍGENOS DE SUPERFICIE DE**  
***Mycoplasma gallisepticum***  
**IDENTIFICADOS MEDIANTE UN ENFOQUE DE**  
**VACUNOLOGÍA REVERSA**

TESIS EN OPCIÓN AL GRADO DE DOCTOR EN CIENCIAS VETERINARIAS

**JOSÉ ANTONIO AGÜERO FERNÁNDEZ**

**Tutora: DMV. Siomara Martínez, DrC.**

Mayabeque  
2016

## SÍNTESIS

Para identificar antígenos de superficie, de *Mycoplasma gallisepticum*, el agente etiológico fundamental de la Enfermedad Respiratoria Crónica de los pollos, la cual causa millonarias pérdidas en la industria avícola, se utilizó la Vacunología Reversa. Esto se realizó con el empleo de la herramienta bioinformática desarrollada *en casa* Bact-Vap. Para evaluar las predicciones se obtuvieron las moléculas predichas, mediante la tecnología del ADN-recombinante. Se demostró que el sistema posee un alto valor predictivo, tanto para candidatos vacunales protectores, como para proteínas de superficie de micoplasmas. Se generó una lista de 291 potenciales antígenos de membrana. Se clonaron diecinueve de estos genes, de los cuales se expresaron trece, que incluyeron siete no caracterizados previamente. Se demostró la inmunogenicidad y la localización en la membrana de las proteínas codificadas por los mismos, así como la antigenicidad de los productos recombinantes. En un aislamiento cubano de *Mycoplasma gallisepticum*, se detectó la presencia de 89 de los primeros genes predichos. Los resultados, permiten contar con una batería de potenciales blancos vacunales de esta bacteria y proteínas que podrían tener valor diagnóstico. Además, se proporciona una plataforma de predicción de potenciales antígenos vacunales válida para el resto de los patógenos (humanos y veterinarios) de origen bacteriano.

## TABLA DE CONTENIDO

## INTRODUCCIÓN

1

1	Capítulo 1. REVISIÓN BIBLIOGRAFICA	7
1.1	Micoplasmas	7
1.1.1	Introducción	7
1.1.2	Cultivo	7
1.1.3	Patogénesis	8
1.1.4	Adhesión	8
1.1.5	Variabilidad Antigénica	9
1.1.6	<i>Mycoplasma gallisepticum</i>	9
1.1.7	Control de la micoplasmosis aviar	10
1.2	Vacunología Reversa	12
1.2.1	El diseño de la próxima generación de vacunas	13
1.2.2	Predicción de la localización sub-celular de las proteínas.	14
1.2.2.1	Métodos basados en búsqueda de similitud de secuencias.	14
1.2.2.2	Métodos basados en búsqueda de motivos/dominios funcionales.	15
1.2.2.3	Métodos basados en secuencias.	16
1.2.2.3.1	Identificación de proteínas integrales de membrana	17
1.2.2.4	Principales métodos para la predicción de LS, basados en secuencia.	17
1.2.3	Inmunoinformática y diseño de vacunas.	18
1.2.3.1	Mapeo epitópico de células T.	19
1.2.3.2	Mapeo epitópico de células B.	20
1.2.4	Determinación de la función de las proteínas.	21
1.2.5	Métodos integradores en la predicción de blancos vacunales.	22
1.3	Sistemas heterólogos de expresión	23
1.3.1	Expresión de proteínas recombinantes en <i>Escherichia coli</i>	23
1.4	Micoplasmas y Vacunología Reversa	25
2	Capítulo 2. PREDICCIÓN <i>IN SILICO</i> CON EL EMPLEO DE Bact-VAP, DE ANTÍGENOS DE SUPERFICIE DE <i>M. gallisepticum</i>	27
2.1	Introducción	27
2.2	Materiales y Métodos	28
2.2.1	Validación del predictor de BBTMs ( $\beta$ -Barrel Predictor)	28
2.2.1.1	Construcción del conjunto de validación	28
2.2.1.2	Evaluación del desempeño de BBTMs	29
2.2.1.3	Comparación con otros predictores BBTM	29
2.2.2	Adiciones al BepiPred	30
2.2.3	Determinación del valor de corte óptimo del programa SPAAN para micoplasmas	31
2.2.4	Algoritmos de predicción y de priorización empleados por Bact-VAP para la predicción de potenciales blancos vacunales	33

2.2.5	Validación de Bact-VAP.	35
2.2.6	Identificación computacional con el empleo de Bact-VAP, de proteínas de superficie antigénicas de <i>M. gallisepticum</i>	37
2.3	Resultados y Discusión	37
2.3.1	Validación del predictor de BBTMs	37
2.3.1.1	Evaluación del desempeño	37
2.3.1.2	Comparación con otros predictores BBTM	38
2.3.2	Determinación del valor de corte óptimo del programa SPAAN para micoplasmas	40
2.3.3	Validación de Bact-VAP	42
2.3.4	Identificación computacional con el empleo de Bact-VAP, de las proteínas de superficie antigénicas de <i>M. gallisepticum</i>	47
3	CAPÍTULO 3. OBTENCIÓN Y CARACTERIZACIÓN TOPOLÓGICA Y ANTIGÉNICA DE LAS PROTEÍNAS OBTENIDAS	50
3.1	Introducción	50
3.2	Materiales y Métodos	50
3.2.1	Vectores, cepas bacterianas y condiciones de cultivo de <i>M. gallisepticum</i> .	50
3.2.2	Diseño de los cebadores para el clonaje posterior de los genes	51
3.2.3	Extracción de ADN genómico de la cepa R(low) de <i>M. gallisepticum</i>	52
3.2.4	Amplificación por PCR de los genes seleccionados	53
3.2.5	Generación de las construcciones recombinantes	54
3.2.5.1	Preparación de los productos de PCR	54
3.2.5.2	Preparación del vector	54
3.2.5.3	Ligazón de los genes con el vector	55
3.2.5.4	Transformación química de DH5α	56
3.2.5.5	Monitoreo de las construcciones recombinantes	57
3.2.6	Obtención de los polipéptidos recombinantes a partir de las construcciones generadas	58
3.2.6.1	Transformación de las cepas de expresión	58
3.2.6.2	Monitoreo de la expresión y extracción por SDS-PAGE y Western-Blot.	59
3.2.6.3	Expresión de los genes recombinantes	59
3.2.7	Purificación de los polipéptidos recombinantes de <i>M. gallisepticum</i>	60
3.2.7.1	Extracción y solubilización de los productos recombinantes	60
3.2.7.2	Purificación por IMAC	61
3.2.8	Obtención de proteínas de membrana externa de la cepa R(low) de <i>M. gallisepticum</i>	62
3.2.8.1	Generación de anticuerpos específicos contra las proteínas de membrana externa.	62
3.2.8.1.1	Cuantificación de las proteínas de membrana externa de <i>M. gallisepticum</i> y los polipéptidos recombinantes purificados.	63
3.2.9	Evaluación de la antigenicidad de los polipéptidos recombinantes purificados en condiciones desnaturalizantes	64
3.2.10	Evaluación de la antigenicidad de los polipéptidos recombinantes purificados en condiciones nativas	65
3.2.11	Evaluación preliminar de la localización en la superficie bacteriana de las proteínas en estudio	65

3.2.12	Secuenciación de los segmentos clonados	66
3.2.13	Identificación de potenciales genes de proteínas de superficie, en aislamiento cubano de <i>M. gallisepticum</i>	66
3.3	Resultados y Discusión	67
3.3.1	Extracción de ADN genómico	67
3.3.2	Amplificación por PCR de los genes identificados	67
3.3.3	Generación de las construcciones recombinantes	70
3.3.3.1	Transformación y monitoreo de las construcciones recombinantes	70
3.3.3.2	Obtención de los polipéptidos recombinantes a partir de las construcciones generadas.	73
3.3.4	Purificación de los polipéptidos recombinantes de <i>M. gallisepticum</i>	76
3.3.4.1	Extracción y solubilización de los productos recombinantes	77
3.3.4.2	Purificación por IMAC	80
3.3.5	Generación de anticuerpos específicos contra las proteínas de membrana externa	81
3.3.6	Evaluación de la antigenicidad de los polipéptidos recombinantes	81
3.3.7	Evaluación preliminar de la localización en la superficie bacteriana de las proteínas de <i>M. gallisepticum</i> en estudio	85
3.3.8	Secuenciación de los segmentos clonados	86
3.3.9	Identificación de potenciales genes de proteínas de superficie en aislamiento cubano de <i>M. gallisepticum</i>	87
4	CAPÍTULO 4. DISCUSIÓN GENERAL	90
	CONCLUSIONES	96
	RECOMENDACIONES	97
	PRODUCCION CIENTÍFICA DEL AUTOR RELACIONADA CON LA TESIS	
	REFERENCIAS BIBLIOGRÁFICAS	
	ANEXOS	



**ABREVIATURAS (Por orden alfabético)**

AA	Aprendizaje automático
AIF	Adyuvante incompleto de Freund
ADN	Ácido desoxirribonucleico
ANN	Redes neuronales artificiales
ARN	Ácido ribonucleico
AUC	Área bajo la curva
BBTM	Barril- $\beta$ transmembrana
BCG	Bacillus Calmette Guerin
BSA	Sero albúmina bovina
CENSA	Centro Nacional de Sanidad Agropecuaria
CENPALAB	Centro Nacional de Producción de Animales de Laboratorio
CI	Cuerpos de inclusión
CTL	Célula T citotóxica
CVN	Conjunto de validación negativo
CVP	Conjunto de validación positivo
DO	Densidad óptica
FN	Falsos negativos
FP	Falsos positivos
HLA	Antígeno leucocitario humano
HMM	Modelos Ocultos de Markov
IMAC	Cromatografía de afinidad por iones metálicos
IPTG	Isopropil- $\beta$ -D-tiogalactosidasa
LB	Luria Bertani
LS	Localización sub-celular
MCC	Coeficiente de correlación de Matthews
MHC	Complejo Mayor de Histocompatibilidad
Mi	Masa del inserto
MSV	Máquinas de Soporte Vectorial
Mv	Masa del vector
NBT/BCIP	nitroazul de tetrazolio/5-Bromo-4-cloro-3-indolil fosfato
OPM	Orientación de Proteínas en la Membrana
ORF	Marco de lectura abierto
PAMP	Patrones Moleculares Asociados a Patógenos
PBS	Tampón fosfato salino
PCR	Reacción en Cadena de la Polimerasa
PDB	Banco de Datos de Proteínas
PDBTM	PDB de proteínas transmembrana
PVC	Candidato vacunal protector
RM	Relación molar entre inserto y vector
RN	Redes Neuronales
ROC	Receiver Operating Characteristic
SB	Super Broth
SDS	Dodecil Sulfato Sódico
SDS-PAGE	Electroforesis en gel de poliacrilamida con SDS
VIH	Virus de la Inmunodeficiencia Humana
VN	Verdaderos negativos
VP	Verdaderos positivos
VPP	Valor predictivo positivo
VR	Vacunología reversa

## INTRODUCCIÓN

Los miembros del género *Mycoplasma*, ubicado dentro de la clase Mollicutes, se han descrito a menudo como las formas de vida autorreplicativa más simples y pequeñas conocidas debido a que poseen un tamaño celular diminuto, carecen de pared celular y presentan un genoma reducido. Una consecuencia de su reducido genoma es la ausencia de un conjunto de ciertas enzimas, lo que les confiere una limitada capacidad metabólica y son dependientes de una amplia gama de nutrientes que toman del medio (Abu-Amro et al., 2000). Los micoplasmas son importantes patógenos de una amplia gama de animales, plantas e insectos en los cuales causan a menudo infecciones crónicas que resultan en enfermedades más mórbidas que mortales. Estas bacterias han adoptado un estilo de vida parasitario y viven en estrecho contacto con su hospedero, con una predilección por la superficie de las mucosas de los tractos respiratorio y genital (Citti et al., 2010). La incidencia de estos microorganismos en la Medicina Veterinaria es alta y generalmente están vinculados a importantes pérdidas económicas a causa de las enfermedades que provocan en vacunos, cabras, ovejas, cerdos y aves, fundamentalmente.

*Mycoplasma gallisepticum* (*M. gallisepticum*) es el agente etiológico fundamental de la Micoplasmosis Aviar, enfermedad de curso crónico que afecta fundamentalmente a pavos y gallinas, aunque puede estar presente en varias especies de aves (patos, gansos, faisanes, codornices, perdices, gallinas de Guinea y palomas) (Evans et al., 2005). Las grandes pérdidas económicas que ocasiona lo ubican como el micoplasma más significativo para la avicultura. Solo en los Estados Unidos las pérdidas han alcanzado los 150 millones de dólares anuales, mientras que a nivel global estas han estado alrededor de los 780 millones de dólares en igual período de tiempo (Hennigan et al., 2012). En Cuba, en el pasado año, la Enfermedad Respiratoria Crónica y la micoplasmosis aviar fueron la 3era causa de pérdidas en la avicultura, con un incremento con respecto al año 2014 (IMV, 2016). Aunque la misma no necesariamente provoca elevada mortalidad, la morbilidad es alta, y provoca reducción de la conversión alimentaria, disminución de la producción de huevos y de la

ganancia de peso, retraso del crecimiento y predisposición a las infecciones secundarias virales y bacterianas (Ley et al., 1997a, Kleven, 1998).

Los antibióticos han demostrado su ineficacia en el control de las infecciones por *M. gallisepticum* (Li et al., 2010, Gerchman et al., 2011). La vacunación con bacterinas o vacunas vivas puede ser una opción en aquellos casos en los cuales el control de la infección no es factible. Aunque las vacunas proporcionan alguna protección contra la enfermedad y disminuyen las pérdidas económicas de los productores, aún ocurren brotes de las enfermedades, a frecuencias demasiado altas (Gaunson et al., 2006).

La elucidación de los mecanismos patogénicos de *M. gallisepticum* y la exacta contribución de cada uno de los factores de virulencia, son críticos en la búsqueda racional de mejores vacunas y estrategias de control. La adhesión de los micoplasmas al epitelio mucosal del tracto respiratorio de las aves es un requisito previo e indispensable para la infección y posterior desarrollo de la enfermedad, y se logra mediante lipoproteínas localizadas en su superficie (Browning et al., 2011, Shimizu, 2015). Las cepas de *M. gallisepticum* no adherentes son incapaces de provocar la enfermedad en animales infestados experimentalmente (Papazisi et al., 2002). Se han descrito algunas citadhesinas como la familia VIhA (Markham et al., 1994), las proteínas GapA (Goh et al., 1998) y PvpA (Boguslavsky et al., 2000) y moléculas relacionadas con la citadherencia como CrmA (Papazisi et al., 2000). Otros ejemplos son OppA cuyo homólogo en *Mycoplasma hominis* ha mostrado una importante contribución a la patogénesis de esta bacteria. El cambio fenotípico de la membrana es otro de los más importantes mecanismos patogénicos de esta bacteria. Las numerosas lipoproteínas que están codificadas por la familia VIhA (51), median la variación de la arquitectura de la superficie de *M. gallisepticum* y se cree que están involucradas en la evasión del sistema inmune del hospedero y en su adaptación a los cambios del entorno (14). Este polipéptido protegió parcialmente a pollos confrontados con una cepa virulenta de *M. gallisepticum* (Saito et al., 1993). Estudios recientes indican que

muchos factores de virulencia aún no han sido identificados, definidos y caracterizados (Szczepanek et al., 2010b, Tulman et al., 2012).

Los métodos convencionales para identificar estas proteínas consumen mucho tiempo y recursos. El surgimiento de la Vacunología Reversa (VR), permitió replantearse la manera de diseñar vacunas, al emplear la información suministrada por el genoma de los patógenos contra ellos mismos (Seib et al., 2012, Bruno et al., 2015, Heinson et al., 2015). Este nuevo concepto se basa en la idea de que el genoma microbiano contiene el repertorio completo de los potenciales antígenos del patógeno, y consiste en predecir, con el empleo de herramientas computacionales, las moléculas con mayores posibilidades de convertirse en candidatos vacunales (Pizza et al., 2000). Desde que la VR comenzó a aplicarse ha evolucionado y ayudado a los investigadores a superar los límites de la vacunología convencional y llevado al descubrimiento y desarrollo de nuevas vacunas (Seib et al., 2012, Bruno et al., 2015). La VR es insuficiente para garantizar la eficacia *in vivo* de los candidatos predichos *in silico* a partir de los datos genómicos. Los candidatos vacunales seleccionados sobre la base de este criterio, necesitan ser validados experimentalmente (Mora et al., 2003). La idea original detrás de la VR era monitorear *in silico* el genoma de un patógeno para identificar los genes que codifican para proteínas con los atributos de un buen blanco vacunal, por ejemplo, proteínas que estén expuestas en la superficie y bien conservadas entre las cepas. Las proteínas seleccionadas se expresan entonces en *Escherichia coli* (*E. coli*) y se emplean para inocular ratones y evaluar la inmunogenicidad y la protección que estos antígenos sean capaces de inducir (Rappuoli, 2000, Seib et al., 2012).

Existen numerosas herramientas disponibles, algunas web otras en versión de escritorio, para la predicción de aquellas características de las proteínas que las convierten en atractivos blancos vacunales, como la localización en la superficie (Krogh et al., 2001, Juncker et al., 2003, Garrow et al., 2005b, Yu et al., 2010, Petersen et al., 2011, Goldberg et al., 2014, Yu et al., 2014), la antigenicidad (Larsen et al., 2006, Nielsen et al., 2007, El-Manzalawy et al., 2008, Ponomarenko et al., 2008, Sweredoski and Baldi, 2008) y aquellas

que permiten hacer estudios de VR substractiva o comparativa (Pearson, 2000, Menlove et al., 2009). Pero no todos los métodos son igual de exactos y además los usuarios finales no siempre están bien informados acerca de los nuevos descubrimientos en el área ni cuentan con experiencia en el uso de herramientas computacionales. El orden en que estas herramientas se emplean también puede hacer la diferencia. Adicionalmente cada uno de ellos tiende a usar diferentes formatos de entrada y no siempre formatos de salida de fácil comprensión (Giombini et al., 2010). Esto hace difícil la integración de los datos. Un empleo exitoso de las mismas frecuentemente requiere de instalación local, ejecución en la línea de comandos y una capacidad de cómputo elevada. Se han obtenido contados resultados encaminados a la combinación de estas herramientas en sistemas integradores dedicados a la VR. Así se han publicado los programas NERVE (New Enhanced Reverse Vaccinology) (Vivona et al., 2006), Vaxign (He et al., 2010b), Jenner-Predict (Jaiswal et al., 2013). Ninguno de los mismos incluye la detección de lipoproteínas y Jenner-Predict carece, además, de la capacidad de predecir citadhesinas. Adicionalmente, en la actualidad ninguno de estos sistemas está disponible o funcionando correctamente.

La VR se ha aplicado de manera limitada a los micoplasmas, en ningún caso, para la exploración del genoma de *M. gallisepticum*. Con la excepción de *Mycoplasma mycoides* (Hamsten et al., 2008, Hamsten et al., 2009, Perez-Casal et al., 2015), algunas especies de micoplasmas porcinos como *Mycoplasma hyopneumoniae* (Siqueira et al., 2013, Reolon et al., 2014), *Mycoplasma flocculare*, *Mycoplasma hyorhinis* (Siqueira et al., 2013) y *Mycoplasma hyosynoviae* (Bumgardner et al., 2014), no se encuentran en la literatura muchos otros ejemplos donde se emplee una metodología computacional exhaustiva con estos propósitos. En ninguno de estos casos se empleó una metodología estandarizada y validada para estos microorganismos. Esto ha limitado, sin dudas, la identificación de antígenos con potencial para ser evaluados como candidatos vacunales contra las múltiples enfermedades ocasionadas por estos microorganismos.

Si se pudiera contar con una herramienta que permitiera predecir con éxito potenciales proteínas de membrana de *M. gallisepticum* y luego desarrollar una serie de experimentos diseñados para permitir la caracterización del papel de las mismas en los mecanismos de patogenicidad del microorganismo o la respuesta inmune que inducen, se estaría en condiciones de ofrecer potenciales blancos para el desarrollo de vacunas o sistemas diagnóstico. En el laboratorio de Genómica Microbiana y Bioinformática del Centro Nacional de Sanidad Agropecuaria (CENSA) se implementó una herramienta computacional denominada Bact-VAP (*Bacterial Vaccine Antigen Predictor*) (<http://baco.censa.edu.cu:8088/predictor>) basada en un algoritmo de predicción (*PipeLine*) que minimiza las carencias antes mencionadas y que se adecua a las características de los micoplasmas (incluye la predicción de importantes factores de virulencia de los micoplasmas como citadhesinas y lipoproteínas en general). La secuenciación del genoma de *M. gallisepticum* (Papazisi et al., 2003) ha creado una base sobre la cual desarrollar numerosos proyectos encaminados a la identificación *in silico* de las principales moléculas expuestas en la superficie de la bacteria o involucradas en los mecanismos de patogenicidad de este organismo, por ende, en la predicción de potenciales candidatos vacunales contra la micoplasmosis aviar.

En resumen, nos enfrentamos al siguiente problema científico: La identificación por métodos convencionales de antígenos de superficie de *Mycoplasma gallisepticum*, como potenciales blancos en el desarrollo de vacunas de nueva generación efectivas para su control, resulta compleja y no se dispone de una herramienta bioinformática única y/o adecuada, como alternativa para este fin.

Teniendo en cuenta lo expresado previamente, el presente trabajo se diseñó y ejecutó basándose en la **hipótesis** de que “la herramienta bioinformática Bact-VAP es efectiva para la identificación de antígenos de superficie de *Mycoplasma gallisepticum*”. Para evaluar esta hipótesis se establecieron los siguientes objetivos:

## OBJETIVO GENERAL

Identificar potenciales antígenos de superficie de *Mycoplasma gallisepticum* mediante el empleo de herramientas bioinformáticas.

## OBJETIVOS ESPECÍFICOS

1. Validar la herramienta Bact-VAP para la identificación de potenciales candidatos vacunales bacterianos y antígenos de superficie de micoplasmas.
2. Predecir antígenos de superficie de *M. gallisepticum* con el empleo de Bact-VAP.
3. Caracterizar las proteínas identificadas en cuanto a su localización sub-celular y su antigenicidad.

## NOVEDAD CIENTÍFICA

1. Por primera vez en Cuba se implementa un sistema de predicción de potenciales antígenos vacunales bacterianos basado en el concepto de la VR.
2. Por primera vez se aplica de manera exhaustiva la VR, a la identificación de antígenos de *M. gallisepticum* con potencial valor vacunal.
3. Se notifican nuevos antígenos de superficie de *M. gallisepticum*, con potencial valor vacunal.

## IMPORTANCIA PRÁCTICA DEL TRABAJO

1. Se cuenta con una herramienta bioinformática para la identificación de potenciales antígenos vacunales bacterianos validada y probada experimentalmente para el caso particular de *M. gallisepticum*.
2. Se identificaron nuevos antígenos de superficie de *M. gallisepticum*, que pudieran ser empleados en futuras investigaciones encaminadas al desarrollo de una vacuna contra la micoplasmosis aviar y/o sistemas de diagnóstico de esta bacteria.

## 1 Capítulo 1. REVISIÓN BIBLIOGRÁFICA

### 1.1 Micoplasmas

#### 1.1.1 Introducción

Los micoplasmas son la forma de vida autorreplicativa más pequeña y simple que se conoce, debido a que poseen un tamaño celular diminuto y un genoma reducido. Fueron calificados durante años como virus, pero a partir de 1930, con la determinación en ese año de la verdadera naturaleza de los virus, se hizo evidente que los micoplasmas no podían ser clasificados como tales porque presentaban ambos tipos de ácidos nucleicos, crecían en medios artificiales y se dividían por fusión binaria al igual que las bacterias. Taxonómicamente, la falta de pared celular de estos microorganismos, se ha usado para separar los micoplasmas de otras bacterias en una clase nombrada Mollicutes (*mollis*, suave; *cutis*, piel, en latín). Los términos “micoplasma” o “mollicutes” se han usado alternativamente para denotar cualquier especie incluida en la clase Mollicutes, los nombres comunes micoplasmas, ureaplasmas, entomoplasmas, esoplasmas, espiroplasmas, acholeplasmas, asteroleplasmas y anaeroplasmas deben ser empleados para los miembros de los géneros correspondientes (Razin, 1985). El rango de tamaño del genoma de los micoplasmas va desde 580 kb (*Mycoplasma genitalium*) (Fraser et al., 1995) hasta 1358 kb (*Mycoplasma penetrans*) (Sasaki et al., 2002).

#### 1.1.2 Cultivo

El mayor impedimento en la investigación y el diagnóstico de laboratorio de los micoplasmas ha sido la dificultad para cultivarlos. Los recientes proyectos sobre el genoma de los micoplasmas han proporcionado las explicaciones genéticas definitivas sobre estas dificultades (Himmelreich et al., 1996). Por estos motivos, para su crecimiento se emplean medios de cultivo complejos. Normalmente estos medios son a base de caldo cerebro corazón, peptona, extracto de levadura, y suero con varios suplementos (Razin and Hayflick, 2010).



### 1.1.3 Patogénesis

La base molecular de la patogenicidad de los micoplasmas permanece es en gran parte desconocida. El cuadro clínico de las infecciones causadas por estas bacterias en humanos y animales sugiere que el daño se debe más a las respuestas inmunes e inflamatorias del hospedero que a los efectos tóxicos directos causados por el patógeno sobre los componentes celulares. Sin embargo, todavía la vieja tesis de que el peróxido de hidrógeno y radicales superóxido generados por los micoplasmas causan lesiones en los tejidos del hospedero por estrés oxidativo continúa ganando apoyo experimental (Razin et al., 1998, Pilo et al., 2005, Hasselbring et al., 2012). Los mecanismos de patogenicidad del género *Mycoplasma* se atribuyen en parte a la competencia con las células del hospedero por sustratos metabólicos precursores de lípidos, purinas y pirimidinas; a su habilidad para adherirse e invadir a las células del hospedero; a la variación de su fenotipo (variación de fase/variación antigénica/variación de talla), probablemente para evadir la respuesta inmune del hospedero (evasión inmune); y a sus efectos citopáticos (Razin et al., 1998).

### 1.1.4 Adhesión

La adhesión de los micoplasmas a las células de su hospedero es un requisito previo para la colonización e infección. La pérdida de la capacidad de adhesión por mutación produce la pérdida de la infectividad, sin embargo la reversión al fenotipo citadherente es acompañado de la recuperación de la infectividad y virulencia (Krause, 1996, Papazisi et al., 2002). Para la colonización de sus hospederos, los micoplasmas poseen un conjunto de citadhesinas que pueden experimentar variaciones de fase y tamaño, lo cual le permite evadir eficientemente su sistema inmunitario (Razin, 1999). La ausencia de pared celular y algunos de sus apéndices, como las fimbrias, asociado con la adherencia de otros procariontes, apunta hacia el hecho de que los componentes celulares responsables de la adhesión, deben ser parte de su membrana citoplasmática (Browning et al., 2011). Además, la falta de pared celular facilita el contacto directo de las citadhesinas con los receptores específicos de las células del hospedero, lo que crea una condición que, en principio, podría llevar a la fusión de las dos membranas, y transferir o intercambiar componentes de ambas (Razin et

al., 1998). El contacto íntimo del micoplasma con la superficie celular del hospedero crea un microambiente que facilita la acumulación de productos finales del metabolismo del micoplasma, como el peróxido y radicales de superóxido, lo que causa daño oxidativo en las membranas del hospedero (Razin, 1985, Davis et al., 1985, Almagor et al., 1985, Razin, 2010). La adhesión de los micoplasmas se ha convertido en el principal campo de investigación desde los años 1970 (Razin and Jacobs, 1992, Balish et al., 2003).

### 1.1.5 Variabilidad Antigénica

El término "Variación Antigénica" o el también empleado "Cambio Fenotípico" se refiere a la habilidad de las especies microbianas de alterar el carácter antigénico de sus componentes de superficie, como flagelos, pilis, proteínas de membrana externa y cápsulas, que incrementan la colonización de los tejidos del hospedero y permiten evadir los fagocitos. Estas moléculas de superficie son los blancos principales del sistema inmune, así que la habilidad de estos microorganismos de cambiar rápidamente su repertorio antigénico de superficie y consecuentemente cambiar la inmunogenicidad de estas estructuras, les permite evadir con efectividad los sistemas de reconocimiento del hospedero (Zhang and Wise, 1997). Estos cambios fenotípicos también podrían suministrarle a las células de micoplasma la flexibilidad para adaptarse a diferentes nichos donde se requieren varios receptores para la adherencia, lo que constituiría un factor clave en las infecciones sistémicas al darle la capacidad a estos microorganismos de colonizar exitosamente las superficies mucosales y entonces propagarse dentro del hospedero (Citti and Rosengarten, 1997).

### 1.1.6 *Mycoplasma gallisepticum*

*M. gallisepticum* es el agente etiológico fundamental del síndrome respiratorio crónico en pollos, el agente primario de la Sinusitis Infecciosa en pavos (Levisohn and Kleven, 2000) y a principios de 1990 emergió como el causante común de conjuntivitis en el pinzón doméstico (Ley et al., 1996, Cecchini et al., 2007). Cada año causa severas pérdidas económicas a la industria avícola a nivel mundial, debido a que produce una reducción en la

ganancia de peso y en la producción de huevo en las aves infectadas (Sundquist et al., 1996, Levisohn and Kleven, 2000). El síndrome respiratorio crónico se caracteriza por la exfoliación de las células epiteliales ciliadas, además de causar la acumulación de un exudado inflamatorio en la tráquea y una severa respuesta inflamatoria en los sacos aéreos (Ley et al., 1997b). La ciliostasis parece ser la primera secuela de los efectos citopáticos en las células del epitelio respiratorio. Algunas especies de micoplasmas exhiben el potencial para invadir las áreas más profundas de los tejidos de la tráquea, es decir la lámina propia. Much y colaboradores demostraron La capacidad de *M. gallisepticum* para cruzar la barrera mucosal (Much et al., 2002), esto probablemente juega un papel fundamental en la propagación sistémica de este patógeno.

#### **1.1.7 Control de la micoplasmosis aviar**

Los antibióticos empleados han demostrado su ineficacia en el control de las infecciones por *M. gallisepticum* (Li et al., 2010, Gerchman et al., 2011). Actualmente las prácticas de control se caracterizan por el uso intensivo de la bioseguridad y el monitoreo serológico de las parvadas de aves (Ley et al., 1997b, Liu et al., 2001). Las medidas regulatorias han sido en su mayor parte exitosas al minimizar los brotes de *M. gallisepticum* en la industria del pollo de ceba y el pavo. En los Estados Unidos el control de *M. gallisepticum* por lo general se basa en el sacrificio de los animales infectados, seguido del reemplazo por una nueva línea de animales negativos a este patógeno y el mantenimiento de condiciones de bioseguridad para que los animales productores y su descendencia estén libres de micoplasmas. Sin embargo, la erradicación de parvadas infectadas a menudo no es factible (Pharr et al., 2006), es por ello que la vacunación con bacterinas o vacunas vivas atenuadas es una importante alternativa. Las bacterinas se consideran más seguras que las vacunas vivas; sin embargo, aunque reducen síntomas clínicos, no previenen la infección, además de que tienen un alto costo y la necesidad de vacunar individualmente cada ave (Whithear, 1996). Por su parte las vivas atenuadas a partir de las cepas ts-11 y 6/85, impiden la infección pero los niveles de protección no son siempre los esperados (Gingerich, 2002). Las basadas en la cepa F son más efectivas pero la virulencia de esta cepa en pollos jóvenes y pavos limita

su empleo (Collier et al., 2015). Además, recientemente se han publicado reportes de casos de campo relacionados con una aparente reversión a la virulencia y la transmisión vertical de la vacuna basada en la ts-11 (El Gazzar et al., 2011, Armour and Ferguson-Noel, 2015). De esta manera tenemos que, aunque las vacunas proporcionan alguna protección contra la enfermedad y disminuyen las pérdidas económicas de los productores, aún ocurren brotes de la enfermedad, a frecuencias demasiado altas (Gaunson et al., 2006).

Por tanto, se hace necesaria la búsqueda de nuevos candidatos vacunales que puedan elevar la protección contra una infección de *M. gallisepticum*. Actualmente el uso de proteínas recombinantes como inmunógenos es la estrategia más prometedora para satisfacer las demandas de la vacunología en el futuro (Prysljak et al., 2011). Las ventajas de vacunas de subunidades incluyen una seguridad incrementada, un costo reducido, menos competición antigénica, y la acción directa sobre el sitio de infección (Potter and Babiuk, 2001). Las vacunas de subunidades cuentan con estrategias compatibles con antígenos predichos por VR que incluyen: vectores no patógenos, sistemas de liberación de antígenos no vivos, y vacunas de ADN (Bowman et al., 2011).

La adherencia de los micoplasmas a las células epiteliales respiratorias del hospedero se realiza mediante lipoproteínas localizadas en la superficie de la bacteria. Aunque se han identificado algunas de las proteínas involucradas en la adherencia como GapA, CrmA (Goh et al., 1998, Papazisi et al., 2002, Mudahi-Orenstein et al., 2003) y PvpA (Yogev et al., 1994), los datos que emergen de los estudios de *M. gallisepticum* indican que algunas citadhesinas o moléculas accesorias relacionadas, están aún por ser identificadas, definidas y caracterizadas (Szczepanek et al., 2010b, Tulman et al., 2012).

Los datos recolectados en los últimos años han subrayado la alta frecuencia de variación espontánea en la expresión y la estructura de las proteínas de superficie de los micoplasmas (Citti et al., 2010). La mayoría de las investigaciones sobre la variación fenotípica de *M. gallisepticum* se han enfocado en VlhA y su respectiva familia de genes (*vlhA*), que está compuesta por poco más de 70 genes en dependencia de la cepa en estudio (Papazisi et al., 2003). La expresión de los miembros de la familia de genes *vlhA* de

*M. gallisepticum* oscila tanto *in vitro* como *in vivo* (Glew et al., 2000), y esta variación podría ser influenciada notablemente por la presencia de anticuerpos VIhA específicos (Glew et al., 1998, Markham et al., 1998). En *M. gallisepticum*, además de la familia VIhA, se han identificado otras proteínas de superficie que también pueden ser expresadas variablemente. En este grupo se incluyen la proteína PvpA perteneciente a la familia de las citadhesinas (Yogev et al., 1994), y GapA (Keeler et al., 1996, Ganapathy and Bradbury, 1998). En contraste con VIhA, estas proteínas son codificadas cada una por un único gen.

## 1.2 Vacunología Reversa

La vacunación es la manera más eficaz y sustentable de prevenir las enfermedades infecciosas. Desde que el trabajo de Jenner fue refinado por Louis Pasteur con el desarrollo de la vacuna contra la rabia en 1885, los científicos se han enfocado en la vacunación como la mejor defensa contra numerosos patógenos bacterianos y virales. Las vacunas contra difteria, tétano, polio, sarampión, paperas, rubéola, neumococo, hepatitis B, y meningitis (*Haemophilus influenzae* y meningococo del serogrupo C, han reducido la incidencia y mortalidad de estas enfermedades en un 97–99% (Moxon and Rappuoli, 2002). Los principios establecidos por Louis Pasteur, es decir, aislamiento, inactivación y administración de los microorganismos causantes de la enfermedad, rigieron el desarrollo de vacunas a lo largo del siglo XX (Bagnoli et al., 2011). La investigación y desarrollo en el campo de la vacunología experimenta un interesante renacimiento en la comunidad científica mundial. Existen cuatro razones fundamentales que justifican este creciente interés: (I) la falta de tratamiento eficaz para un gran número de infecciones devastadoras; (II) la emergencia de bacterias resistentes a múltiples drogas; (III) la necesidad de mejorar la seguridad de más vacunas patentadas obtenidas por métodos tradicionales; y finalmente, (IV) el gran compromiso de la comunidad científica por innovar en el diseño e investigación en vacunas, con la convergencia de las ciencias ómicas, dígame genómica, proteómica e inmunómica con la vacunología (Bagnoli et al., 2011).

### 1.2.1 El diseño de la próxima generación de vacunas

El desarrollo en la investigación de vacunas en su conjunto, puede categorizarse en tres generaciones: las vacunas de "primera generación" fueron desarrolladas esencialmente basadas en los principios básicos de Pasteur que consiste en usar patógenos inactivos por completo o formas vivas atenuadas como vacuna, por ejemplo, *Bacillus Calmette Guerin* (BCG), tosferina, polio, rabia y viruela (Moxon and Rappuoli, 2002, Kaushik and Sehgal, 2008); las vacunas de "segunda generación" están compuestas por componentes celulares microbianos purificados, y se conocen como vacunas de subunidades, dígame polisacáridos o proteínas usados como antígenos contra el tétano, la difteria, el ántrax, la neumonía, la influenza, las hepatitis B, y la Enfermedad de Lyme (Plotkin, 2005); esta metodología se ha apoyado más recientemente en los nuevos adelantos de la tecnología del ADN recombinante y la química de polisacáridos. Estos métodos empíricos convencionales empleados en el diseño de vacunas muestran severas limitaciones: (I) no pueden ser aplicados a microbios que no se pueden cultivar *in vitro*, por ejemplo hepatitis C, virus del papiloma serotipos 16 y 18 y *Mycobacterium leprae*; (II) a menudo estas vacunas no son capaces de proporcionar una protección a largo plazo contra patógenos antigénicamente hipervariables, tal es el caso de meningococo serogrupo B, gonococo, malaria, y virus de la inmunodeficiencia humana (VIH) (Rappuoli, 2004); (III) no se adaptan a patógenos oportunistas como *Staphylococcus aureus* (T et al., 2012); (IV) son lentos en la lucha contra patógenos que evolucionan rápidamente como VIH; y por último pero no menos importante, (V) el desarrollo de vacunas por métodos tradicionales consume mucho tiempo y recursos.

La VR surgió a finales de la última década del siglo pasado como un método innovador que emplea algoritmos bioinformáticos para identificar antígenos en el proteoma bacteriano que pueden representar potenciales candidatos vacunales (Mora et al., 2003, Vivona et al., 2008), Esto permite reducir el número de candidatos a ser evaluados experimentalmente. Además, este método tiene la ventaja de descubrir antígenos independientemente de su abundancia y sin la necesidad de crecer el microorganismo *in vitro*. Los antígenos seleccionados pueden expresarse en sistemas heterólogos, purificarse y entonces evaluar

dichos antígenos respecto a su potencial como candidatos vacunales (Rappuoli, 2001). Para predecir con exactitud la eficacia *in vivo* de los candidatos antigénicos los datos genómicos por si solos son insuficientes. Los candidatos vacunales seleccionados *in silico*, necesitan ser validados experimentalmente (Mora et al., 2003). Desde que la VR fue concebida y aplicada hace diez años, esta tecnología ha progresado de forma espectacular. La localización sub-celular (LS) sigue siendo considerada como uno de los criterios principales en la predicción de blancos vacunales, así como la antigenicidad y la función de las proteínas. Sin embargo, se han añadido nuevos criterios a estos análisis. A menudo el número de dominios transmembrana para un blanco vacunal se considera en el filtrado bioinformático porque las proteínas de membrana externa que contienen más de una hélice transmembrana son, en general, difíciles de clonar y purificar (Casadevall, 2003). La predicción de los candidatos vacunales que se conservan en todos los genomas de una misma especie es cada vez más común (Maione et al., 2005, Naz et al., 2015). Si los genomas de las cepas no patógenas de las especies están también disponibles, se buscan los blancos vacunales que están presentes en los genomas de las cepas patógenas pero están ausentes en las cepas no virulentas (Moriel et al., 2010). Para evitar la autoinmunidad, los blancos vacunales predichos no deben poseer similitud de secuencia con proteínas del hospedero (Jaiswal et al., 2013).

## **1.2.2 Predicción de la localización sub-celular de las proteínas**

### **1.2.2.1 Métodos basados en búsqueda de similitud de secuencias**

Como una posible consecuencia de la evolución molecular, las proteínas que comparten una alta identidad de secuencia, por lo general se dirigen a un mismo compartimento sub-celular. Esto se ha descrito en varios estudios computacionales (Nair and Rost, 2002, Pierleoni et al., 2006, Yu et al., 2006). Por lo tanto, con pocas excepciones, la mejor manera de anotar la LS de una secuencia es hacerlo a partir de una búsqueda de homología. Aunque las anotaciones basadas en similitud a menudo son las más precisas, en algunos casos estas pueden perder señales cortas y/o dispersas que son capaces de cambiar la ubicación de

una proteína sin afectar en gran medida la similitud global. Por otra parte, para muchas cadenas de proteínas, las anotaciones basadas en la similitud no pueden llevarse a cabo, particularmente cuando se tienen en cuenta especies poco estudiadas (Pierleoni et al., 2007).

#### **1.2.2.2 Métodos basados en búsqueda de motivos/dominios funcionales**

Un enfoque menos común para la anotación de LS se basa en la búsqueda de motivos y/o dominios de secuencia que se correlacionan con una determinada localización. Estos métodos también se podrían, en teoría, aplicar a las secuencias parciales; sin embargo, sufren de las mismas limitaciones de las búsquedas basadas en homología, ya que no son capaces de clasificar cada secuencia. Además, los dominios, las familias y los motivos son difíciles de detectar cuando están involucradas homologías de secuencias remotas (Casadio et al., 2008). Existen métodos que se basan en diferentes bases de datos de familias y dominios. PROSITE (*Protein Site Database*) (Hulo et al., 2006) ofrece herramientas para el análisis de secuencias de proteínas y detección de motivos de proteínas; es parte de los servidores de análisis de proteómica de ExPASy (*Expert Protein Analysis System*). pTARGET (Guda, 2006) busca dominios PFAM (*Protein Families Database*) relacionados con LS especiales y discrimina hasta 10 localizaciones. Cuando no se encuentra ningún dominio adecuado en PFAM las predicciones entonces se complementan con métodos basados en aprendizaje automático (AA). PSLT2 (*Protein Subcellular Localization T2*) (Scott et al., 2005) es un predictor especializado en hongos que procesa, por medio de una red Bayesiana, varias características que derivan de la búsqueda de dominios transmembrana, péptidos señal y motivos en InterPro (*Integrated Protein Database*). Sobre bases similares se desarrolló un método para la búsqueda de co-ocurrencia de dominios SMART (*Simple Modular Architecture Research Tool*) (Mott et al., 2002). El continuo crecimiento de las bases de datos de motivos conocidos asociados con las diferentes localizaciones, posiblemente ayudará en las tareas de predicción.



### 1.2.2.3 Métodos basados en secuencias

Los métodos basados en AA son particularmente adecuados para predecir la LS de una proteína a partir de su secuencia de residuos, sin depender de la existencia de secuencias homólogas anotadas ni del reconocimiento de motivos y dominios específicos (Schneider and Fechner, 2004). Se ha desarrollado un gran número de predictores de LS y se ha adoptado un amplio espectro de técnicas de AA, que van desde Modelos Ocultos de Markov (HMM) a Redes Neuronales (RN) y Máquinas de Soporte Vectorial (MSV) (Schneider and Fechner, 2004). Desafortunadamente, la mayoría de los predictores disponibles se entrenan con bases de datos de secuencias que comparten hasta un 90% de identidad; por esta razón, los rendimientos alcanzados apenas describen la capacidad de generalización real del predictor y su utilidad en el tratamiento de los problemas del "mundo real". Los métodos entrenados en bases de datos redundantes alcanzan rendimientos comparables (o incluso inferiores) a los obtenidos con una simple búsqueda de similitud (Pierleoni et al., 2006), ya que dos secuencias con más de un 30% de identidad, suelen compartir el mismo compartimento sub-celular (Nair and Rost, 2002).

Las herramientas computacionales que buscan señales de localización explícitas tales como péptidos señal o péptidos de tránsito, también se consideran como predictores de LS basados en secuencia. Entre ellos, SignalP (Petersen et al., 2011), LipoP (Juncker et al., 2003), TargetP (Emanuelsson et al., 2007) y PProwler (*Protein Prowler*) (Hawkins and Boden, 2006), predicen la presencia de un péptido señal para la secreción o el tránsito para dirigir la proteína a un organelo. En un estudio realizado por Bagos y colaboradores en el 2008 quedó evidenciada la superioridad del SignalP con respecto a otros predictores de péptido señal (Bagos et al., 2008). Adicionalmente, en un estudio realizado en particular en micobacterias se determinó que, entre nueve algoritmos de predicción diferentes, el algoritmo de Modelos Ocultos de Markov (HMM) del SignalP era el que mejor comportamiento tenía para la predicción de la presencia de péptido señal en proteínas de este tipo de microorganismo (Leversen et al., 2009). En el caso del LipoP, el estudio realizado por Bagos demostró que, a pesar de que este programa no había sido diseñado

para predicción en bacterias Gram-positivas (y por ende no entrenado con proteínas provenientes de esta clase de bacterias) (Juncker et al., 2003), tenía un comportamiento muy cercano al mostrado por el PRED-LIPO y superior al mostrado por otros predictores de lipopéptido señal (Bagos et al., 2008).

#### **1.2.2.3.1 Identificación de proteínas integrales de membrana**

Las proteínas integrales de membrana se dividen en dos clases estructurales, hélice- $\alpha$  y barriles- $\beta$  transmembrana (BBTMs). Ambas clases realizan funciones críticas y tienen un gran potencial para emplearse en vacunas o en el desarrollo de fármacos. Para la predicción de las primeras, TMHMM es uno de los programas de mejor desempeño. En el 2001, en un estudio comparativo mostró el mejor rendimiento entre 14 programas (Moller et al., 2001). En otra evaluación en la que se analizaron 27 métodos, se encontró que la versión 1.0 del TMHMM se encontraba entre los procedimientos más eficaces (no se evaluó la versión 2.0) (Chen et al., 2002). Más recientemente en un trabajo se ubicó al TMHMM como el número 3 entre otros 10 programas (Sadovskaya et al., 2006). Una última evaluación determinó que PolyPhobius y MEMSAT-SVM eran más eficaces que el TMHMM, pero el primero necesitaba entre 4 y 11 minutos para predecir una proteína y el segundo requería aún de más tiempo, mientras el TMHMM podía predecir más de 100 proteínas por minuto (ordenadores de un núcleo, AMD Opteron 2431) (Reeb et al., 2014). Entre los sistemas de identificación de BBTMs el algoritmo desarrollado por Freeman y Wimley, según sus autores, posee una precisión comparable a la de la predicción de proteínas hélice- $\alpha$  (Freeman and Wimley, 2010), superior a la de programas como BOMP (*Beta-Barrel Outer Membrane Protein Predictor*) (Berven et al., 2004) y TMB-Hunt (Garrow et al., 2005a, Garrow et al., 2005b).

#### **1.2.2.4 Principales métodos para la predicción de LS, basados en secuencia**

Cuando diferentes predictores están disponibles, es difícil de determinar su rendimiento relativo, sobre todo si se tiene en cuenta que se han implementado a través de los años con diferentes conjuntos de datos y que algunos de los conjuntos de entrenamiento pueden

haber sido redundantes con respecto al conjunto de prueba empleado. Por lo tanto, se necesitan evaluaciones independientes y uniformes. Recientemente se han publicado diferentes trabajos que abordan este problema. En el 2005 Klee y Ellis (Klee and Ellis, 2005), realizaron una evaluación general que indicó que, entre seis métodos, TargetP fue el mejor predictor de proteínas eucariotas secretadas. En el año 2006 Sprenger y col. (Sprenger et al., 2006) realizaron una comparación más general de predictores de LS que consideró cinco métodos y nueve tipos de LS (incluyendo la membrana). El Wolf PSORT (Horton et al., 2007) fue superior a los otros métodos basados en secuencia, especialmente al discriminar localizaciones dotadas con el mayor número de secuencias conocidas: núcleo, citosol, la mitocondria y regiones extracelulares. El mismo año se implementó BaCelLo (*Balanced subCellular Localization predictor*) (Pierleoni et al., 2006) y se comparó el comportamiento predictivo de 11 métodos, para discriminar cinco clases de LS. BaCelLo, TargetP (Emanuelsson et al., 2007) y LocTree3 (Nair and Rost, 2005) emergieron como los mejores predictores.

### 1.2.3 Inmunoinformática y diseño de vacunas

Las primeras herramientas inmunoinformáticas para el diseño de vacunas se desarrollaron en la década de 1980 por DeLisi, Berzofsky y otros (He et al., 2010a). Los algoritmos de mapeo epitópico son los más empleados en las herramientas informáticas empleadas en el diseño de vacunas. Dado que los epítomos T se unen en forma lineal al antígeno leucocitario humano (HLA, por sus siglas en inglés), la interfaz entre ligandos y células T ahora se puede modelar con precisión. En consecuencia se ha desarrollado un gran número de algoritmos de mapeo de epítomos T (De Groot and Berzofsky, 2004, De Groot and Moise, 2007). Tal información es inmensamente valiosa para el desarrollo de nuevas vacunas, con fines de diagnóstico y para el estudio de la patología de las enfermedades infecciosas (De Groot et al., 1997, Bond et al., 2001, De Groot et al., 2002, Dong et al., 2004, McMurry et al., 2005, De Groot, 2006, Inaba et al., 2006, Koita et al., 2006).

### 1.2.3.1 Mapeo epitópico de células T

Aunque frecuentemente la inmunidad se atribuye únicamente al desarrollo de anticuerpos protectores, la respuesta inmune a los virus atenuados y las vacunas de subunidades es, en gran medida, dependiente del reconocimiento por parte de las células T, de epítomos peptídicos unidos al Complejo Mayor de Histocompatibilidad (MHC, por sus siglas en inglés). Las células T citotóxicas (CTLs) generalmente tienen un papel importante en la contención de las infecciones virales y bacterianas (Plotnicky et al., 2003), y la prevalencia de las CTL comúnmente se correlaciona con la tasa de aclaramiento del patógeno.

Los algoritmos actuales de predicción de epítomos T logran un alto grado de precisión en la predicción (en el rango de 90 a 95% de Valor Predictivo Positivo). Entre las herramientas disponibles a través de internet se incluyen el sitio SYFPEITHI (Rammensee et al., 1999) y una herramienta de predicción de unión al HLA, disponible en los Institutos Nacionales de Salud (BIMAS) (Parker et al., 1994). A través de la base de datos “Immune Epitope” está disponible un conjunto de herramientas recientemente desarrolladas. Cada una de estas herramientas ha sido descrita y validada (Wang et al., 2008). En la industria farmacéutica se emplea activamente el programa propietario, EpiMatrix, (De Groot and Martin, 2009). Aunque ninguno de estos sitios genera exactamente las mismas predicciones, todas estas son bastante exactas, sobre todo si se compara con los resultados obtenidos con herramientas de mapeo temprano de epítomos (por ejemplo, SYFPEITHI y BIMAS) (Wang et al., 2008, De Groot and Martin, 2009). En general, los algoritmos más nuevos y con mayor nivel de mantenimiento, tienden a superar a los viejos métodos predictivos estadísticos. Los predictores no lineales superan a los basados en matrices, y la mayoría de los predictores pueden mejorarse mediante transformaciones no lineales de sus puntuaciones primarias de predicción (Lin et al., 2008a). Mientras se ha logrado que un buen desempeño para las predicciones de MHC de clase I, los resultados en la predicción de epítomos de HLA de clase II todavía son limitados (Gowthaman and Agrewala, 2008, Lin et al., 2008b).

### 1.2.3.2 Mapeo epitópico de células B

Actualmente los recursos bioinformáticos disponibles para identificar antígenos de células B, son limitados. Si bien la respuesta humoral representa la primera línea de defensa contra la mayoría de los patógenos virales y bacterianos, la modelación *in silico* de la proteína blanco de este mecanismo de defensa es muy compleja y en algunos casos imposible de lograr (Liljeroos et al., 2015). Varias herramientas de predicción de epítomos B, incluyendo 3DEX (3D-Epitope-Explorer), CEP y Pepito, se encuentran en distintas etapas de desarrollo (Enshell-Seijffers et al., 2003, Kulkarni-Kale et al., 2005, Schreiber et al., 2005, Sweredoski and Baldi, 2008). Desafortunadamente, la complejidad de modelaje y los recursos computacionales que se requieren para la predicción de epítomos B son enormes. Esta complejidad se debe, en parte, a la flexibilidad inherente a las regiones determinantes de la complementariedad del anticuerpo y, en parte, atribuible a modificaciones postraduccionales tales como glicosilación, todo lo cual puede resultar en la modificación de los epítomos B (Liljeroos et al., 2015).

Aunque la mayoría de los epítomos son discontinuos (Barlow et al., 1986), la detección experimental de estos es principalmente para epítomos lineales. Existen herramientas para la predicción de epítomos B lineales, pero en general su capacidad predictiva no es alta (Blythe and Flower, 2005, Reimer, 2009). En un estudio se evaluaron cinco predictores contra un conjunto de epítomos B lineales identificados experimentalmente y todos mostraron un pésimo desempeño (Reimer, 2009). Uno de los programas empleados con este fin es BepiPred (Larsen et al., 2006). Si bien es cierto que existen otros programas como el BEOracle (B-Cell Epitope Oracle), que han mostrado mejor rendimiento que este (Wang et al., 2011), los mismos no están disponibles para su uso como programas de escritorio. Es importante tener en cuenta que a pesar de que este programa representa una mejora significativa con respecto a sus predecesores aún posee limitaciones. En general, es necesario mejorar el poder predictivo antes de que los predictores de epítomos B lineales ofrezcan predicciones confiables (Larsen et al., 2006). A pesar de esto, BepiPred continua

siendo un programa empleado con frecuencia (Bowman et al., 2011, Maritz-Olivier et al., 2012).

#### **1.2.4 Determinación de la función de las proteínas**

La predicción de la función de las proteínas es una tarea muy importante y difícil dentro de la bioinformática. Los métodos basados en homología estructural utilizan la estructura de la proteína, que se compara, con el empleo de técnicas de alineamiento estructural, con la de proteínas de función conocida, prediciendo la función a partir de la proteína que posea mayor similitud (Tiwari and Srivastava, 2014). Para la comparación de secuencias de aminoácidos se han propuesto BLAST (Altschul et al., 1990) y FASTA (Fast All) (Pearson and Lipman, 1988). Cuando se requiere la detección, entre proteínas distantes, de similitudes débiles, pero relevantes biológicamente, se emplea el PSI-BLAST (PairwiseStatSig-BLAST) (Altschul et al., 1997).

Las estructuras de las proteínas están más conservadas que las secuencias, de modo que cuando la predicción de la función basada en la secuencia no se puede lograr con una alta precisión, la estructura tridimensional de las proteínas se utiliza con este objetivo. Las similitudes conformacionales, globales y locales, entre las proteínas, son útiles para inferir funciones en las proteínas (Harrison et al., 2003). Capra y colaboradores (Capra et al., 2009) desarrollaron un método para predecir sitios de unión molecular que integra estimaciones de conservación de secuencia con métodos basados en estructura, para identificar cavidades en la superficie de las proteínas, paquetes de unión a ligandos, residuos de unión a ligandos individuales, sitios catalíticos y los bolsillos de unión a drogas. Las simulaciones de dinámica molecular se utilizaron con algoritmos de predicción de función basados en la estructura, para encontrar los sitios de unión para la predicción de la función de proteínas (Glazer et al., 2009). La principal limitación de este tipo de métodos de predicción es la poca disponibilidad de datos estructurales de alta resolución, de las proteínas o sus homólogos.

Los métodos basados en secuencia y estructura utilizan relaciones de homología entre las proteínas para la predicción de la función. Cuando la homología basada en la secuencia

falla, entonces la homología basada en estructura se utiliza para predecir la función de las proteínas. Sin embargo, estos métodos tienen problemas en la predicción debido a la escasa disponibilidad de datos adecuados de proteínas homólogas y a que estas pueden tener diferentes funciones. Así que cuando no se pueden establecer las relaciones de homología para las proteínas blanco, estos métodos fallan (Whisstock and Lesk, 2003). El alcance de la predicción de la función de la proteína basada en la estructura se ha restringido debido a que en las bases de datos solo está disponible un número limitado de estructuras y plegamientos. Para aquellas proteínas que tienen poca o ninguna similitud de secuencia con proteínas de función conocida, la predicción de la función a partir de la secuencia es un gran desafío. Las técnicas de inteligencia computacional han demostrado ser útiles en la predicción de la función, mediante el uso de propiedades derivadas de las secuencias, pero independientes de la similitud con otras secuencias de proteínas (Han et al., 2006).

#### **1.2.5 Métodos integradores en la predicción de blancos vacunales**

A pesar de que la VR se ha empleado por una década, a menudo este enfoque no es accesible para todos los laboratorios, debido a la carencia de programas fáciles de utilizar e implementar. Aunque muchos programas específicos están disponibles para ayudar en la predicción de blancos vacunales (Grandi, 2003, Grandi, 2004, Serruto et al., 2004, Oldstone, 2005, Sachdeva et al., 2005, Pierleoni et al., 2006, Nielsen et al., 2007, Chang et al., 2008, Ponomarenko et al., 2008, Sweredoski and Baldi, 2008, Zhou et al., 2008, Agrawal and Huang, 2009, Rubinstein et al., 2009, Sun et al., 2009, Yu et al., 2010, Giaco et al., 2012, Hayat and Elofsson, 2012, Magnus et al., 2012, Soares et al., 2012, Yao et al., 2012, Goldberg et al., 2014, Yu et al., 2014, Dobson et al., 2015, Tsirigos et al., 2015), estos están desarrollados de forma individual para diferentes propósitos y contienen dispares formatos de datos y ajustes de programación. Esto hace difícil la integración de las herramientas existentes y los datos generados. Adicionalmente, el uso exitoso de estos programas a menudo requiere de instalación local, ejecución en la línea de comandos y de un

considerable poder computacional, sin contar el hecho de que muchas de estas herramientas no están optimizadas para el procesamiento de datos de alto flujo.

En los últimos años se han realizado esfuerzos con el objetivo de superar las limitaciones antes expuestas. Así, en este entorno de la VR, se han desarrollado sistemas capaces de integrar diferentes programas independientes, con el objetivo de predecir blancos vacunales. Los principales sistemas publicados son Vaxign (He et al., 2010b), NERVE (New Enhanced Reverse Vaccinology Environment) (Vivona et al., 2006) y Jenner-Predict (Jaiswal et al., 2013). A excepción de NERVE, que es una aplicación de escritorio, el resto son aplicaciones disponibles a través de internet. Actualmente es imposible el empleo de estas herramientas, NERVE ya no recibe soporte y la aplicación no corre sobre los sistemas operativos actuales, y el resto tiene problemas de funcionamiento lo que imposibilita su empleo. Esto hace necesario el desarrollo de nuevas herramientas que sustituyan las mencionadas y mejoren su rendimiento.

### **1.3 Sistemas heterólogos de expresión**

#### **1.3.1 Expresión de proteínas recombinantes en *Escherichia coli***

Numerosos sistemas se han utilizado para la expresión de genes foráneos. Estos productos expresados fuera de su sistema antólogo, muchas veces no muestran sus características nativas, por lo cual se requiere encontrar no solo el sistema adecuado sino las condiciones en que la expresión se debe llevar a cabo (Dicker and Strasser, 2015, Henrich et al., 2015).

Las células de insectos constituyen un sistema efectivo para la expresión de proteínas heterólogas que requieren procesamiento post-traduccionales tales como proteólisis, glicosilación o secreción. Estos sistemas utilizan a los baculovirus como vectores y a pesar de las ventajas anteriores se les han reportado bajos niveles de expresión (Kroemer et al., 2015, Osz-Papai et al., 2015). Las levaduras también se han utilizado ampliamente como sistema de expresión, debido a su fácil manipulación genética y a las características de crecimiento similar a la de los organismos procariontes, conjuntamente con la presencia de los mecanismos post-traduccionales de los organismos superiores. Además, exhiben



determinadas preferencias por los codones utilizados por estos últimos (Aghcheh and Kubicek, 2015, Routledge et al., 2016). Sin embargo los productos se secretan a muy bajos niveles debido al fenómeno de hiperglicosilación que también puede afectar las propiedades nativas y su actividad biológica (Jacobs et al., 2008, Lau and Fong, 2008, Mahalik et al., 2014). La expresión de genes heterólogos en bacterias es un método simple y factible para la obtención de grandes cantidades de proteínas. *Escherichia coli* (*E. coli*) es el sistema de biosíntesis de proteínas más empleado y barato, esto satisface económicamente la experimentación con nuevas construcciones génicas y el escalado de la proteína recombinante para su producción (Meloni et al., 2006).

Se ha podido comprobar que el producto de muchos genes clonados no adquiere la estructura tridimensional adecuada al ser expresados en *E. coli*. En muchos casos, producto de la elevada expresión, la proteína recombinante, es secuestrada dentro de agregados ubicados en el citoplasma, conocidos como cuerpos de inclusión (Singh et al., 2015). Las proteínas activas deben ser, por tanto, recuperadas. Con este fin se usan agentes desnaturizantes que afectan el plegamiento nativo de la cadena polipeptídica y los métodos de renaturalización no son siempre efectivos (Correa and Oppezzo, 2015, Singh et al., 2015).

El sistema basado en la inducción de la expresión con Isopropil- $\beta$ -D-1-tiogalactopiranosido (IPTG), es el más empleado para producir proteínas recombinantes en *E. coli* (Ping et al., 2013). Sin embargo, en años recientes Studier desarrolló el método de autoinducción con lactosa (Studier, 2005). La autoinducción permite el monitoreo eficiente de la expresión y la solubilidad de varios clones en paralelo. Los cultivos se inoculan e incuban hasta la saturación (no hay necesidad de realizar monitoreos intermedios para determinar el momento óptimo de adición de IPTG) y los rendimientos de la proteína blanco son típicamente varias veces superiores a los que se obtienen por la inducción convencional con IPTG. En este sistema la lactosa se emplea como inductor y los niveles indeseados de expresión basal de la proteína blanco en los momentos iniciales del cultivo, se suprimen por la adición de glucosa al medio.

## 1.4 Micoplasmas y Vacunología Reversa

Aunque en investigaciones en el campo de la micoplasmosis se han aplicado algunas herramientas computacionales asociadas a la VR (Hamsten et al., 2008, Hamsten et al., 2009, Siqueira et al., 2013, Bumgardner et al., 2014, Reolon et al., 2014, Perez-Casal et al., 2015) el empleo de este concepto aún es muy escaso e insuficiente en los trabajos con este género de bacterias. Cuando se ha aplicado la VR, generalmente se ha hecho de manera muy limitada y nunca de manera estandarizada. Por ejemplo, para la identificación en una primera oportunidad, de proteínas de superficie de *Mycoplasma mycoides*, Hamsten y colaboradores solo usaron el programa LipoP para predecir lipoproteínas y TMHMM para hélices- $\alpha$  transmembrana (Hamsten et al., 2008), mientras, en un trabajo posterior, utilizaron el programa SignalP para predecir péptidos señal y nuevamente TMHMM (Hamsten et al., 2009). Por otra parte, para identificar potenciales blancos para el desarrollo de vacunas contra *Mycoplasma hyopneumoniae*, Siqueira y colaboradores, detectan hélices- $\alpha$  transmembrana combinando cuatro predictores diferentes (TMHMM, SCAMPI, SVMtm Transmembrane Domain Predictor y HMM-TM), además emplean Psortb y buscan ortólogos por homología (Siqueira et al., 2013). El grupo de Reolon también hizo esto último y empleó dos programas diferentes (LIPOPREDICT y LIPO CBU) para la detección de lipoproteínas (Reolon et al., 2014). Para identificar factores de virulencia que pudieran ser críticos en el desarrollo de vacunas contra *Mycoplasma hyosynoviae*, Bumgardner y colaboradores solo utilizaron una búsqueda de homólogos con el BLAST (Bumgardner et al., 2014). Por último, en una investigación reciente se emplea una metodología más completa que incluye PSORT para identificar proteínas de membrana y extracelulares, TMHMM, SPAAN (Software for Prediction of Adhesins and Adhesin-like proteins using Neural networks) que es el único programa que existe para la identificación de citadhesinas, LipoP y Vaxign para la identificación de antígenos vacunales (Perez-Casal et al., 2015).

Como se puede apreciar no existe consistencia en estos análisis, los algoritmos que se emplean son muy variados, y estos adolecen, en la mayoría de los casos, de carencia de herramientas para la detección de los más importantes factores de virulencia de los

micoplasmas. En la mayor parte de los trabajos no se tiene en cuenta la detección de citadhesinas (Hamsten et al., 2008, Hamsten et al., 2009, Siqueira et al., 2013, Bumgardner et al., 2014, Reolon et al., 2014), en algunos tampoco se predicen lipoproteínas (Hamsten et al., 2009, Siqueira et al., 2013, Bumgardner et al., 2014, Reolon et al., 2014), e incluso, cuando se utiliza un esquema de predicción más completo, que incluye la detección de este tipo de moléculas, el mismo presenta redundancias innecesarias en su análisis (Perez-Casal et al., 2015). En este último caso, después de utilizar Psort (que en su esquema de predicción incluye TMHMM), se emplea TMHMM. Esta redundancia en los análisis también se observa con la detección de citadhesinas, después de aplicar SPAAN, el esquema de predicción incluye Vaxign que utiliza SPAAN en su algoritmo. Adicionalmente, se emplea SPAAN en proteínas que fueron identificadas como extracelulares, lo cual es contradictorio, porque las proteínas extracelulares no pueden funcionar como citadhesinas (Perez-Casal et al., 2015).

## **2 Capítulo 2. PREDICCIÓN DE ANTÍGENOS DE SUPERFICIE DE *M. gallisepticum* CON EL EMPLEO DE Bact-VAP**

### **2.1 Introducción**

La identificación, con el apoyo de herramientas bioinformáticas, de nuevos antígenos de superficie de *M. gallisepticum*, es una alternativa atractiva, a la identificación experimental de los mismos, que permitiría un ahorro sustancial de tiempo y recursos. El sistema Bact-VAP desarrollado en el laboratorio de Genómica Microbiana y Bioinformática del CENSA, podría permitir alcanzar este objetivo. Validar el desempeño de esta herramienta es importante para garantizar resultados confiables.

Objetivos:

1. Validar la herramienta Bact-VAP para la identificación de potenciales candidatos vacunales bacterianos y antígenos de superficie de micoplasmas.
2. Predecir antígenos de superficie de *M. gallisepticum* con el empleo de Bact-VAP.

Tareas específicas:

1. Validar la capacidad predictiva de Bact-VAP para identificar potenciales blancos vacunales bacterianos, así como proteínas de superficie de *M. gallisepticum*, mediante la determinación de sus parámetros de desempeño
2. Predecir potenciales proteínas de superficie antigénicas de *M. gallisepticum* con el empleo de Bact-VAP.

## 2.2 Materiales y Métodos

### 2.2.1 Validación del predictor de BBTMs ( $\beta$ -Barrel Predictor)

Para predecir las proteínas de tipo barril- $\beta$  se utilizó una aplicación que se desarrolló *en casa* y se implementó en <http://baco.censa.edu.cu:8088/predictor/betaB/index>, que se basa en un algoritmo desarrollado por Wimley y Feeman (Freeman and Wimley, 2010). En una primera aproximación no se le realizaron modificaciones a este algoritmo, pero con el objetivo de mejorar el desempeño de la aplicación, se le incluyó el filtro de abundancia de aminoácidos relativos que emplea el BOMP (Berven et al., 2004), que permite eliminar falsos positivos (predictor de BBTMs-Plus). Para la evaluación se generó un conjunto de validación (Anexo 1), que fue empleado para determinar los parámetros de desempeño del programa y realizar una comparación objetiva con otros predictores de barriles- $\beta$ .

#### 2.2.1.1 Construcción del conjunto de validación

El conjunto de validación se construyó con 535 proteínas cuyas estructuras se confirmaron en el *Protein Data Bank* (PDB): <http://www.rcsb.org/pdb/home/home.do>. Se generaron dos subconjuntos: el conjunto de validación positivo (CVP), integrado por 102 proteínas, y el negativo (CVN), formado por 433 proteínas. Para la creación del CVP se eligieron proteínas con estructura BBTM conocida; el CVN se construyó con proteínas que no poseían estructura BBTMs. El CVP se creó inicialmente con secuencias de proteínas bacterianas de 24 superfamilias de barriles-beta transmembrana a partir de la base de datos *Orientations of Proteins in Membrane* (OPM), la cual se encuentra disponible en <http://opm.phar.umich.edu/>. Además se incluyeron otros BBTMs diferentes con estructura resuelta. Estas últimas BBTMs surgieron como resultado de una búsqueda actualizada en las bases de datos de los principales predictores BBTMs. El CVN se extrajo del PDB de proteínas transmembrana (PDBTM), disponible en <http://pdbtm.enzim.hu/>, al filtrar todas las secuencias de proteínas

que fueran barril-beta. Además, ambos conjuntos se inspeccionaron de forma manual mediante una búsqueda en el PDB.

Para evitar la existencia de secuencias redundantes, todas las secuencias de ambos conjuntos con una identidad > 50% se filtraron con la herramienta BlastClust (<http://toolkit.tuebingen.mpg.de/blastclust>).

En ambos casos, tanto para el CVP como para el CVN, el filtrado de secuencias restantes se refinó aún más por la exclusión de las proteínas que estuvieran fuera del rango entre 60 y 4000 residuos de longitud, para cumplir con las restricciones de longitud de cadena del algoritmo de predicción de Freeman- Wimley.

#### **2.2.1.2 Evaluación del desempeño de BBTMs**

Para la evaluación del desempeño del predictor de BBTMs, se construyó una curva del tipo *Receiver Operating Characteristic* (ROC), las cuales son herramientas útiles para evaluar clasificadores en aplicaciones de bioinformáticas (Robin et al., 2011).

#### **2.2.1.3 Comparación con otros predictores BBTM**

Para evaluar su desempeño, el predictor de BBTMs se comparó con otros programas de predicción de BBTMs que permitían procesar gran cantidad de proteínas, incluso, proteomas completos. Así el sistema implementado se comparó, mediante el cálculo de algunos parámetros tales como: sensibilidad (probabilidad de que una proteína positiva sea identificada como tal), especificidad (probabilidad de que una proteína negativa sea identificada como tal), coeficiente de correlación de Matthews (MCC, por sus siglas en inglés) (medida estadística de la asociación entre dos variables binarias), valor predictivo positivo (VPP) (probabilidad de que una proteína sea positiva si el resultado es positivo) y exactitud, con el programa BOMP (<http://services.cbu.uib.no/tools/bomp>) (Berven et al.,

2004) y la versión de escritorio de la aplicación TMB-Hunt (Garrow et al., 2005a, Garrow et al., 2005b).

Estos parámetros se calcularon según las fórmulas que aparecen a continuación:

$$\text{Sensibilidad} = \frac{VP}{VP+FN}$$

$$\text{Especificidad} = \frac{VN}{VN+FP}$$

$$MCC = \frac{VP \times VN - FP \times FN}{\sqrt{(VP+FN)(VP+FP)(VN+FN)(VN+FP)}}$$

$$\text{Exactitud} = \frac{VP+VN}{VP+VN+FP+FN}$$

$$VPP = \frac{VP}{VP+FP}$$

Donde: VP: verdaderos positivos predichos; VN: verdaderos negativos predichos; FP: falsos positivos predichos; FN: falsos negativos predichos; MCC: coeficiente de correlación de Matthews; VPP: valor predictivo positivo.

### 2.2.2 Adiciones al BepiPred

Aunque aproximadamente el 90% de los epítomos B son de naturaleza conformacional (Yasser and Honavar, 2010), y la efectividad de los programas existentes para la predicción de epítomos B lineales es aún baja, el hecho de trabajar en la predicción de potenciales candidatos vacunales hacía imprescindible incluir algún programa que abordara este campo.

En el caso del BepiPred, en su versión de escritorio (Larsen et al., 2006), se consideró necesario adicionar algunos cálculos complementarios a la salida del sistema, diseñados en el grupo de Genómica Microbiana y Bioinformática del CENSA, para calcular el valor total de antigenicidad de cada proteína. El programa identifica los posibles aminoácidos involucrados en la formación de determinantes antigénicos. Esta información por sí misma, si bien útil, no permitía identificar que proteínas eran las más antigénicas. La adición, de este modo

permite organizar las proteínas según su antigenicidad. Este valor global de antigenicidad se calculó de la siguiente manera:

$$\begin{aligned} \text{Puntuación Proteína} &= \sum_{i=1}^a \text{Puntuación Péptido}_i \\ \text{Puntuación Péptido} &= \frac{\sum_{i=1}^b \text{Puntuación Epítipo}_i}{\text{Nº Epítipos-Péptido}} \\ \text{Puntuación Epítipo} &= \frac{\sum_{i=1}^c \text{Puntuación Amino Ácido}_i}{\text{Nº Amino Ácidos-Epítipo}} \\ \text{Nº Epítipos} &= \text{int} \left[ 1 + \frac{(\text{aaP} - 12)}{4} \right] \end{aligned}$$

Donde:

- a**  $\Rightarrow$  es el número de péptidos antigénicos en la proteína (según salida del BepiPred)
- b**  $\Rightarrow$  es el número de epítipos en cada péptido antigénico
- c**  $\Rightarrow$  es el número de amino ácidos en cada epítipo identificado
- aaP**  $\Rightarrow$  es el número de amino ácidos del péptido antigénico
- 12**  $\Rightarrow$  es la longitud media del epítipo B
- 4**  $\Rightarrow$  es la ventana de desplazamiento para considerar la formación de un nuevo epítipo
- Si  $\text{aaP} - 12 \leq 0$ , entonces  $\text{Nº epítipos} = 1$
- Si el péptido identificado por el BepiPred contiene menos de ocho amino ácidos, entonces se desecha (los epítipos B se encuentran entre cuatro y 12 aminoácidos, aunque por regla general, no tienen menos de ocho (Buus et al., 2012)).

### 2.2.3 Determinación del valor de corte óptimo del programa SPAAN para micoplasmas

Con el objetivo de determinar el valor de corte óptimo de este programa para identificar correctamente citadhesinas de micoplasmas se emplearon dos conjuntos de datos. El positivo, integrado por 40 citadhesinas de micoplasmas identificadas experimentalmente (Tabla 1) y el negativo, por 1071 proteínas de micoplasmas, identificadas como citoplasmáticas en la base de datos UniProt (<http://www.uniprot.org/uniprot/?query=mycoplasma+keyword:963#>) (Anexo 2). De estas, una proteína no fue procesada por el sistema, ya que genera un error a partir del análisis de



la misma. Por esta razón la validación se realizó con los resultados obtenidos para 1070 proteínas negativas. Para evitar la existencia de secuencias redundantes en el conjunto de evaluación, todas las secuencias, con una identidad > 50%, se filtraron con la herramienta BlastClust (<http://toolkit.tuebingen.mpg.de/blastclust>).

**Tabla 1.** Citadhesinas empleadas como conjunto de datos positivos, para evaluar el programa SPAAN

Gen	Proteína	ID GenBank	Referencias
vlhA.4.03a	VlhA.4.03 variable lipoprotein family protein N-terminal domain	ADB96857.1	(Papazisi et al., 2003)
vlhA.4.10	VlhA.4.10 variable lipoprotein family protein	AAP56553.1	(Papazisi et al., 2003)
vlhA.5.01a	VlhA.5.01 variable lipoprotein family domain protein	AAP56686.2	(Papazisi et al., 2003)
vlhA.5.01b	VlhA.5.01 variable lipoprotein family domain protein	ADB96873.1	(Papazisi et al., 2003)
vlhA.5.01c	VlhA.5.01 variable lipoprotein family domain protein	ADB96874.1	(Papazisi et al., 2003)
vlhA.5.04	VlhA.5.04 variable lipoprotein family protein	AAP56689.1	(Papazisi et al., 2003)
vlhA.5.10b	VlhA.5.10 variable lipoprotein family protein C-terminal domain	ADB96875.1	(Papazisi et al., 2003)
vlhA.1.02	VlhA.1.02 variable lipoprotein family protein	AAP56773.1	(Papazisi et al., 2003)
vlhA.1.05	VlhA.1.05 variable lipoprotein family protein	AAP56776.2	(Papazisi et al., 2003)
vlhA.1.08a	VlhA.1.08 variable lipoprotein family protein N-terminal domain	ADB96882.1	(Papazisi et al., 2003)
vlhA.2.01	VlhA.2.01 variable lipoprotein family protein	AAP56805.2	(Papazisi et al., 2003)
vlhA.2.02	VlhA.2.02 variable lipoprotein family protein	AAP56806.2	(Papazisi et al., 2003)
vlhA.3.0.1	VlhA.3.0.1 variable lipoprotein family domain protein	ADB96891.1	(Papazisi et al., 2003)
mgpA	adhesin P1	ADK87038.1	(Xiao et al., 2015)
MAG_2410	P40, lipoprotein	YP_001256383.1	(Bird et al., 2008)
P97	protein p97; cilium adhesin	AAV27763.1	(Deutscher et al., 2010)
P102	protein p102	AAV27762.1	(Seymour et al., 2011b)
mhp493	conserved hypothetical protein, p97 paralog	AAV27917.1	(Wilton et al., 2009)
p146	p146 adhesin like-protein, p97 paralog	AAV28029.1	(Bogema et al., 2012)
mhp108	conserved hypothetical protein, p102 paralog	AAV27417.1	(Seymour et al., 2010)
MHP7448_0662	hypothetical protein	YP_288047.1	(Bogema et al., 2011)
mgpC	adhesion protein MG192	AFU86193.1	(Burgos et al., 2012)
mgpA	MgPa adhesin	AAC71410.1	(Zeng et al., 2012)
P30	adhesin protein	ABR09215.1	(Hasselbring et al., 2012)
mhp384	conserved hypothetical protein, p102 paralog	AAV27855.1	(Deutscher et al., 2012)
mhp385	conserved hypothetical protein, p97 paralog	AAV27856.1	(Deutscher et al., 2012)
vsaG	VsaG	AAF70134.1	(Bolland and Dybvig, 2012)
mhp107	conserved hypothetical protein, p97 paralog	AAV27416.1	(Seymour et al., 2011a)
eno	Enolase	AAP56855.1	(Chen et al., 2011)
crmA	cytadherence related molecule A (CrmA)	AAP56531.1	(Papazisi et al., 2002)
plpA	Pneumoniae-like protein A (plpA)	AAP56670.2	(Papazisi et al., 2003)
gapA	cytadhesin protein GapA	AAP56530.2	(Papazisi et al., 2000)
pvpAb	phase-variant protein A (pvpA) domain protein	ADB96886.1	(Bogoslavsky et al., 2000)
hmw1	HMW1 cytadherence high molecular weight protein 1	ADK86732.1	(Schmidl et al., 2009)
hmw2	HMW2 Cytadherence high molecular weight protein 2	P75471 *	(Schmidl et al., 2009)
hmw3	HMW3 cytadherence high molecular weight protein 3	ADK86713.1	(Schmidl et al., 2009)
tuf	EFTU Elongation factor Tu	P23568 *	(Balasubramanian et al., 2006)
p110	P110 protein	AAC71411.2	(Musatovova et al., 2006)
mgpB	P140	AGH30699.1	(Burgos et al., 2006)
vspN	variable surface lipoprotein VspN	YP_004056589.1	(Thomas et al., 2005)

Los parámetros empleados para realizar la comparación fueron los mismos usados para validar el Predictor de BBTMs (epígrafe 2.2.1.3): sensibilidad, especificidad, MCC, precisión y VPP. Estos parámetros se calcularon para cuatro valores de corte diferentes (0,5; 0,55;

0,6; y 0,7), que se movían entre los dos valores de corte que mejores resultados mostraron en diferentes evaluaciones que se le realizaron al programa en el proceso de estandarización (Sachdeva et al., 2005).

#### **2.2.4 Algoritmos de predicción y de priorización empleados por Bact-VAP para la predicción de potenciales blancos vacunales**

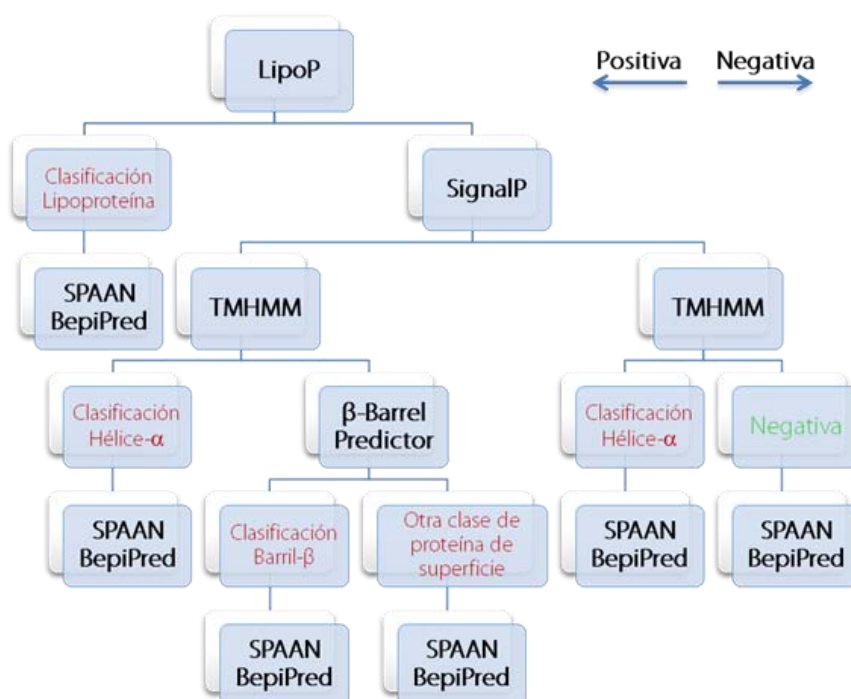
Bact-VAP integra herramientas de código abierto disponibles, y programas desarrollados *in house* (en el laboratorio de Genómica Microbiana y Bioinformática del CENSA), que permiten predecir clases de proteínas y características de las mismas como su tipo, función y antigenicidad. Este algoritmo incluyó los siguientes componentes:

- **LipoP v 1.0** - Herramienta que permite predecir lipoproteínas y diferenciar lipopéptidos señal, otros péptidos señales y hélices de membrana n-terminal en bacterias Gram Negativas. Este predictor está basado en Modelos Ocultos de Markov (HMM por sus siglas en inglés) (Juncker et al., 2003). La versión de escritorio de este programa (lipop-1.0a) se obtuvo del sitio <http://www.cbs.dtu.dk/services/LipoP>.
- **SignalP v 4.0** - Herramienta que permite predecir la presencia y localización del sitio de anclaje de péptidos señal en secuencias de aminoácidos de diferentes organismos: Gram positivas procariotas, Gram negativas procariotas y eucariotas. Este predictor está basado en Redes Neuronales Artificiales (ANN por sus siglas en inglés) (Petersen et al., 2011). La versión de escritorio de este programa (signalp-4.0d) se obtuvo del sitio <http://www.cbs.dtu.dk/services/SignalP>.
- **TMHMM v 2.0** - Aplicación que está diseñada para predecir y modelar la localización y topología de proteínas hélice- $\alpha$ . Está basado en HMM y ANN (Krogh et al., 2001). La versión de escritorio de este programa (tmhmm-2.0c) se obtuvo del sitio <http://www.cbs.dtu.dk/services/TMHMM-2.0>.
- **Predictor de BBTMs** - Wimley-Freeman es un algoritmo para la predicción de las estructuras de tipo barril beta en proteínas transmembrana (BBTMs) basado en las

características físico-químicas de las proteínas, este ofrece una capacidad de precisión de hasta 99% con un MCC de 0,806, superando algoritmos previamente publicados (Freeman and Wimley, 2010). El algoritmo se implementó *in house* (Laboratorio de Genómica Microbiana y Bioinformática, CENSA).

- **SPAAN v 1.0** - Aplicación que permite la predicción de adhesinas, o de proteínas similares a adhesinas, basada en ANN (Sachdeva et al., 2005). La versión de escritorio de este programa (spaan\_64\_bit) se obtuvo del sitio <ftp://203.195.151.45>. El valor de corte de la probabilidad de ser citadhesina, se fijó en  $\geq 0,7$ .
- **BepiPred v 1.0** - Herramienta utilizada para la predicción de la localización de epítopes B lineales en proteínas. Basada en HMM y escala de propensión (Larsen et al., 2006). La versión de escritorio de este programa (BepiPred-1.0b) se obtuvo del sitio <http://www.cbs.dtu.dk/services/BepiPred>.

Estos componentes se combinaron como se describe en la Figura 1. Este sistema se implementó como servicio en el CENSA y está disponible en <http://baco.censa.edu.cu:8088/predictor/>.



**Figura 1.** Algoritmo empleado por Bact-VAP para la predicción de antígenos bacterianos de superficie. El sistema incluye la predicción de lipoproteínas y citadhesinas (importantes factores de virulencia bacterianos y en particular de los micoplasmas), así como de la estructura de las proteínas integrales de membrana, lo que puede emplearse para la identificación de las regiones proteicas expuestas. Estas son las que interactúan con el entorno, en particular, con el sistema inmune del hospedero.

Para ordenar las proteínas en el listado de salida se empleó el siguiente algoritmo de priorización:

1. Proteínas de superficie, que fueran citadhesinas y antigénicas. Ordenadas de manera tal que las primeras fueran lipoproteínas y luego, según el valor del SPAAN (si hay dos o más con el mismo valor, el lugar se decide por el valor de antigenicidad).
2. Proteínas de superficie y citadhesinas. Ordenadas de manera tal que las primeras fueran lipoproteínas y luego, según el valor del SPAAN (si hay dos o más con el mismo valor, el lugar se decide por el orden en el genoma).
3. Citadhesinas y antigénicas. Ordenadas según el valor del SPAAN (si hay dos o más con el mismo valor, el lugar se decide por el valor de antigenicidad).
4. Citadhesinas. Ordenadas según el valor del SPAAN (si hay dos o más con el mismo valor, el lugar se decide por el orden en el genoma).
5. Proteínas de superficie, y antigénicas. Ordenadas de manera tal que las primeras fueran lipoproteínas y luego, según el valor de antigenicidad (si hay dos o más con el mismo valor, el lugar se decide por el orden en el genoma).
6. Proteínas de superficie. Ordenadas de manera tal que las primeras fueran lipoproteínas y luego, según el orden en el genoma.
7. Proteínas que no se identificaron como de superficie pero que son antigénicas. Ordenadas según el valor de antigenicidad (si hay dos o más con el mismo valor, el lugar se decide por el orden en el genoma).
8. Y por último, las proteínas que no se identificaron como de superficie.

### **2.2.5 Validación de Bact-VAP**

#### **2.2.5.1 Validación de capacidad de predicción de candidatos vacunales protectores**

La validación de Bact-VAP se llevó a cabo en dos pasos. En una primera etapa se investigó la capacidad del sistema de identificar sensible y específicamente, potenciales candidatos vacunales bacterianos. Para esto se comparó el desempeño de Bact-VAP en estudio, con

respecto al que mostraron los programas NERVE, Vaxign, VaxiJen y Jenner-Predict. En este estudio se emplearon los mismos conjuntos de datos que cuando estos programas se evaluaron previamente, con candidatos vacunales protectores (PVCs, por sus siglas en inglés) identificados experimentalmente (Jaiswal et al., 2013). Estos conjuntos de validación se subdividen como se describe a continuación: 26 antígenos protectivos de *Streptococcus pneumoniae* (gram +); 18 de *Escherichia coli* (gram -); 177 provenientes de más de 40 bacterias, curados en función de su localización en la superficie celular y de la similitud entre ellos, a partir de la base de datos Protegen (Yang et al., 2011); por último un CVP de 88 y un CVN de 33 proteínas, obtenidos de los conjuntos de entrenamiento empleados en el desarrollo de VaxiJen (Doytchinova and Flower, 2007), luego de curarlos en función de su localización en la superficie celular. Esto permitió estudiar la capacidad de predecir antígenos vacunales del mismo y contrastarlo con la de los programas descritos en la literatura que realizan las mismas funciones (Vivona et al., 2006, Doytchinova and Flower, 2007, He et al., 2010b, Jaiswal et al., 2013).

#### **2.2.5.2 Validación de capacidad de predicción de proteínas de superficie de micoplasmas.**

En un segundo paso se estudió el potencial de la metodología para identificar proteínas de superficie de micoplasmas. Los resultados se compararon con los obtenidos con otros tres predictores de localización celular: Psort v3.0 (<http://www.psort.org/psortb>) (Yu et al., 2010), CELLO (<http://cello.life.nctu.edu.tw>) (Yu et al., 2014) y LocTree3 (<https://www.rostlab.org/services/loctree3>) (Goldberg et al., 2014). En todos los casos, para hacer válida la comparación, se empleó un único conjunto de datos generado a partir de 13 especies diferentes de micoplasmas, representativos desde el punto de vista filogenético (Anexo 2). El conjunto positivo estaba formado por 274 secuencias de proteínas de superficie de estas bacterias, algunas identificadas experimentalmente y recopiladas en una base de datos de factores de virulencia bacterianos (<http://www.mgc.ac.cn/cgi-bin/VFs/genus.cgi?Genus=Mycoplasma>) y otras extraídas, por curación manual de la

literatura, con información validada experimentalmente (Anexo 2) . El conjunto negativo, por otro lado, por 1071 secuencias de proteínas de micoplasmas obtenidas de UniProt (<http://www.uniprot.org/uniprot/?query=mycoplasma+keyword:963#>), e identificadas como citoplasmáticas en esta base de datos (Anexo 2). Para evitar la existencia de secuencias redundantes, todas las secuencias de ambos conjuntos con una identidad > 50% se filtraron con la herramienta BlastClust (<http://toolkit.tuebingen.mpg.de/blastclust>).

Los parámetros empleados para realizar la comparación fueron los mismos usados para validar el Predictor de BBTMs (epígrafe 2.2.1.3): sensibilidad, especificidad, MCC, precisión y VPP.

#### **2.2.6 Identificación de proteínas de superficie antigénicas de *M. gallisepticum* con el empleo de Bact-VAP**

La secuencia deducida de proteínas de la cepa R(low) de *M. gallisepticum*, se descargó de la base de datos del NCBI (Nº Acceso: AE015450.2, GI: 284811830). El proteoma deducido de este microorganismo se procesó con Bact-VAP.

### **2.3 Resultados y Discusión**

#### **2.3.1 Validación del predictor de BBTMs**

##### **2.3.1.1 Evaluación del desempeño**

Los gráficos ROC muestran el rendimiento de un método de clasificación binaria con salida ordinal continua o discreta (Robin et al., 2011). Se calculó la curva ROC para el predictor de BBTMs (Figura 2), con el empleo del conjunto de validación construido (Anexo 1). El valor del área bajo la curva (AUC) obtenido fue de 0,93. En el contexto ROC, el AUC mide el rendimiento de un clasificador. Este parámetro se encuentra entre cero y uno. Un valor de

AUC cercano a uno significa una mejor clasificación o rendimiento (Robin et al., 2011). La calidad del valor de AUC es relativa, en dependencia del contexto en que se calcule. Si se tiene en cuenta que para herramientas bioinformáticas similares en el campo de localización sub-celular, los valores de AUC oscilan en lo fundamental entre 0,5 y 0,9, con muy pocos por encima de este último y varios por debajo de 0,5 (Casadio et al., 2008, Shen and Burger, 2010), se puede clasificar el valor que alcanza el predictor de BBTMs como excelente. Además, para un 3% de falsos positivos, el valor de corte fue de 90,39. Las proteínas cuyo valor  $\beta$ -barrel es mayor o igual a este valor, se consideraron BBTMs.



**Figura 2.** Evaluación del rendimiento del Predictor de BBTMs. Curva ROC calculada para el sistema, con un valor de 0,93. Para un 3% de falsos positivos, el valor de corte calculado es 90,39

### 2.3.1.2 Comparación con otros predictores BBTM

Para continuar la evaluación del rendimiento del sistema, se calcularon algunos parámetros: sensibilidad, especificidad, VPP, la exactitud y el MCC (Tabla 2). Como se puede apreciar, con el empleo del algoritmo desarrollado por Freeman y Wimley (Freeman and Wimley, 2010) el Predictor de BBTMs mostró un rendimiento superior al del TMB-Hunt, pero inferior al del programa BOMP (con la excepción de la sensibilidad). Esto se debió, probablemente, a los dos análisis adicionales que lleva a cabo el BOMP: el reconocimiento de patrones del extremo C-terminal y la eliminación de falsos positivos, el cual se realiza mediante el procedimiento de filtrado final de los valores de abundancia de aminoácidos relativos, para discriminar resultados positivos de negativos (Berven et al., 2004).

Si se tiene en cuenta que el predictor de barriles emplea la última versión del algoritmo de Freeman-Wimley, que mostró resultados superiores a cualquier otro algoritmo publicado previamente para distinguir BBTMs de otras proteínas (Freeman and Wimley, 2010), debía esperarse que, si se le adicionan el reconocimiento de patrones del extremo C-terminal y el filtro para limitar el número de falsos positivos (o solo alguno de ellos), el Predictor de BBTMs ofreciera un mejor desempeño que el resto de los programas empleados en la comparación. Como se observa en la Tabla 2, esto fue confirmado con los resultados obtenidos. Con la única excepción de la sensibilidad, que es ligeramente inferior que la que del sistema sin el filtro, en el resto de los parámetros la adición del filtro (Predictor de BBTMs-Plus) elevó los valores por encima del resto de los programas. La disminución de una centésima de sensibilidad con respecto al sistema sin el filtro puede explicarse por el hecho de que el filtro incorporado elimina falsos positivos, pudiendo eliminar, por error, algún verdadero positivo. En compensación, se ganó en especificidad, lo que es deseable en las predicciones que luego tienen que ser evaluadas en el laboratorio ya que concentra los esfuerzos experimentales en los candidatos más promisorios.

**Tabla 2.** Comparación del desempeño de los sistemas Beta Predictor, BOMP y TMB-Hunt

	VP	FN	VN	FP	Sensibilidad	Especificidad	MCC	VPP	Exactitud
<b>Predictor de BBTMs</b>	87	15	412	21	0,853	0,952	0,787	0,806	0,933
<b>Predictor de BBTMs-Plus</b>	86	16	431	2	0,843	0,995	0,889	0,977	0,966
<b>BOMP</b>	78	24	430	3	0,765	0,993	0,830	0,963	0,95
<b>TMB-Hunt</b>	81	21	404	29	0,794	0,933	0,707	0,736	0,907

Donde VP-verdaderos positivos predichos; FN-falsos negativos predichos; VN-verdaderos negativos predichos; FP-falsos positivos predichos; MCC-coeficiente de correlación de Matthews y VPP valor predictivo positivo.

Las proteínas barriles-beta tienen potencial para emplearse en vacunas o componentes de vacunas como algunas proteínas de *Mannheimia haemolytica* (Ayalew et al., 2010), la Omp87 de *Pasteurella multocida* serogrupo B:2 (Kumar et al., 2013) y la BmaA de la cepa altamente patógena Nagasaki (serovar 5) de *Haemophilus parasuis* (Pina-Pedrero et al., 2012). En muchos casos, las BBTMs interactúan con las células del sistema inmune del hospedero y se pueden considerar como patrones moleculares asociados a patógenos (PAMP) debido a su capacidad para señalar por la vía de las moléculas receptoras de tipo



Toll y otros receptores de reconocimiento de patrones (McClean, 2012). Un ejemplo de esto es la familia OmpA de *Escherichia coli* (Confer and Ayalew, 2013).

El algoritmo de Freeman-Wimley se implementó en un programa de código abierto altamente eficiente y flexible. La construcción de un conjunto de validación de alta calidad garantizó que la comparación del sistema con otros existentes fuera confiable. Teniendo en cuenta que TMB-Hunt y BOMP están entre los mejores predictores de BBTMs capaces de procesar proteomas completos (Hu and Yan, 2008) y que la versión final del programa tiene un comportamiento superior a ambos, aún sin implementar el reconocimiento de patrones en el extremo C-terminal, puede decirse que el Beta-Predictor-Plus se encuentra entre los mejores programas disponibles para la predicción de BBTMs en proteomas completos.

Además el Predictor de BBTMs se puede utilizar como parte de algoritmos de predicción que tengan como objetivo la identificación de proteínas bacterianas expuestas en la superficie de estos microorganismos y, consecuentemente, de potenciales antígenos para estudios encaminados al desarrollo de vacunas.

### **2.3.2 Determinación del valor de corte óptimo del programa SPAAN para micoplasmas**

Para definir el valor de corte óptimo a emplear para micoplasmas, este se evaluó entre 0,5, similar al 0,51 que mejor permitió segregar las citadhesinas de las no-citadhesinas, y 0,7, tal y como proponen los autores del programa SPAAN para disminuir la detección de falsos positivos (Sachdeva et al., 2005). Desafortunadamente, cuando se obtenían los mejores valores de sensibilidad, con un valor de corte de 0,5, el MCC y la exactitud caían a niveles muy bajos. Por el contrario, para un valor de corte de 0,7, se obtenía una exactitud de 1 y el mejor de todos los MCC, pero la sensibilidad caía a 0,475. Teniendo en cuenta que en este caso, además del MCC y la exactitud mencionados, el VPP era también el mayor y la especificidad alcanzaba 1 (Tabla 3), se decidió seleccionar 0,7 como valor de corte en nuestro sistema. Esta decisión era, además, consistente con la intención de priorizar que las

proteínas positivas identificadas por nuestro sistema, lo fueran en realidad, sobre la posibilidad de perder algunas en el proceso (especificidad sobre sensibilidad), ya que esto incrementa el nivel de confianza de los resultados, al disminuir el número de falsos positivos llevados al laboratorio.

**Tabla 3.** Comportamiento de los parámetros de desempeño en función del valor de corte empleado en la evaluación del programa SPAAN. Alto nivel de exactitud y el mejor MCC, solo se logra con un valor de corte de 0,7, pero para este valor la sensibilidad es solo de 0,45. No obstante, como para 0,7 se obtiene el mayor valor de VPP y la especificidad es de 1, se seleccionó este valor de corte para elevar el nivel de seguridad de los resultados.

	VP	FN	VN	FP	Sensibilidad	Especificidad	MCC	VPP	Exactitud
<b>SPAAN</b> <b>Valor de Corte</b> <b>(Pad)-0,5</b>	33	7	985	85	0,825	0,921	0,451	0,917	0,280
<b>SPAAN</b> <b>Valor de Corte</b> <b>(Pad)-0,55</b>	32	8	1023	47	0,800	0,956	0,548	0,950	0,405
<b>SPAAN</b> <b>Valor de Corte</b> <b>(Pad)-0,6</b>	28	12	1048	22	0,700	0,979	0,611	0,969	0,560
<b>SPAAN</b> <b>Valor de Corte</b> <b>(Pad)-0,65</b>	23	17	1059	11	0,575	0,990	0,611	0,975	0,676
<b>SPAAN</b> <b>Valor de Corte</b> <b>(Pad)-0,7</b>	19	21	1070	0	0,475	1	0,683	0,981	1

Donde Pad - Probabilidad de una proteína de ser adhesina

Estos resultados difieren ligeramente de los obtenidos por los autores del SPAAN (Sachdeva et al., 2005), quienes para un valor de corte de 0,51 alcanzaron una sensibilidad de 0,89, una especificidad de 1 y un MCC de 0,94 (el mayor para todo el espectro de valores de corte). Estas diferencias pueden deberse a que ellos emplean unos conjunto de entrenamiento y evaluación muy reducidos, y que, en ningún caso, contenían adhesinas o proteínas de micoplasmas. Los blancos de las adhesinas pueden ser tan variados como hospederos hay (Chaudhuri and Ramachandran, 2014), e incluso, dentro de cada hospedero las células blanco de la adherencia pueden ser diferentes (Chaudhuri et al., 2011). Esto hace suponer que las características de las adhesinas, pueden ser también de naturaleza diversa. Así las citadhesinas de micoplasmas pueden tener características particulares que no estaban contenidas ni en el conjunto de entrenamiento ni en el de evaluación del SPAAN. Esto puede explicar las diferencias encontradas.

### 2.3.3 Validación de Bact-VAP

#### 2.3.3.1 Validación de capacidad de predicción de candidatos vacunales protectores

Con el empleo de Bact-VAP se predijeron 15 de 18 conocidos PVCs de *S. pneumoniae* mientras que el programa NERVE y los servidores Vaxign, VaxiJen y Jenner-Predict, predijeron solo 7; 6; 3 y 10 PVCs, respectivamente. En *E. coli*, aunque las diferencias no fueron tan marcadas, los resultados fueron similares. Con Bact-VAP se identificaron 25 PVC de un total de 28 proteínas mientras que con NERVE, Vaxign, VaxiJen y Jenner-Predict, solo se detectaron 21; 18; 21 y 23 respectivamente. Bact-VAP mostró los mejores resultados con los conjuntos de datos de *S. pneumoniae* y *E. coli*. En el caso de la evaluación con las proteínas provenientes de la base de datos Protegen, de 177 proteínas, con Bact-VAP se identificaron 135, este resultado fue igualmente superior al obtenido con NERVE, Vaxign, VaxiJen (121; 89 y 97 respectivamente), pero ligeramente inferior a las 137 detectadas con Jenner-Predict. Por último, al llevar a cabo el análisis con las secuencias empleadas en el desarrollo de VaxiJen, de nuevo Bact-VAP fue capaz de detectar más PVCs que ninguno de los otros predictores 68 de un total de 83, por 53; 41; 46 y 59 de NERVE, y los servidores, Vaxign, VaxiJen y Jenner-Predict respectivamente. Sin embargo, al analizar el conjunto de datos negativos, Bact-VAP mostró un nivel de falsos positivos superior a los otros sistemas, 12 de 33, por 8; 5; 3 y 2 de Vaxign, VaxiJen y Jenner-Predict respectivamente. Todos estos resultados se resumen en la Tabla 4.

**Tabla 4.** Resultados de la predicción obtenidos con Bact-VAP y los programas similares existentes: NERVE, Vaxign, Vaxijen y Jenner, con el empleo de bases de datos de PVCs identificados experimentalmente.

	<i>S. pneumoniae</i>		<i>E. coli</i>		Protegen		VaxiJe			
	VP	FN	VP	FN	VP	FN	VP	FN	VN	FP
NERVE	7	11	21	7	121	56	53	30	25	8
Vaxign	6	12	18	10	89	88	41	42	28	5
VaxiJen	3	15	21	7	97	80	46	27	30	3
Jenner-Predict	10	8	23	5	137	40	59	24	31	2
Bact-VAP	15	3	25	3	135	42	68	15	21	12

Donde VP-verdaderos positivos predichos; FN-falsos negativos predichos; VN-verdaderos negativos predichos; FP-falsos positivos predichos

Los parámetros de desempeño de los diferentes métodos de predicción de PVC publicados, se presentan en la Tabla 5. Los niveles de sensibilidad de Bact-VAP son superiores al de

resto de los sistemas evaluados, con independencia del conjunto de datos empleado. La falta de disponibilidad de un subconjunto negativo de PVCs en los proteomas de *S. pneumoniae* y *E. coli* (Jaiswal et al., 2013), así como de en la base de datos de Protegen (Yang et al., 2011), imposibilitó el cálculo del resto de los parámetros de desempeño en tres estos casos. Cuando se empleó la base de datos utilizada para VaxiJen (Doytchinova and Flower, 2007), Bact-VAP también mostró el mejor valor de sensibilidad, pero el nivel de falsos positivos fue más alto que el del resto de los programas, lo que provocó que tanto la especificidad como la exactitud del sistema se ubicaran últimas con respecto al resto de las herramientas empleadas en el estudio. A pesar de esto, sus MCC y VPP fueron superiores a los de NERVE, Vaxign y VaxiJen.

**Tabla 5.** Evaluación del desempeño de Bact-VAP contra los programas similares existentes: NERVE, Vaxign, Vaxijen y Jenner

	CONJUNTOS DE DATOS DE EVALUACIÓN							
	<i>S. pneumoniae</i> Sensibilidad	<i>E. coli</i> Sensibilidad	Protegen Sensibilidad	VaxiJen				
				Sensibilidad	Especificidad	MCC	VPP	Exactitud
<b>NERVE</b>	0,39	0,75	0,68	0,64	0,76	0,36	0,67	0,87
<b>Vaxign</b>	0,33	0,64	0,5	0,49	0,85	0,32	0,6	0,89
<b>VaxiJen</b>	0,17	0,75	0,55	0,55	0,91	0,42	0,66	0,94
<b>Jenner-P</b>	0,56	0,82	0,77	0,71	0,94	0,59	0,78	0,97
<b>Bact-VAP</b>	0,83	0,9	0,76	0,82	0,64	0,44	0,77	0,85

Donde MCC-coeficiente de correlación de Matthews y VPP valor predictivo positivo.

El objetivo final del desarrollo de este tipo de herramientas es la de proveer un servicio que permita identificar candidatos vacunales confiables, e información relevante acerca de su potencial, de manera tal que el desarrollo de vacunas por subunidades se pueda llevar a cabo de manera acelerada. Esto, ciertamente, pasa por generar salidas de proteínas positivas, lo más reducidas posible, es decir, alcanzar altos niveles de especificidad en los resultados (Altindis et al., 2015). En este sentido Bact-VAP aparentemente presenta algunas limitaciones, ya que con este conjunto de datos negativos su especificidad fue inferior a la de los otros sistemas evaluados en el estudio. Esto puede estar motivado por las insuficiencias que aún presenta el algoritmo empleado. Por ejemplo, aún no se incluye en el análisis la comparación con proteínas del hospedero para evitar potenciales fenómenos de autoinmunidad, ni tampoco se hacen estudios de comparación entre diferentes cepas para buscar proteínas conservadas, como está implementado en el servidor Jenner-Predict

(Jaiswal et al., 2013). Ambos análisis podrían reducir el número de falsos positivos del sistema.

No obstante existen otros aspectos que deben tenerse en cuenta en este análisis. Por un lado, el conjunto de datos negativos empleado en Vaxijen es muy pequeño; esto puede disminuir estadísticamente, la fiabilidad del resultado. Por otra parte, a diferencia de los antígenos protectores, las evidencias experimentales de la incapacidad de algunas moléculas de inducir protección, pueden llevar a conclusiones erróneas. En muchos casos los modelos animales empleados para evaluar los antígenos protectores están sesgados por el hecho de que los animales son inoculados con grandes cantidades de bacterias y solo los antígenos que inducen anticuerpos que bloquean la bacteriemia resultan protectores. Otros antígenos que pueden tener un papel importante en la inducción de respuesta protectora se pierden en el proceso (Altindis et al., 2015). Es necesario constantemente tener en cuenta este aspecto antes de excluir de la lista de antígenos protectores a alguna molécula. Todos estos factores refuerzan la necesidad de contar con bases de datos para la evaluación de la capacidad protectora de las proteínas más amplias y representativas. Por estas razones los resultados relacionados con la especificidad de Bact-VAP, deberían ser tomados con cautela, aunque sin dejar de prestarles atención y trabajar en las mencionadas limitaciones del sistema.

Puede resumirse que el algoritmo de predicción empleado, aún con limitaciones, permite predecir PVCs bacterianos, de modo confiable y seguro. A todo esto es importante agregar que, actualmente, ninguno de los predictores mencionados, se encuentra disponible o en estado funcional, por lo que el Bact-VAP se convierte en el único habilitado para este tipo de predicción.

#### **2.3.3.2 Validación de capacidad de predicción de proteínas de superficie de micoplasmas**

El último objetivo de este capítulo es aplicar esta herramienta al estudio de los micoplasmas y debido a las características únicas de estos, lo ideal era validarlo con conjuntos de

evaluación compuestos con proteínas provenientes de estos microorganismos. Al no poseer un conjunto de datos de PVCs de micoplasmas, unido al hecho de que los predictores mencionados no están funcionales, para determinar cuan efectivo era el sistema para identificar PVCs de micoplasmas, se limitó la evaluación a validar la capacidad del sistema de identificar proteínas de superficie de este género bacteriano. Esta decisión se basó en que este grupo de moléculas son claves en la patogenicidad de estos microorganismos (Browning et al., 2011, Hoelzle et al., 2014, Ni et al., 2015, Hegde et al., 2015, Burki et al., 2015) y, en general, median la interacción de los microorganismos con el sistema inmune (Walters and Mobley, 2010, Almeida et al., 2011, Zhang et al., 2013, Singh et al., 2014, Zhang et al., 2014, Carlos et al., 2015, Iverson-Cabral et al., 2015, Vishnu et al., 2015).

Los resultados de este análisis revelan que el número de falsos negativos, 52, que se obtuvo con el empleo de Bact-VAP, fue inferior al que se obtuvo con el resto de los programas empleados, que llegaron a ser en el caso del servidor LocTree3, de 206. Así mismo, los falsos positivos generados por Bact-VAP fueron solo 24, de un total de 1071 proteínas, ligeramente superior solamente, a los 20 que generó el servidor LocTree3 (Tabla 6).

**Tabla 6.** Resultados de la evaluación del desempeño de Bact-VAP, para la predicción de proteínas de superficie. Comparación con el rendimiento de los sistemas de predicción de localización celular PsortB, CELLO

	VP	FN	VN	FP	Sensibilidad	Especificidad	MCC	VPP	Exactitud
<b>PsortB</b>	173	101	1040	31	0,631	0,971	0,676	0,902	0,848
<b>CELLO</b>	117	157	1017	54	0,427	0,950	0,455	0,843	0,684
<b>LocTree3</b>	68	206	1051	20	0,248	0,981	0,374	0,832	0,772
<b>Bact-VAP</b>	222	52	1047	24	0,810	0,978	0,821	0,943	0,902

Donde VP-verdaderos positivos predichos; FN-falsos negativos predichos; VN-verdaderos negativos predichos; FP-falsos positivos predichos; MCC-coeficiente de correlación de Matthews y VPP valor predictivo positivo.

La carencia de un conjunto de PVCs de micoplasmas obligó a evaluar la capacidad del programa de procesar estos microorganismos, solo en lo referido a la localización en la superficie de sus proteínas. En este sentido el desempeño de Bact-VAP es superior al de los programas con los que se comparó. Demostró una sensibilidad casi cuatro veces mayor a LocTree3, prácticamente el doble que CELLO y superior al PsortB (0,81 por 0,631). En el apartado de especificidad todos los sistemas tuvieron un desempeño similar. Bact-VAP

ocupó el segundo lugar por detrás de LocTree3, por un margen muy estrecho, con un excelente valor de 0,978 por 0,981 del LocTree3. Y su MCC fue superior al del resto de los programas. El VPP, por su parte también mostró un mejor resultado para el sistema en estudio, con un valor de 0,943 seguido del resultado del PsortB con 0,902. El resto estuvo por debajo de 0,90. Por último el Bact-VAP fue el único de los programas empleados en la comparación, con un valor de exactitud por encima de 0,90. Es importante destacar el alto nivel de especificidad que se alcanza. Como ya se ha señalado, lo ideal es que el número de blancos propuestos a partir del análisis computacional, para el trabajo en el laboratorio, sea el más pequeño posible, para garantizar que la mayor cantidad de proteínas propuestas sea positiva. En resumen, en la predicción de proteínas de superficie, el desempeño de Bact-VAP es superior al del resto de los predictores de localización celular en cada uno de los parámetros a medir, a excepción de la especificidad, que es solo inferior a la de LocTree3.

Una posible causa de las diferencias entre los resultados obtenidos para Bact-VAP y los obtenidos con tres de los principales predictores de localización sub-celular bacteriana existentes (Goldberg et al., 2014), podría estar en las diferencias entre el algoritmo de predicción que se emplea en Bact-VAP, con respecto a los empleados por estos sistemas. Como ha sido bien documentado, tanto los programas empleados, como el orden de los mismos, tienen un importante efecto en el rendimiento del algoritmo final (Zhou et al., 2008). En su algoritmo de predicción, PsortB combina varios módulos: I.- SCL-BLAST para predicción de localización sub-celular por análisis de homología, II.- PROSITE para la detección de motivos específicos que correlacionan con localizaciones sub-celulares específicas, III.- S-TMHMM para la identificación de proteínas hélice- $\alpha$ , IV- un predictor de péptido señal que combina un algoritmo de HMM y otro de MSV y V- un algoritmo de MSV para clasificar las proteínas en citoplasmáticas o no citoplasmáticas (Yu et al., 2010). Mientras, CELLO emplea un sistema de clasificación formado por dos niveles de MSV (Yu et al., 2006) y LocTree3 combina el empleo de PSI-Blast y un sistema jerárquico de MSV (Goldberg et al., 2014). Aparentemente, nuestra elección, tanto de programas, como de la

distribución de los mismos en el algoritmo de predicción, fue la correcta. Otra posibilidad es que en los conjuntos de datos empleados para el entrenamiento de estos sistemas no se emplearon proteínas de micoplasmas (Yu et al., 2006, Yu et al., 2010, Goldberg et al., 2014). Previamente CELLO y Psortb han mostrado pobres desempeños, que se han vinculado a carencias de los conjuntos de datos de entrenamiento (Zhou et al., 2008).

No obstante, a pesar de que el sistema ya muestra un buen desempeño, el mismo posee algunas limitaciones que deben ser atendidas en el futuro mediante la incorporación de nuevos componentes, como: I- determinación de la similitud de secuencias entre las proteínas predichas y el proteoma del hospedero; II- conservación entre diferentes genomas de diferentes cepas de la especie de interés; III- exclusión de secuencias presentes en las cepas no patógenas; IV- predicción de epítomos de MHC de clase I y clase II; V- identificación de características estructurales/funcionales recurrentes en antígenos protectivos bacterianos conocidos, el llamado "espacio Protectómico", y uso de este tipo de "firmas de protección" para el descubrimiento de antígenos protectivos.

En resumen, nuestros resultados evidencian que Bact-VAP es capaz de predecir blancos vacunales bacterianos y antígenos de superficie de micoplasmas. Este sistema proporciona un punto de partida para la predicción de candidatos vacunales a partir del conocimiento del genoma de la bacteria y algunos indicios acerca de la posible actividad biológica o la estructura de dichos candidatos.

#### **2.3.4 Identificación computacional con el empleo de Bact-VAP, de las proteínas de superficie antigénicas de *M. gallisepticum***

Además del papel crucial de adhesinas de los micoplasmas en los mecanismos de virulencia de estos microorganismos, se conoce que sus proteínas de superficie son altamente inmunogénicas y generan una fuerte respuesta de anticuerpos. La identificación de las proteínas de la superficie de *M. gallisepticum* es, por tanto, crucial para poder desarrollar métodos serológicos de diagnóstico y para el diseño de nuevas vacunas. Como parte de los



esfuerzos para identificar nuevas proteínas de superficie de *M. gallisepticum*, potencialmente implicadas en los mecanismos de patogenicidad de este microorganismo se empleó un nuevo procedimiento bioinformático, basado en un enfoque de VR.

De un total de 763 proteínas, se identificaron 291 como de superficie (Anexo 3), para un 38,14%, proporción superior a la media de 30% predicho por las metodologías clásicas de VR (Altindis et al., 2015). Este aumento puede explicarse porque para microorganismos de tan reducido genoma, la reducida capacidad de codificación de estos microorganismos, no es a expensas de los genes de proteínas de superficie (Browning et al., 2011, Citti and Blanchard, 2013), sino de la eliminación de otros genes no esenciales, ya que la interacción con el entorno garantiza los metabolitos que no son capaces de generar, interacción que está mediada por las proteínas de membrana externa (Browning et al., 2011). Y así, la proporción de los genes que codifican para estas últimas con respecto al total de regiones codificantes del genoma de *M. gallisepticum*, es superior a la que presentan otros microorganismos.

Las 291 proteínas se subdividieron como sigue, 175 (23%) hélices- $\alpha$  transmembrana, 16 (2%) proteínas del tipo barril- $\beta$ , 52 (7%) lipoproteínas, otras 36 (5%) podrían estar ancladas por los extremos N o C-terminal, y 80 (10%) citadhesinas (la mayor parte de estas ya incluidas en los subgrupos mencionados).

En los últimos años las lipoproteínas de membrana han atraído mucha atención, debido a que su abundancia en la superficie de los micoplasmas es significativa, en contraste con el número limitado de lipoproteínas en las membranas de otras eubacterias (Razin et al., 1998) y por el papel que desempeñan en los mecanismos patogénicos de los micoplasmas (Bencina, 2002, Browning et al., 2011, Wu et al., 2012, Reolon et al., 2014). A diferencia de algunas lipoproteínas y citadhesinas de este microorganismo, nunca se han identificado y/o caracterizado, ni las proteínas hélice- $\alpha$  ni las de tipo barril- $\beta$  de *M. gallisepticum*. En este trabajo se predicen, por primera vez, proteínas de estas categorías para esta bacteria. Si se tiene en cuenta la relación que existe entre la estructura y la función de las proteínas (Tiwari

and Srivastava, 2014), este resultado sienta las bases para una futura caracterización de la función de algunas de estas moléculas.

### **Conclusiones parciales**

1. Bact-VAP es efectivo tanto en la predicción de PVCs bacterianos, como de proteínas de superficie de micoplasmas.
2. El proteoma de *M gallisepticum* es rico en antígenos de superficie, potenciales PVCs, un alto número de los cuales no se han caracterizado hasta el momento.

### 3 CAPÍTULO 3. OBTENCIÓN Y CARACTERIZACIÓN TOPOLÓGICA Y ANTIGÉNICA DE LAS PROTEÍNAS OBTENIDAS

#### 3.1 Introducción

Todas las predicciones computacionales deben ser experimentalmente verificadas antes de arribar a conclusiones (Mora *et al.*, 2003). La obtención de proteínas vía ADN recombinante es una de las alternativas utilizadas con el objetivo de estudiar su función en los mecanismos de patogenicidad y evasión de la respuesta inmune por parte de los microorganismos (Stryjewska *et al.*, 2013).

Objetivo:

1. Caracterizar las proteínas obtenidas en cuanto a su localización sub-celular y su antigenicidad.

Tareas específicas:

1. Amplificar por PCR los genes seleccionados *in silico*.
2. Obtener las proteínas recombinantes purificadas.
3. Determinar si los productos obtenidos son antigénicos y evaluar la posible localización en la superficie bacteriana de las proteínas en estudio

#### 3.2 Materiales y Métodos

##### 3.2.1 Vectores, cepas bacterianas y condiciones de cultivo de *M. gallisepticum*

La cepa DH5 $\alpha$  (F<sup>-</sup> *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20*  $\phi$ 80*dlacZ* $\Delta$ M15  $\Delta$ (*lacZYA-argF*)U169, *hsdR17*(*r<sub>K</sub><sup>-</sup>m<sub>K</sub><sup>+</sup>*),  $\lambda^-$ ), se obtuvo del banco de cepas de *E. coli* del Centro Nacional de Sanidad Agropecuaria (CENSA) y la cepa C41pRospLysS fue gentilmente donada por el Laboratorio Max F. Perutz, Universidad de Viena. La cepa

ISM612 de *E. coli* fue suministrada por el Departamento de Microbiología Veterinaria y Medicina Preventiva, Colegio de Medicina Veterinaria, Universidad Estatal de Iowa, Ames, EUA. La cepa R(low) de *M. gallisepticum* fue donada por el Instituto Friedrich-Loeffler, Centro Federal de Investigaciones para la Salud Animal, Jena, Alemania. El vector empleado para los estudios de clonaje molecular y la posterior expresión de los genes clonados fue el vector de fusión con poli-His, pQE30 (Qiagen) (Anexo 4). La cepa R(low) de *M. gallisepticum* se cultivó en medio Hayflick caldo (Anexo 5). El cultivo se incubó estáticamente a 37°C, en presencia de CO<sub>2</sub> al 5% hasta que alcanzó una coloración amarilla. Posteriormente se subcultivó en medio Hayflick agar (Anexo 5) donde se incubó durante siete días a 37°C, en presencia de CO<sub>2</sub> al 5%. Se seleccionó una colonia aislada y se sembró nuevamente en medio Hayflick caldo. El cultivo se incubó estáticamente a 37°C, en presencia de CO<sub>2</sub> al 5% hasta que alcanzó una coloración amarilla.

### 3.2.2 Diseño de los cebadores para el clonaje posterior de los genes

Los cebadores se diseñaron con el empleo del programa Vector NTI Advance v 11.0 (Invitrogen). Como blanco del diseño se empleó la secuencia genómica de la cepa R(low) de *M. gallisepticum*, que se descargó de la base de datos del NCBI (Nº Acceso: AE015450.2, GI: 284811830). Se diseñaron cebadores (sentido y antisentido) para los genes codificantes para las primeras 91 proteínas de la lista que no están caracterizadas (Anexo 3, señaladas en rojo), así como para algunas de las proteínas de membrana y factores de virulencia de *M. gallisepticum*, más conocidos y caracterizados [GapA (Papazisi *et al.*, 2002), CrmA (Papazisi *et al.*, 2002), CrmC (Papazisi *et al.*, 2003), VlhA.5.02 (Markham *et al.*, 1994), Mgc2 (Hnatow *et al.*, 1998) y OppA (Papazisi *et al.*, 2002)] (Anexo 3).

Para definir que enzimas podían ser empleadas en los clonajes, se realizó un estudio con el objetivo de determinar cuáles de los sitios de restricción presentes en el sitio de clonaje múltiple del plásmido pQE30 (Anexo 4) estaban ausentes en las secuencias a clonar (Anexo 3). Se seleccionaron *Sma* I y *Hind* III para el clonaje de *vlhA.5.02*, *Sac* I y *Kpn* I para el gen

*oppA* (MGA\_0237) y *Sph* I y *Sma* I para los restantes 54 genes. Los sitios de restricción correspondientes a estas enzimas se incorporaron a los cebadores sentido y antisentido (según resultados del estudio).

En el diseño también se tuvo en cuenta eliminar del producto de amplificación, la región correspondiente a los péptidos-señal identificados con el empleo del SignalP y el LipoP. Debido a la hidrofobicidad de las regiones transmembrana, estas son, en general, difíciles de clonar y purificar (Pizza *et al.*, 2000), además, en la mayoría de los casos son removidos de las proteínas maduras (Schlegel *et al.*, 2014, Zuckert, 2014, Costa *et al.*, 2015).

Por último, los cebadores se diseñaron de manera tal que todos los productos de amplificación quedaran en el marco de lectura correcto en el plásmido receptor. Esto se garantizó al realizar *in silico* el clonaje de cada uno de los productos de PCR (también generados *in silico*) y evaluar si se generaban las regiones codificantes esperadas. En el caso particular de *oppA* y *vlhA.5.02*, el diseño se realizó de manera tal que el fragmento resultante no incluyera ni potenciales regiones transmembrana ni codones opal. Todos los análisis de PCR, Clonaje y Expresión *in silico*, se realizaron con el empleo del Vector NTI Advance v 11.0.

### **3.2.3 Extracción de ADN genómico de la cepa R(low) de *M. gallisepticum***

El ADN genómico se purificó según el protocolo empleado por Kang y colaboradores (Kang *et al.*, 2012), con algunas modificaciones. Se centrifugaron por 30 min a 10 000xg, 10mL de un cultivo de la cepa R(low) de *M. gallisepticum*, La masa celular se resuspendió en 5mL de tampón de lisis (Anexo 5). La suspensión resultante se incubó 10 min a 50°C, seguidos de 10 min a 37°C. Se adicionaron 50µL de proteinasa K (100mg/mL) (Anexo 5) y se incubó por 40 min a 37°C. A esta mezcla se le añadieron 10µL de KAc 5M (Anexo 5) y se centrifugó 15 min a 14 000xg. El sobrenadante se transfirió a un nuevo tubo y se adicionaron 3550µL de isopropanol, y la mezcla se centrifugó a 14 000xg por 30 min. El precipitado se lavó con etanol al 70% y se centrifugó 15 min a 14 000xg. El precipitado se dejó secar media hora a

temperatura ambiente y se resuspendió de manera pasiva en 1mL de H<sub>2</sub>O<sub>dd</sub> (Agua bidestilada), dejándolo por 60 min a 4°C.

Los resultados de la purificación se monitorearon por electroforesis en gel de agarosa al 1% (peso/volumen, p/v) (Anexo 5). La misma se llevó a cabo a 100 V (voltaje constante). Como agente de tinción se empleó una solución de bromuro de etidio (Sigma) (Anexo 5). La cuantificación del ADN purificado se realizó en un equipo GeneQuant II (Pharmacia-Biotech).

#### **3.2.4 Amplificación por PCR de los genes seleccionados**

Los genes o fragmentos de genes a clonar, se obtuvieron por amplificación mediante la PCR (Innis *et al.*, 2012), con el empleo de ADN cromosomal de la cepa R(low) de *M. gallisepticum*, como molde de la reacción. En todas las amplificaciones se empleó la polimerasa de alta fidelidad de copia (BioLabs). Todas las condiciones de la PCR fueron ajustadas para lograr la amplificación de todos los blancos con el empleo de un único programa, por lo que se seleccionó una temperatura de unión inferior a la menor temperatura fusión de los cebadores y un tiempo de elongación suficiente para amplificar de hasta 3 000 bases, superior a la talla de los mayores fragmentos esperados. Las secuencias de los fragmentos a amplificar se detallan en el anexo 6.

Para la amplificación de los genes por PCR se adicionó 1µL del ADN cromosomal de la cepa R(low) (200ng/µL) a la mezcla que contenía 5µL de tampón de la ADN polimerasa 10x (BioLabs), dNTPs (BioLabs) 200µM (concentración final), ambos cebadores 400nM (concentración final) y 0,2U de la Vent ADN polimerasa (BioLabs) en un volumen total de 50µL. El programa de PCR fue el siguiente: 1 ciclo de 94°C por 5 min, seguido por 30 ciclos de 94°C por 1 min, 45°C por 1 min y 72°C por 3 min y por último un ciclo de 72°C por 7 min.

Los resultados de la PCR se monitorearon por electroforesis en gel de agarosa, como se describe en el epígrafe 3.2.3. En todos los casos se aplicaron 3µL de patrón de peso molecular (1kb leader, Fermentans).

### 3.2.5 Generación de las construcciones recombinantes

#### 3.2.5.1 Preparación de los productos de PCR

Se digirieron 4µg de los productos de PCR *vlhA.5.02*, *oppA* y los correspondientes al resto de los genes, con 10U de *Sma* I (10U/µL), *Sac* I (10U/µL) y *Sph* I (10U/µL) respectivamente, en un volumen total de 20µL, en presencia de 2µL de tampón Tango 10x, a 37°C durante 2h. Pasado ese tiempo se añadieron 10U de *Hind* III, *Kpn* I y *Sma* I (10U/µL), respectivamente y la mezcla se incubó por 2 horas más a 30°C (Todas las enzimas y el tampón Tango eran de la compañía Fermentans). Las enzimas se inactivaron incubando la mezcla 20 min a 65°C.

Los productos de PCR, antes y después de la digestión se purificaron con el empleo del juego comercial “QIAquick PCR Purification Kit” (Qiagen), siguiendo el protocolo del fabricante. La cuantificación de los productos de PCR purificados se realizó en un equipo GeneQuant II (PharmaciaBiotech).

Los resultados se monitorearon por electroforesis en gel de agarosa, como se describe en el epígrafe 3.2.4. En todos los casos se aplicaron 3µL de patrón de peso molecular (1kb leader, Fermentans).

#### 3.2.5.2 Preparación del vector

Se inocularon 5mL de medio Luria Bertani (LB) (Anexo 5), suplementado con ampicilina 100µg/mL, con una colonia de la cepa DH5α transformada con el plásmido pQE30. El cultivo se incubó en agitación toda la noche a 37°C. Con 4mL de este cultivo se inocularon 200mL de medio LB (100µg/mL de ampicilina). Este cultivo también se incubó en agitación toda la noche a 37°C.

El vector pQE30 se purificó con el juego comercial Plasmid Midi Kit (QIAGEN), según el protocolo descrito por el fabricante. El plásmido resultante obtenido de los 200mL de cultivo, se resuspendió en un volumen final de 300µL de H<sub>2</sub>O<sub>dd</sub>.

Para el clonaje de *vlhA.5.02* se digirieron 5µg del vector. La digestión se llevó a cabo con 10U de *Sma* I (10U/µL), en un volumen total de 20µL, en presencia de 2µL de tampón Tango 10x, a 37°C durante 2h. Pasado ese tiempo se añadieron 10U de *Hind* III (10U/µL) y la mezcla se incubó por 2 horas más a 37°C. Para el clonaje de *oppA* se digirieron 5µg del vector. La digestión se realizó con 10U de *Sac* I (10U/µL), en un volumen total de 20µL, en presencia de 2µL de tampón Tango 10x, a 37°C durante 2h. Pasado ese tiempo se añadieron 10U de *Kpn* I (10U/µL) y la mezcla se incubó por 2 horas más a 37°C. Para el clonaje del resto de los genes se digirieron 100µg del vector. La digestión se realizó con 100U de *Sma* I (10U/µL), en un volumen total de 300µL, en presencia de 30µL de tampón Tango 10x, a 37°C durante 2h. Pasado ese tiempo se añadieron 100U de *Sma* I (10U/µL) y la mezcla se incubó por 2 horas más a 37°C. Todas las enzimas así como el tampón Tango eran de la compañía Fermentans. Las enzimas se inactivaron incubando la mezcla 20 min a 65°C.

Los productos de digestión se purificaron con el empleo del juego comercial QIAquick Gel Extraction Kit (QIAGEN) según protocolo del fabricante.

Los resultados se monitorearon por electroforesis en gel de agarosa, como se describe en el epígrafe 3.2.4. En todos los casos se aplicaron 3µL de patrón de peso molecular (1kb leader, Fermentans). La cuantificación del ADN plasmídico se realizó en un equipo GeneQuant II (Pharmacia-Biotech) luego de verificar por electroforesis la ausencia total de ARN en las muestras.

### **3.2.5.3 Ligazón de los genes con el vector**

Para obtener las construcciones recombinantes, los productos de PCR digeridos se ligaron (Scharf, 1990), con 50 ng del vector pQE30 digerido, en una relación de 3 moles de inserto por cada mol de vector. Se emplearon en cada caso 400U de T4 ADN ligasa (400U/µL) (BioLabs) y 2µL de tampón T4 ADN ligasa 10x (BioLabs), en un volumen total de 20µL y la



ligazón se realizó a 22°C durante toda la noche. La cantidad de inserto en cada reacción se calculó según la fórmula:

$$Mi = Mv \times \frac{Ti}{Tv} \times RM$$

Donde Mi es la masa del inserto, Mv la masa del vector, Ti la talla del inserto, Tv la talla del vector y RM la relación molar entre ambos

#### 3.2.5.4 Transformación química de DH5α

Células quimio-competentes de la cepa DH5α se transformaron (Froger and Hall, 2007) con el producto de la ligazón de los genes con el vector pQE30. Se empleó la cepa DH5α debido a que ésta facilita la propagación y posterior purificación de los plásmidos.

Las células quimio-competentes se prepararon según el siguiente protocolo. Incubar en agitación a 37°C toda la noche 10mL de un cultivo de DH5α en LB. Inocular 1L de LB con los 10mL de cultivo e incubar en agitación a 37°C. Cuando el cultivo alcanzó una DO<sub>600</sub> entre 0,35 y 0,4. Se colocó en hielo por 30 minutos. Se centrifugó a 3 000xg durante 15 minutos a 4°C. El sedimento se resuspendió en 400mL de MgCl<sub>2</sub> 100mM a 4°C. Se centrifugó de nuevo a 2 000xg por 15 minutos a 4°C. El sedimento se resuspendió en 200mL de MgCl<sub>2</sub> 100mM a 4°C. Se dejó reposar por 15 min en hielo. Se volvió a centrifugar a 2 000xg durante 15 minutos a 4°C. El sedimento se resuspendió en 200mL de MgCl<sub>2</sub> 100mM a 4°C y se dejó reposar 15 minutos en hielo. Se centrifugó la suspensión a 2 000xg durante 15 minutos. El sedimento se resuspendió en 50mL de MgCl<sub>2</sub> 85mM, glicerol 15% a 4°C. Se centrifugó la suspensión a 1 000xg durante 15 minutos a 4°C. El sedimento se resuspendió en 2mL de MgCl<sub>2</sub> 85mM, glicerol 15% a 4°C. Se prepararon fracciones de 100μL en tubos de microcentrifuga de 1,5mL. Las células competentes se conservaron a -80°C.

Para la transformación se descongelaron las células competentes en hielo por 10-15 min, a 100μL de células competentes se le adicionaron 10μL del producto de la ligazón y se mezclaron suavemente, como control se empleó el plásmido pQE30 (Anexo 4), del cual se adicionaron 20ng también a 100μL de células competentes. Se incubó en hielo durante

20min, para incrementar la eficiencia de la transformación se dio un choque térmico a 42°C durante 45seg, seguido de una incubación en hielo de 5 min. A cada alícuota se le adicionó 1mL de LB y se incubó 1h a 37°C con agitación a 50rpm. En placas de LB suplementadas con ampicillina (100µg/mL) se sembraron 150µL de los cultivos de la cepa transformada con los diferentes productos de ligazón y en otra se sembraron 50µL del cultivo del plásmido control. Las placas se incubaron toda la noche a 37°C.

### 3.2.5.5 Monitoreo de las construcciones recombinantes

En un primer paso de filtrado se seleccionaron los plásmidos con una talla superior a la talla del pQE30 (potenciales clones recombinantes). Para esto primeramente se sembraron las colonias en 3mL de caldo LB suplementado con ampicillina (100µg/mL) y se incubaron toda la noche a 37°C. Se purificaron los plásmidos por el método de lisis alcalina. Se tomaron 1,4mL de los cultivos, se centrifugaron 5 min a 10 000xg y se resuspendió totalmente el sedimento celular en 110µL de tampón de resuspensión (Anexo 5). Se añadieron 100µL de tampón de lisis (Anexo 5) y se mezcló suavemente por inversión 5-6 veces. Posteriormente se adicionaron 200µL de tampón de neutralización y se mezcló suavemente por inversión 5-6 veces, luego se incubó la mezcla por 3 min a temperatura ambiente. Los restos celulares se eliminaron por centrifugación a 12 000xg durante 15 min. El sobrenadante se transfirió a un nuevo tubo y el ADN plasmídico se precipitó con la adición de 250µL de isopropanol, se mezcló por inversión 10 veces y se incubó por 5 min a temperatura ambiente. Por último la mezcla se centrifugó por 5 min a 12 000xg. El precipitado se lavó con 500µL de etanol al 70%. Se centrifugó por 1 min a 12 000xg y se dejó secar por 10 min a temperatura ambiente. El precipitado final se resuspendió en 30µL de H<sub>2</sub>O<sub>dd</sub>.

La presencia de los insertos en los plásmidos de mayor talla que el pQE30 se confirmó por PCR, según el protocolo descrito en el epígrafe 3.2.4. Para los PCR se tomaron 2µL de cada uno de los plásmidos purificados, se empleó en esta ocasión la enzima *Taq*-Pol (Promega) a una concentración de 1U/reacción y se utilizaron los mismos cebadores que para la amplificación inicial de los genes a partir del ADN genómico de *M. gallisepticum*.

Los resultados se monitorearon por electroforesis en gel de agarosa, como se describe en el epígrafe 3.2.4.

### **3.2.6 Obtención de los polipéptidos recombinantes a partir de las construcciones generadas**

Para la expresión de los genes clonados (a excepción de *vlhA.5.02* y *oppA*) se empleó la cepa ISM612 de *E. coli*, que es una cepa mutante para la expresión de genes con codones opal (codificantes para triptófano en micoplasmas) en *E. coli*. En el caso de *vlhA.5.02* y *oppA*, no fue necesario emplear esta cepa, ya que se clonaron fragmentos que no poseían este triplete en su secuencia. En estos casos se utilizó la cepa C41pRospLysS de *E. coli*.

#### **3.2.6.1 Transformación de las cepas de expresión**

Tanto para la preparación de las células competentes como para la transformación de las cepas ISM612 y C41pRospLysS con las construcciones recombinantes, se siguieron los protocolos descritos en el epígrafe 3.2.5.4. Las placas con LB sólido se suplementaron adicionalmente con cloranfenicol (Sigma) (35µg/mL).

La presencia de las construcciones recombinantes en las células transformadas se monitoreó por PCR. Se sembraron las colonias en 1mL de caldo SB o LB (Anexo 5) (según fuera el caso) suplementado con ampicilina (100µg/mL) y cloranfenicol (35µg/mL), y se incubaron en agitación toda la noche a 37°C. Se tomaron 100µL de los cultivos, se centrifugaron 5 min a 1 000xg y se reconstituyeron en 100µL de H<sub>2</sub>O<sub>dd</sub>, la suspensión se calentó 10 min a 95°C y se volvió a centrifugar 5 min a 10 000xg. Según el protocolo descrito en el epígrafe 3.2.4, se le realizó PCR a 5µL de cada uno de los sobrenadantes, se empleó en esta ocasión la enzima *Taq*-Pol (Promega) a una concentración de 1U/µL. Se emplearon los mismos cebadores que para la amplificación inicial de los genes a partir del ADN genómico de *M. gallisepticum*.

Los resultados se monitorearon por electroforesis en gel de agarosa, como se describe en el epígrafe 3.2.4.

### **3.2.6.2 Monitoreo de la expresión y extracción por SDS-PAGE y Western-Blot.**

La electroforesis se realizó bajo condiciones reductoras (Hellman, 2000) y se corrió a corriente constante con 15mA, en geles de acrilamida al 12% (p/v) (Anexo 5). Por último los geles se tiñeron con solución de Azul de Coomasie (Anexo 5) durante 10 minutos en agitación lenta. Posteriormente los geles se destiñeron hirviéndolos en H<sub>2</sub>O<sub>dd</sub> hasta observar las bandas de proteínas definidas.

Para el Western-Blot se hizo una transferencia semiseca (Laurière, 1993) de las proteínas del gel de poliacrilamida a una membrana de nitrocelulosa. La transferencia se realizó a corriente constante y esta se calculó multiplicando el área del gel que se iba a transferir por 2,5. La membrana se bloqueó durante 1 hora en una solución de PBS-Tween 20 0,05%-BSA (sero albúmina bovina) 1% (Anexo 5) y luego se incubó 2 horas a 37°C con un anticuerpo monoclonal contra poli-histidina (Sigma) diluido 1/2 000 en PBS-Tween 20 0,05%-BSA 1%. Seguidamente se incubó 1 hora a 37°C con el conjugado enzimático anti IgG de ratón conjugada con fosfatasa alcalina (Sigma) diluido 1/5 000 en PBS-Tween 20 0,05%-BSA 1%, y por último la reacción se reveló con NBT/BCIP (nitroazul de tetrazolio/5-Bromo-4-cloro-3-indolil fosfato) (Sigma) en tampón para fosfatasa alcalina (Anexo 5).

### **3.2.6.3 Expresión de los genes recombinantes**

La expresión en la cepa ISM612, se realizó mediante inducción de la misma con isopropil- $\beta$ -D-tiogalactopiranosido (IPTG) (Sigma). Se prepararon inóculos de las construcciones recombinantes en la cepa ISM612, en 10mL de caldo SB (Anexo 5) suplementado con ampicilina (Sigma) (100 $\mu$ g/mL) y cloranfenicol (Sigma) (35 $\mu$ g/mL), estos se incubaron en agitación toda la noche a 37°C. Los inóculos se añadieron a 500mL de caldo SB y se incubaron en agitación a 37°C. Cuando estos alcanzaron una densidad óptica (DO) a 600nm

de 0,4÷0,6, se indujo la expresión con IPTG 1mM. La inducción se prolongó por 4 horas en agitación a 37°C.

Para la expresión de los genes *vlhA.5.02* y *oppA*, se empleó el método de autoinducción con lactosa desarrollado por Studier (Studier, 2005). Una ventaja de usar la cepa C41pRospLysS es la presencia en su genoma del gen que codifica para la lisozima que facilita la lisis celular para el aislamiento de las proteínas expresadas. Se prepararon inóculos que se cultivaron toda la noche a 37°C en medio líquido LB suplementado con ampicilina (100µg/mL) y cloranfenicol (35µg/mL). Al día siguiente, con 100µL de este cultivo se inocularon 100mL (en un erlenmeyer de 1L, se emplearon cinco por cada producto recombinante) de medio de auto-inducción (triptona 1%, extracto de levadura 0,5%, Na<sub>2</sub>HPO<sub>4</sub> 25mM; KH<sub>2</sub>PO<sub>4</sub> 25mM; NH<sub>4</sub>Cl 50mM; Na<sub>2</sub>SO<sub>4</sub> 5mM; glicerol 0,5%; glucosa 0,05%; alfa-lactosa monohidratada 0,2%; MgSO<sub>4</sub> 7H<sub>2</sub>O 2mM; FeCl<sub>3</sub> 0,05mM) suplementado con ampicilina (100µg/mL) y cloranfenicol (35µg/mL), y se cultivó por 16 horas a 37°C con agitación a 250rpm. Se recolectaron las células por centrifugación a 5 000 rpm por 10min y se conservaron a -20°C.

La evaluación del resultado de la inducción de la expresión se realizó por Western-Blot según se describe en el epígrafe 3.2.6.2 Las muestras se centrifugaron 10min a 6 000xg, de 1mL de los cultivos, luego la masa celular se resuspendió en 100µL de H<sub>2</sub>O<sub>dd</sub>, Finalmente de estos se tomaron 10µL a los que se le adicionaron 2µL de tampón de muestra 6x (Promega); todas las muestras se hirvieron durante 10min antes de aplicar en el gel.

### **3.2.7 Purificación de los polipéptidos recombinantes de *M. gallisepticum***

#### **3.2.7.1 Extracción y solubilización de los productos recombinantes**

La masa celular obtenida después de la inducción de cada cultivo de 500mL se resuspendió en 75mL de PBS (Anexo 5). Las células se lisaron en un equipo Vibracell (Sonics & Materials), al someterlas a seis ciclos de sonicación de 10 segundos con intervalos de 10

segundos entre cada ciclo (60 kHz, corriente constante). El producto de la lisis se centrifugó 30 minutos a 14 000xg.

El sobrenadante de cada muestra se desechó y los restos celulares sedimentados correspondientes a la expresión de *oppA* y *vlhA.5.02* se resuspendieron en 75mL de PBS-SDS-0,1% y el resto obtenidos en ISM612, en PBS-SDS-1%. La mezcla se incubó, en agitación moderada, por 1 hora a temperatura ambiente. Posteriormente se centrifugó por 30 min a 14 000xg. El sedimento se desechó y el sobrenadante se incubó por 30 min a 4°C. La mezcla se centrifugó por 30 min a 14 000xg. Los restos celulares sedimentados se resuspendieron en 75mL de PBS-SDS-1%.

La presencia de las proteínas de interés en la fracción soluble se monitoreó por SDS-PAGE y Western-Blot según se describe en el epígrafe 3.2.6.2. Las muestras se prepararon con 10µL (tanto de las fracciones insolubles como de las solubles) a los que se le adicionaron 2µL de tampón de muestra 6x (Promega); todas las muestras se hirvieron durante 10min antes de aplicar en el gel.

### **3.2.7.2 Purificación por IMAC**

Para purificar los productos recombinantes por cromatografía de interacción con quelatos metálicos (IMAC), se cargaron las columnas con 1mL de *Chelating Sepharose Fast Flow* (Pharmacia-Biotech) y la purificación se llevó a cabo según instrucciones del fabricante. Las columnas se equilibraron con 5 volúmenes (5mL) de tampón de equilibrio (Anexo 5). Se aplicaron 20mL de cada uno de los extractos de proteínas por columna. Se lavaron con 10 volúmenes de tampón de lavado (Anexo 5). Las proteínas se eluyeron con 5mL tampón de elución (Anexo 5) (se colectaron fracciones de 1,5mL). Todas las soluciones se pasaron por filtros de 0,45µm. El proceso se desarrolló con un flujo determinado por la gravedad.

La evaluación de los resultados del proceso de purificación se realizó por SDS-PAGE y Western-Blot según se describe a en el epígrafe 3.2.6.2. Las muestras se prepararon

tomando 10µL de cada fracción, a los que se le adicionaron 2µL de tampón de muestra 6x (Promega); todas las muestras se hirvieron durante 10min antes de aplicar en el gel.

### **3.2.8 Obtención de proteínas de membrana externa de la cepa R(low) de *M. gallisepticum***

Para extraer las proteínas de membrana de la cepa R(low) de *M. gallisepticum* se empleó un método de partición con Triton X-114 (Riethman *et al.*, 1987). La solución acuosa de Triton X-114 es homogénea a 0°C pero alrededor de los 20°C se separa en una fase acuosa y otra anfifílica. Esta propiedad puede emplearse para separar las proteínas integrales de membrana de los micoplasmas, de las proteínas hidrofílicas solubles (Riethman *et al.*, 1987). El procedimiento se repitió tres veces, siempre a partir de la fracción del detergente obtenida, con el objetivo de minimizar la presencia de proteínas citoplasmáticas en la fase final del detergente.

#### **3.2.8.1 Generación de anticuerpos específicos contra las proteínas de membrana externa.**

Se emplearon dos conejos machos de la raza Gigante Blanco (Centro Nacional de Producción de Animales de Laboratorio, CENPALAB) (688 y 689). Antes de la primera inoculación se extrajo sangre de ambos conejos, a partir de la cual se obtuvo suero negativo. Posteriormente se inmunizaron según el siguiente esquema: Inmunización con 200µg de la mezcla de proteínas de membrana de *M. gallisepticum* obtenida en el epígrafe anterior, en 500µL con el empleo de adyuvante completo de Freund, por vía subcutánea; a los 14 días se administraron 100µg de la mezcla, en 500µL con adyuvante incompleto de Freund (AIF) (250µL por vía subcutánea y 250µL por vía intradérmica); en el día 28 100µg en 500µL en AIF por vía subcutánea; en el día 56, 50µg en 500µL en AIF subcutáneamente. Se extrajo sangre a los 42 días (T1), 56 días (T2). Por último, los animales fueron desangrados al día 70 (T3) después de la primera inoculación.

La titulación de los sueros se realizó por ELISA. Las placas de ELISA (MaxiSorp, NUNC) se sensibilizaron con 100µL de las proteínas de membrana diluidas en tampón de cubrimiento a una concentración de 1µg/mL (cada muestra por duplicado), durante 3 horas a 37°C. La placas se bloquearon durante 1 hora con 150µL de una solución de PBS-Tween 20 0,05%-BSA 1% a 37°C y luego se incubaron 2 horas a 37°C con 100µL de los sueros policlonales obtenidos en los conejos 688 y 689 a los tiempos T1, T2 y T3, contra las proteínas de membrana de la cepa R(low), con el empleo de los siguientes factores de dilución: 1 000; 2 000; 5 000; 10 000; 20 000 y 50 000, en PBS-Tween 20 0,05%-BSA 1%. Seguidamente se incubaron 1 hora a 37°C con 100µL de un conjugado enzimático anti IgG de conejo, conjugado con peroxidasa (Sigma) diluido 1:5 000 en PBS-Tween 20 0,05%-BSA 1%. Por último la placa se reveló con 100µL de tetra metil bencidina (TMB) (Sigma) en tampón sustrato peroxidasa a una concentración de 0,01% y se detuvo la reacción con 50µL de solución de parada antes de leerla a 450nm. El protocolo para preparar todas las soluciones empleadas en el desarrollo del ELISA se describen en el Anexo 5. El título de los sueros se definió como la última dilución de estos, en la cual la relación entre la densidad óptica del suero positivo y la del negativo (DO+/DO-) era mayor que dos.

#### **3.2.8.1.1 Cuantificación de las proteínas de membrana externa de *M. gallisepticum* y los polipéptidos recombinantes purificados.**

Para cuantificar la cantidad de proteínas totales en las muestras de las proteínas de membrana externa de *M. gallisepticum* y de los polipéptidos recombinantes purificados, se utilizó el método de Bradford adaptado a microplacas. Para esto se preparó una curva patrón de BSA en un rango de concentraciones entre 0,031µg/mL hasta 2mg/mL en PBS - Imidazol 250mM -NaCl 300mM. Se añadieron 10µL de cada punto de la curva, de PBS - Imidazol 250mM - NaCl 300mM (blanco) y de las muestras, en los pozos de la placa (2 réplicas de cada punto de la curva) contentivos de 200µL del reactivo de Bradford (Anexo 5) y se mezclaron con este. Por último se determinó la densidad óptica (DO) de los pozos leyendo la placa a 620nm. EL análisis de los datos se realizó de la siguiente manera.



Siempre se restó el valor del blanco a cada uno de los puntos (tanto de la curva como de las muestras), se realizó un análisis de regresión lineal (concentración vs. DO) con los datos de la curva y se determinó la ecuación de la recta que describía la relación de estos parámetros (el análisis de regresión lineal y el posterior cálculo de la concentración de las muestras se realizó empleando el GraphPad Prism versión 5.00 para Windows, GraphPad Software, San Diego California USA, [www.graphpad.com](http://www.graphpad.com)). Luego se calculó la concentración de proteínas totales solubles en las muestras con el empleo de la fórmula:

$$C = \frac{DO - b}{a}$$

Donde: C – Concentración en mg/mL; DO - densidad óptica a 620nm; b - intercepto de la recta en el eje Y; a - pendiente de la recta

### **3.2.9 Evaluación de la antigenicidad de los polipéptidos recombinantes purificados en condiciones desnaturalizantes**

Primeramente se realizó una electroforesis bajo condiciones reductoras (Hellman, 2000). Se corrió a corriente constante con 15mA, en geles de acrilamida al 12% (p/v) (Anexo 5).

Para el Western-Blot se hizo una transferencia semiseca (Laurière, 1993) de las proteínas del gel de poliacrilamida a una membrana de nitrocelulosa. La transferencia se realizó a corriente constante y esta se calculó multiplicando en el área del gel que se iba a transferir por 2,5. La membrana se bloqueó durante 1 hora en una solución de PBS-Tween 20 0,05%-BSA 1% (Anexo 5) y luego se incubó 2 horas a 37°C con el antisuero policlonal específico a proteínas totales de la cepa R(low) de *M. gallisepticum* (Donado por el Instituto Bacteriología y Micología de la Universidad de Medicina Veterinaria de Viena, Austria), diluido 1/2 000 en PBS-Tween 20 0,05%-BSA 1%. Seguidamente se incubó 1 hora a 37°C con el conjugado enzimático anti IgG de ratón conjugada con fosfatasa alcalina (Sigma) diluido 1/5 000 en PBS-Tween 20 0,05%-BSA 1%, y por último la reacción se reveló con NBT/BCIP (Sigma) en tampón para fosfatasa alcalina (Anexo 5).

### **3.2.10 Evaluación de la antigenicidad de los polipéptidos recombinantes purificados en condiciones nativas**

Este experimento se realizó por ELISA para evaluar si las proteínas recombinantes eran capaces de reaccionar en condiciones nativas, con el antisuero contra la cepa R(low) de *M. gallisepticum*.

Los ensayos se realizaron con las proteínas purificadas, el antisuero empleado fue un suero policlonal obtenido en conejos contra proteínas totales de la cepa R(low) de *M. gallisepticum* (Donado por el Instituto Bacteriología y Micología de la Universidad de Medicina Veterinaria de Viena, Austria) y el conjugado enzimático fue un anticuerpo contra IgG de conejo, conjugado con peroxidasa (Sigma).

El protocolo que se siguió para el ELISA fue similar al descrito en el epígrafe 3.2.8.1. Las únicas diferencias fueron que el anticuerpo empleado fue un suero policlonal obtenido en conejos contra la cepa R(low) de *M. gallisepticum* diluido 1:500 y que se consideraron positivas aquellas muestras cuya DO fue superior a la DO del blanco (constituido por un lisado de *E. coli* sometido al mismo proceso de purificación al que se sometieron los productos recombinantes) más 2 veces la desviación estándar (DE) de las DO de este último.

### **3.2.11 Evaluación preliminar de la localización en la superficie bacteriana de las proteínas en estudio**

Para confirmar la localización en la superficie bacteriana de las proteínas en estudio, se investigó si eran reconocidas por los anticuerpos contra proteínas de membrana de *M. gallisepticum* obtenidos según se describió en el epígrafe 3.2.8.1. La evaluación se realizó por ELISA según el protocolo descrito en el epígrafe anterior, pero con el empleo de los anticuerpos contra proteínas de membrana.

### 3.2.12 Secuenciación de los segmentos clonados

Los insertos de las construcciones pQE30-gapA-17, pQE30-MGA\_0648-26, pQE30-MGA\_0993-1, pQE30-MGA\_0267-8, pQE30-MGA\_0416-3, pQE30-MGA\_0319-5, pQE30-MGA\_0843-2, pQE30-MGA\_0398-13, pQE30-crmA-16, pQE30-crmC-17, pQE30-mgc2-32, pQE30-oppA y pQE30-vlhA.5.02, se secuenciaron por la compañía MacroGen Inc. Se utilizó el método Sanger, y los cebadores de secuenciación de los vectores pQE: 5' CC CGAAAAGTGC CACCTG 3' (sentido) y 5' GTTCTGAGGTCATTACTGG 3' (antisentido) (QIAGEN, 2003). Los plásmidos para secuenciación se purificaron empleando el protocolo descrito en el epígrafe 3.2.5.2.

La calidad de las secuencias obtenidas con los cebadores sentido y antisentido se analizó con el programa Gap v4.10. Las zonas de baja calidad se eliminaron con el programa seqret del paquete EMBOSS, corrido en Linux. El ensamblaje de las secuencias corregidas se realizó con la herramienta ContigExpress del programa Vector NTI Advance v 11.0. Las zonas pertenecientes al vector pQE30 presentes en los productos de los ensamblajes se eliminaron empleando el programa seqret. El alineamiento de las secuencias resultantes, con las secuencias amplificadas de los genes *gapA*, MGA\_0648, MGA\_0993, MGA\_0267, MGA\_0416, MGA\_0319, MGA\_0843, MGA\_0398, *crmA*, *crmC*, *mgc2*, *oppA* y *vlhA.5.02* de la cepa R(low) de *M. gallisepticum*, se realizó con la herramienta Blastn corrida localmente en Linux.

### 3.2.13 Identificación de potenciales genes de proteínas de superficie, en aislamiento cubano de *M. gallisepticum*

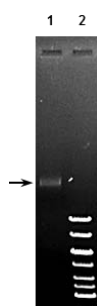
Se evaluó por PCR la posible presencia de genes codificantes para potenciales proteínas de superficie [las primeras 97 de la lista de priorización, (Anexo 3)], en un aislamiento de campo (MYCOLAB, CENSA) obtenido en una granja de la provincia de Mayabeque, a partir de animales con un cuadro agudo de la Enfermedad Crónica Respiratoria. Este estudio se realizó con el empleo de los cebadores diseñados para el clonaje de estos genes en el

epígrafe 3.2.2 (Anexo 6). El ADN del aislamiento se obtuvo como se describe en el epígrafe 3.2.3 y las condiciones de la PCR fueron las mismas descritas en el epígrafe 3.2.4. El resultado de la reacción se evaluó por electroforesis en geles de agarosa, según las indicaciones descritas en el epígrafe 3.2.4..

### 3.3 Resultados y Discusión

#### 3.3.1 Extracción de ADN genómico

Como se puede apreciar en la Figura 3, con el empleo del protocolo descrito, el ADN de la cepa R(low) de *M. gallisepticum* se obtuvo con la calidad necesaria para los siguientes pasos de PCR, no degradado y puro. Se obtuvieron 500μL del mismo, con una concentración de 200ng/μL, suficientes para las necesidades de los siguientes experimentos.



**Figura 3.** Purificación de ADN genómico de *M. gallisepticum*. Electroforesis de agarosa al 1%. **Línea 1:** ADN de la cepa R(low) de *M. gallisepticum*. **Línea 2:** Patrón de Peso Molecular de ADN, 1kb

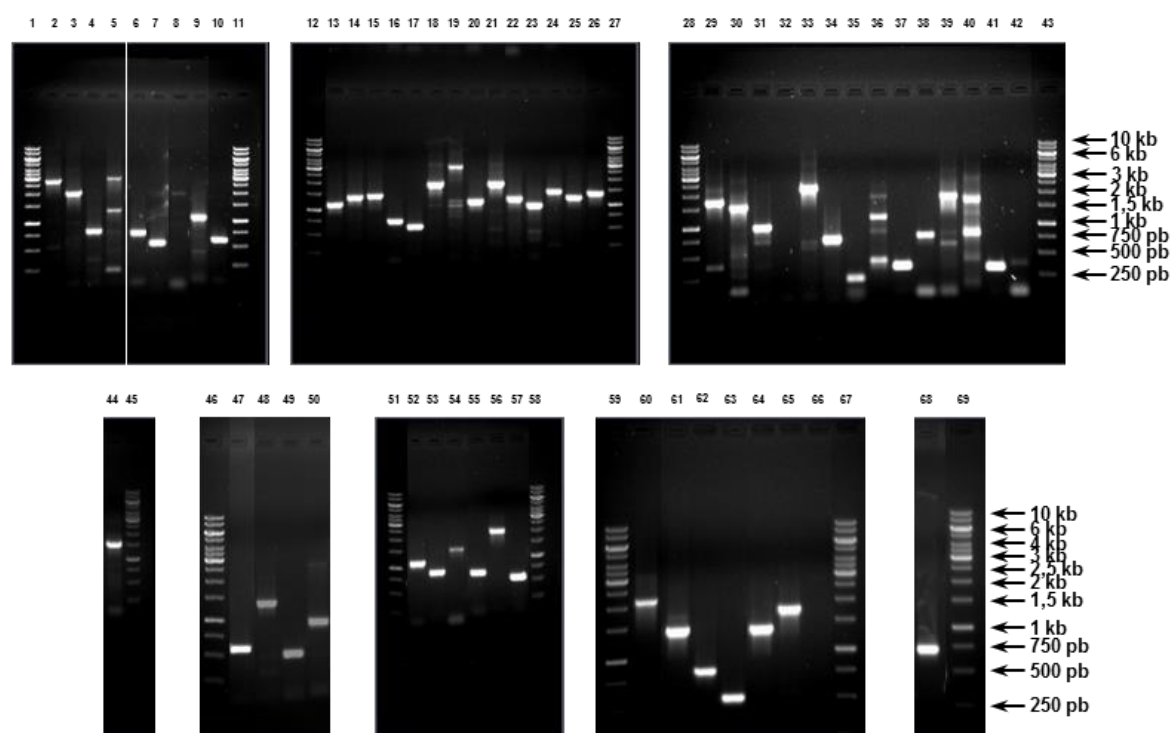
#### 3.3.2 Amplificación por PCR de los genes identificados

Las secuencias de los cebadores diseñados se pueden encontrar en el anexo 6. Al emplear como molde el ADN genómico de la cepa R(low) de *M. gallisepticum*, se lograron amplificar 54 de los 56 genes o fragmentos de genes evaluados, con la excepción de MGA\_0565 y MGA\_1079 (Figura 4). Diferencias en la secuencia genómica de la cepa R(low) con la que se trabajó, con respecto a la secuencia genómica de la cepa R(low) que se secuenció y se depositó en el GeneBank con N° Acceso: AE015450.2, pudieron haber provocado este resultado, en caso de que los cambios hubieran afectado los sitios de hibridación de los

cebadores diseñados a partir de la secuencia. Estas diferencias pudieron haber estado relacionadas o con errores de secuenciación o por fenómenos de recombinación ocurridos en la cepa R(low) con la que trabajamos que son comunes en los micoplasmas (Lee et al., 2008, Allam et al., 2010, Spuesens et al., 2010, Dordet-Frisoni et al., 2014, Zimmerman et al., 2014). En todos los casos de amplificación exitosa se obtuvieron productos con tallas que se correspondían con las tallas esperadas (Anexo 6). Es de significar que en los casos de MGA\_0008, MGA\_0250, MGA\_0281, MGA\_0310, MGA\_0416, MGA\_0485, MGA\_0648, MGA\_0801, MGA\_0993 y MGA\_1161 además del producto esperado, se obtuvieron uno o más productos adicionales (Figura 4). Este resultado se puede explicar porque se empleó un programa único, con una baja temperatura de unión que tenía como objetivo lograr la amplificación de todos los fragmentos, con independencia de la temperatura de fusión de los múltiples cebadores empleados. En las PCR, bajas temperaturas de unión con respecto a las temperaturas de fusión de los cebadores puede provocar inespecificidad en los ensayos, las primeras no deben estar por debajo de los 5°C con respecto a las segundas (Ishmael and Stellato, 2008). En nuestro caso se empleó como temperatura de unión 45°C, mientras la temperatura de fusión de los cebadores específicos a partir de los cuales se generaron productos inespecíficos, se encontraban en el intervalo entre 60,8°C y 81,9°C. Tales diferencias pudieron haber provocado los resultados observados. También podrían estar provocados por fenómenos de recombinación (Lee et al., 2008, Allam et al., 2010, Spuesens et al., 2010, Dordet-Frisoni et al., 2014, Zimmerman et al., 2014). Otra posible explicación es la presencia de repeticiones en el genoma de *M. gallisepticum*; tales repeticiones son bien conocidas y se han documentado frecuentemente tanto para esta bacteria (Jiang et al., 2009, Chen et al., 2010, Tulman et al., 2012, Yong et al., 2015), como para otros micoplasmas (Algire et al., 2012, Zimmerman et al., 2013, Lee et al., 2014, Dos Santos et al., 2015, Xiao et al., 2015).

Como resultado de estas amplificaciones inespecíficas se pueden obtener construcciones recombinantes con fragmentos de la secuencia esperada, o insertos cuyas secuencias no están relacionadas con los genes de interés. En el primer caso el polipéptido recombinante

podría conservar la utilidad de la proteína completa. Previamente existen ejemplos de clonaje de fragmentos de genes de micoplasmas para mapear su antigenicidad y sus dominios funcionales (Schurwanz *et al.*, 2009). Adicionalmente el hecho de que los productos inespecíficos sean minoritarios con respecto al esperado, hace que las posibilidades de clonar uno de estos fragmentos sean inferiores a las de clonar las secuencias completas.

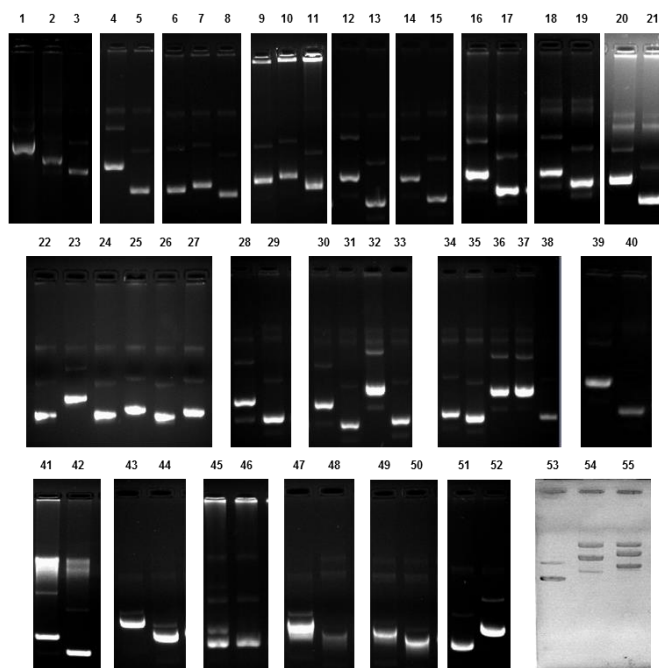


**Figura 4 .** Obtención de los genes a clonar. Electroforesis en agarosa 1% de los productos de PCR con los cebadores específicos para los genes de las primeras 50 proteínas de la lista de priorización y 6 genes de proteínas controles. **Línea 1:** MPM-1kb, **Línea 2:** *crmA*, **Línea 3:** *crmC*, **Línea 4:** *mgc2*, **Línea 6:** MGA\_0008, **Línea 7:** MGA\_0023, **Línea 8:** MGA\_0063, **Línea 9:** MGA\_0084, **Línea 10:** MGA\_0111, **Línea 11:** MPM-1kb, **Línea 12:** MPM-1kb, **Línea 13:** MGA\_0126, **Línea 14:** MGA\_0173, **Línea 15:** MGA\_0211, **Línea 16:** MGA\_0213, **Línea 17:** MGA\_0217, **Línea 18:** MGA\_0226, **Línea 19:** MGA\_0241, **Línea 20:** MGA\_0250, **Línea 21:** MGA\_0267, **Línea 22:** MGA\_0281, **Línea 23:** MGA\_0298, **Línea 24:** MGA\_0310, **Línea 25:** MGA\_0319, **Línea 26:** MGA\_0335, **Línea 27:** MPM-1kb, **Línea 28:** MPM-1kb, **Línea 29:** MGA\_0398, **Línea 30:** MGA\_0416, **Línea 31:** MGA\_0485, **Línea 32:** MGA\_0565, **Línea 33:** MGA\_0648, **Línea 34:** MGA\_0656, **Línea 35:** MGA\_0702, **Línea 36:** MGA\_0801, **Línea 37:** MGA\_0812, **Línea 38:** MGA\_0913, **Línea 39:** MGA\_0993, **Línea 40:** MGA\_1161, **Línea 41:** MGA\_1216, **Línea 42:** MGA\_1318, **Línea 43:** MPM-1kb, **Línea 44:** MGA\_0403, **Línea 45:** MPM-1kb, **Línea 46:** MPM-1kb, **Línea 47:** MGA\_0472, **Línea 48:** MGA\_0487, **Línea 49:** MGA\_0512, **Línea 50:** *oppA*, **Línea 51:** MPM-1kb, **Línea 52:** MGA\_0760, **Línea 53:** MGA\_0789, **Línea 54:** MGA\_0797, **Línea 55:** MGA\_0843, **Línea 56:** MGA\_0849, **Línea 57:** MGA\_0867, **Línea 58:** MPM-1kb, **Línea 59:** MPM-1kb, **Línea 60:** MGA\_0879, **Línea 61:** MGA\_0954, **Línea 62:** *vlhA.5.02*, **Línea 63:** MGA\_0957, **Línea 64:** MGA\_1014, **Línea 65:** MGA\_1037, **Línea 66:** MGA\_1079, **Línea 67:** MPM-1kb, **Línea 68:** MGA\_1315, **Línea 69:** MPM-1kb.

### 3.3.3 Generación de las construcciones recombinantes

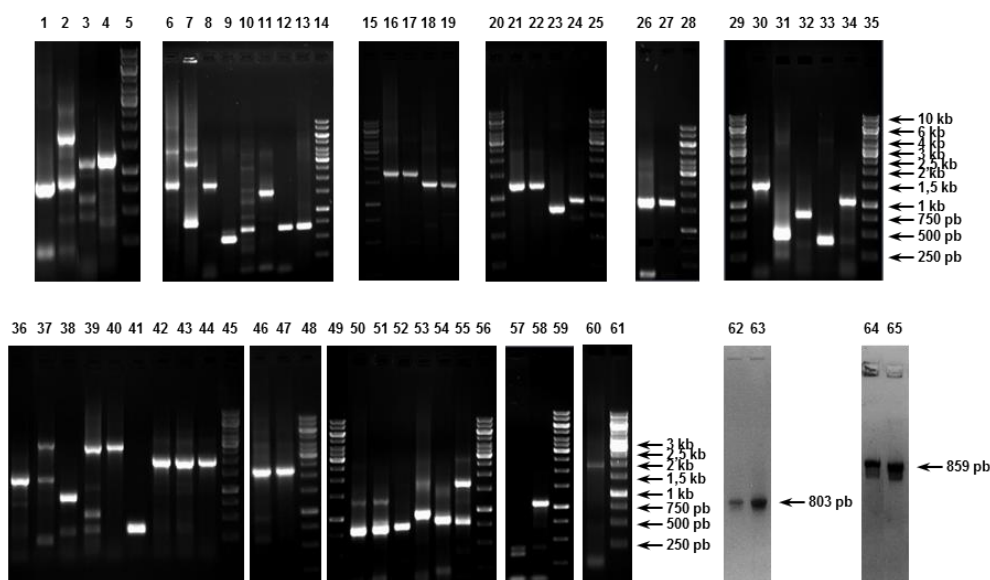
#### 3.3.3.1 Transformación y monitoreo de las construcciones recombinantes

La evaluación preliminar de la transformación inicial, que se realizó con propósitos de monitoreo del clonaje y para, posteriormente, contar con clones de propagación de los plásmidos recombinantes, se realizó mediante la comparación de la talla de las construcciones obtenidas, con la del plásmido pQE30 nativo. En esta primera inspección, se encontraron 30 posibles productos recombinantes, derivados de 19 de los genes en estudio (Figura 5).



**Figura 5.** Monitoreo preliminar de los clones recombinantes. Electroforesis en agarosa 1%, de los plásmidos obtenidos de los minialcalinos con resultados positivos. **Línea 1:** pQE30-MGA\_1161-2, **Línea 2:** pQE30-MGA\_0281-3, **Línea 3:** pQE30, **Línea 4:** pQE30-MGA\_0993-1, **Línea 5:** pQE30, **Línea 6:** pQE30-MGA\_0416-1, **Línea 7:** pQE30-MGA\_0416-5, **Línea 8:** pQE30, **Línea 9:** pQE30-MGA\_0843-2, **Línea 10:** pQE30-MGA\_0993-11, **Línea 11:** pQE30, **Línea 12:** pQE30-MGA\_0267-8, **Línea 13:** pQE30, **Línea 14:** pQE30-MGA\_0416-3, **Línea 15:** pQE30, **Línea 16:** pQE30-MGA\_0319-5, **Línea 17:** pQE30, **Línea 18:** pQE30-MGA\_0485-5, **Línea 19:** pQE30, **Línea 20:** pQE30-MGA\_0398-13, **Línea 21:** pQE30, **Línea 22:** pQE30, **Línea 23:** pQE30-MGA\_0797-29, **Línea 24:** pQE30, **Línea 25:** pQE30-mgc2-10, **Línea 26:** pQE30, **Línea 27:** pQE30-MGA\_0084-2, **Línea 28:** pQE30-gapA-17, **Línea 29:** pQE30, **Línea 30:** pQE30-crmA-13, **Línea 31:** pQE30, **Línea 32:** pQE30-crmA-16, **Línea 33:** pQE30, **Línea 34:** pQE30-crmC-12, **Línea 35:** pQE30, **Línea 36:** pQE30-crmC-17, **Línea 37:** pQE30-crmC-18, **Línea 38:** pQE30, **Línea 39:** pQE30-MGA\_0648-26, **Línea 40:** pQE30, **Línea 41:** pQE30-MGA\_1161-78, **Línea 42:** pQE30, **Línea 43:** pQE30-MGA\_1161-92, **Línea 44:** pQE30, **Línea 45:** pQE30-mgc2-32, **Línea 46:** pQE30, **Línea 47:** pQE30-mgc2-44, **Línea 48:** pQE30, **Línea 49:** pQE30-MGA\_0084-51, **Línea 50:** pQE30, **Línea 51:** pQE30, **Línea 52:** pQE30-MGA\_0760-103, **Línea 53:** pQE30, **Línea 54:** pQE30-vlhA.5.02, **Línea 55:** pQE30-oppA.

Al monitorear por PCR la inserción de las secuencias en el vector, en aquellas construcciones cuya talla había sido mayor que la del plásmido control, se confirmó que las 30 construcciones derivadas de 19 de los genes en estudio, eran recombinantes (Figura 6).



**Figura 6.** . Monitoreo de los clones recombinantes. Electroforesis en agarosa 1%, de los PCR de las potenciales construcciones recombinantes. **Línea 1:** pQE30-MGA\_1161-2, **Línea 2:** MGA\_1161, **Línea 3:** pQE30-MGA\_0281-3, **Línea 4:** MGA\_0281, **Línea 5:** MPM-1kb, **Línea 6:** pQE30-MGA\_0993-1, **Línea 7:** pQE30-MGA\_0993-11, **Línea 8:** MGA\_0993, **Línea 9:** pQE30-MGA\_0416-1, **Línea 10:** pQE30-MGA\_0416-5, **Línea 11:** MGA\_0416, **Línea 12:** pQE30-MGA\_0843-2, **Línea 13:** MGA\_0843, **Línea 14:** MPM-1kb, **Línea 15:** MPM-1kb, **Línea 16:** pQE30-MGA\_0267-8, **Línea 17:** MGA\_0267, **Línea 18:** pQE30-MGA\_0416-3, **Línea 19:** MGA\_0416, **Línea 20:** MPM-1kb, **Línea 21:** pQE30-MGA\_0319-5, **Línea 22:** MGA\_0319, **Línea 23:** pQE30-MGA\_0485-5, **Línea 24:** MGA\_0485, **Línea 25:** MPM-1kb, **Línea 26:** pQE30-MGA\_0398-13, **Línea 27:** MGA\_0398, **Línea 28:** MPM-1kb, **Línea 29:** MPM-1kb, **Línea 30:** pQE30-MGA\_0797-29, **Línea 31:** pQE30-mgc2-10, **Línea 32:** *mgc2*, **Línea 33:** pQE30-MGA\_0084-2, **Línea 34:** MGA\_0084, **Línea 35:** MPM-1kb, **Línea 36:** pQE30-gapA-17, **Línea 37:** gapA, **Línea 38:** pQE30-crmA-13, **Línea 39:** pQE30-crmA-16, **Línea 40:** *crmA*, **Línea 41:** pQE30-crmC-12, **Línea 42:** pQE30-crmC-17, **Línea 43:** pQE30-crmC-18, **Línea 44:** *crmC*, **Línea 45:** MPM-1kb, **Línea 46:** pQE30-MGA\_0648-26, **Línea 47:** MGA\_0648, **Línea 48:** MPM-1kb, **Línea 49:** MPM-1kb, **Línea 50:** pQE30-mgc2-32, **Línea 51:** pQE30-mgc2-44, **Línea 52:** *mgc2*, **Línea 53:** pQE30-MGA\_1161-78, **Línea 54:** pQE30-MGA\_1161-92, **Línea 55:** MGA\_1161, **Línea 56:** MPM-1kb, **Línea 57:** pQE30-MGA\_0084-51, **Línea 58:** MGA\_0084, **Línea 59:** MPM-1kb, **Línea 60:** pQE30-MGA\_0760-103, **Línea 61:** MPM-1kb, **Línea 62:** pQE30-vlhA.5.02, **Línea 63:** *vlhA.5.02*, **Línea 64:** pQE30-*oppA*, **Línea 65:** *oppA*.

Como se puede apreciar, no en todos los casos, la talla del producto de amplificación a partir de los plásmidos recombinantes, coincide con la talla del fragmento original correspondiente (Anexo 6), en estos casos se encontraron insertos tanto de mayor como de menor talla. No obstante, en el caso de 15 de los 19 genes obtenidos, se obtuvo al menos una construcción recombinante con la talla correcta (MGA\_1161, MGA\_0281, MGA\_0993, MGA\_0843, MGA\_0267, MGA\_0416, MGA\_0319, MGA\_0398, *crmA*, *crmC*, MGA\_0648, *mgc2*, MGA\_0760, *oppA* y *vlhA.5.02*) (Tabla 7).



**Tabla 7** Monitoreo de los posibles clones recombinantes. Comparación de la talla de los productos de PCR obtenidos con la talla esperada

Construcción	Clon	Talla esperada	Talla obtenida
pQE30-MGA_0084	2	1084	400
pQE30-MGA_0084	51	1084	400
pQE30-MGA_0267	8	1813	1813
pQE30-MGA_0281	3	1219	1219
pQE30-MGA_0319	5	1243	1243
pQE30-MGA_0398	13	1687	1687
pQE30-MGA_0416	1	1513	550 *
pQE30-MGA_0416	5	1513	650 *
pQE30-MGA_0416	3	1513	1513
pQE30-MGA_0485	5	964	800 *
pQE30-MGA_0648	26	2200	2200
pQE30-MGA_0760	103	1078	1900 *
pQE30-MGA_0797	29	1501	1501
pQE30-MGA_0843	2	679	679
pQE30-MGA_0993	1	1813	1813
pQE30-MGA_0993	11	1813	850; 3100 *
pQE30-MGA_1161	2	1687	750
pQE30-MGA_1161	78	1687	1000 *
pQE30-MGA_1161	92	1687	750
pQE30-crmA	13	1984	900 *
pQE30-crmA	16	2734	2734
pQE30-crmC	12	1984	350 *
pQE30-crmC	17	1984	1984
pQE30-crmC	18	1984	1984
pQE30-gapA	17	2926	1400 *
pQE30-mgc2	10	796	500 *
pQE30-mgc2	32	796	796
pQE30-mgc2	44	796	796
pQE30-oppA	2	859	859
pQE30-vlhA.5.02	6	800	800

\* Construcciones recombinantes con un inserto de talla diferente a la esperada

Se observa que las variantes son diversas y, en términos de talla, se mueven en ambas direcciones. Este resultado es consecuencia directa del fenómeno de los múltiples productos que se obtuvieron en el paso de amplificación de los genes a partir del ADN cromosomal, que ya fue analizado y explicado. Esto pudo haberse evitado purificando los fragmentos a clonar a partir de los geles de agarosa, como es común cuando se trabaja con pocos blancos, pero esto no es factible, tampoco, en las tecnologías de alto flujo. Esta técnica habría complicado y demorado enormemente el proceso de clonaje, por lo que era más

factible afrontar las consecuencias ya descritas. Esta variedad de fragmentos llegó hasta la etapa de ligazón y clonaje, y probablemente es la causa de la obtención de construcciones con tallas diferentes a las esperadas. Algunas de las construcciones en las cuales se detectó la presencia de insertos de talla diferente a la esperada, se generaron a partir de productos de PCR en los cuales no se observó este fenómeno, pero puede ser, que aunque no se observaran, los fragmentos de otras tallas si estuvieran presentes y no se detectaran por estar a muy bajas concentraciones, por debajo del límite de detección del sistema de visualización. Para confirmar los resultados, las potenciales construcciones recombinantes generadas sobre la base de estos errores, serian fácilmente filtradas más adelante por el propio proceso, al no generar proteínas recombinantes, o no ser estas reconocidas por los anticuerpos contra *M. gallisepticum*.

### **3.3.3.2 Obtención de los polipéptidos recombinantes a partir de las construcciones generadas.**

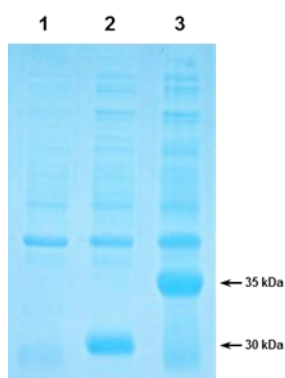
Para los experimentos de expresión se excluyeron algunas de las construcciones recombinantes, aunque se garantizó la presencia de todos los genes clonados. No se continuó trabajando con pQE30-MGA\_0084-1 y pQE30-MGA\_1161-2 ya que a partir de estos se generaban productos de PCR mucho más pequeños que lo esperado y se contaba con otro u otros clones derivados de estos genes, con insertos de la misma talla. Algo muy parecido sucedió con pQE30-MGA\_0416-1, pQE30-crmA-13, pQE30-crmC-12 y pQE30-mgc2-10, todos daban lugar a productos de menor talla pero en el caso de estos se contaba con al menos una, y hasta dos construcciones que tenían la talla adecuada (Tabla 7).

Las 24 construcciones seleccionadas para los siguientes experimentos de expresión se muestran en la Tabla 8.

**Tabla 8. Construcciones seleccionadas para realizar los experimentos de expresión**

Construcción	Clon	Construcción	Clon	Construcción	Clon
pQE30-MGA 0084	51	pQE30-MGA 0648	26	pQE30-gapA	17
pQE30-MGA 0267	8	pQE30-MGA 0760	103	pQE30-crmA	16
pQE30-MGA 0281	3	pQE30-MGA 0797	29	pQE30-crmC	17
pQE30-MGA 0319	5	pQE30-MGA 0843	2	pQE30-crmC	18
pQE30-MGA 0398	13	pQE30-MGA 0993	1	pQE30-mgc2	32
pQE30-MGA 0416	3	pQE30-MGA 0993	11	pQE30-mgc2	44
pQE30-MGA 0416	5	pQE30-MGA 1161	78	pQE30-oppA	-
pQE30-MGA 0485	5	pQE30-MGA 1161	92	pQE30-vihA.5.02	-

Los resultados de la expresión de los fragmentos clonados con el empleo del método de autoinducción con lactosa mostraron lo acertado de la selección de la cepa y el método para inducir la expresión. Los dos polipéptidos de 289 y 308 aa correspondientes a los fragmentos de las proteínas VihA.5.02 y OppA, respectivamente, se obtuvieron con las tallas esperadas. Las correspondientes bandas de aproximadamente 31kDa y 35kDa se observaron en los lisados de los cultivos inducidos por SDS-PAGE con un gel al 10 % (p/v) (Figura 7).



**Figura 7.** SDS-PAGE 10% de las muestras resultantes de la autoinducción de las células C41pRospLysS transformadas con los plásmidos recombinantes pQE30-VihA5.02 y pQE30-OppA. Línea 1: C- lisado de la cepa C41pRosPLysS. Línea 2: pQE30- VihA5.02. Línea 3: pQE30-OppA

Los niveles de expresión obtenidos mediante la autoinducción fueron similares a los alcanzados con el método tradicional de inducción con IPTG, aunque los rendimientos totales fueron superiores (aproximadamente 3 ó 4 veces) debido al mayor crecimiento de la masa celular que se genera por autoinducción (Studier, 2005). En adición a esto, el IPTG es un reactivo costoso que en este caso se sustituye  $\alpha$ -lactosa, que es mucho más económica. En este sistema la lactosa se emplea como inductor y los niveles basales de expresión de la proteína blanco en los momentos iniciales del cultivo, se evitan por la adición de glucosa al

medio. La glucosa impide la inducción por lactosa mediante mecanismos muy bien estudiados (Mobbs et al., 2007, Santillan and Mackey, 2008). En el protocolo original las cepas de expresión se cultivan toda la noche hasta la saturación en medio de no-inducción (Studier, 2005), en este trabajo este medio se sustituyó por LB común porque los rendimientos fueron similares y el LB es mucho más sencillo y barato (Agüero et al., 2008).

La expresión de las proteínas de micoplasmas en *E. coli* es una tarea complicada y difícil, por el uso de codones de los mismos. Debido a los bajos niveles de expresión que se obtuvieron con el empleo de la cepa ISM612, fue imposible monitorear el proceso de expresión por SDS-PAGE, y fue necesario emplear el Western-Blot con un anticuerpo monoclonal específico para péptidos de poli-histidina, por ser este un sistema más sensible. En este caso, a partir de las 22 cepas recombinantes (en ISM612), se pudo detectar la presencia de productos recombinantes en 15 de ellas, correspondientes a 11 genes diferentes (MGA\_0993, MGA\_0416, MGA\_0843, MGA\_0267, MGA\_0319, MGA\_0398, gapA, crmA, crmC, MGA\_0648 y mgc2) (Figura 8). En general se obtuvo expresión en 17 construcciones recombinantes, correspondientes a 13 genes de los genes evaluados.

Como se puede apreciar, en prácticamente todos los casos donde se observa expresión, se obtiene más de un fragmento. La cepa ISM612, que ha sido empleada para expresar epítomos de la proteína P1 de *M. pneumoniae* (Smiley and Minion, 1993) y para evaluar una librería genética de *M. hyopneumoniae* (Minion et al., 1995), con la cual se logró incrementar el reconocimiento de productos recombinantes de micoplasmas con anticuerpos (Smiley and Minion, 1993), contiene un supresor opal inducible (trpT), así como la mutación prfB3 del factor 2 de liberación. A pesar de estas modificaciones genéticas, con esta cepa no siempre se logra la expresión de los genes recombinantes, y cuando se consigue, generalmente es a bajos niveles, además, es muy común la generación no sólo de productos de la talla esperada, sino de tallas intermedias (Smiley and Minion, 1993, Minion et al., 1995). La terminación prematura de la traducción a causa de la presencia de los codones UGA, observada en estas investigaciones previas, explica los resultados obtenidos en este trabajo.

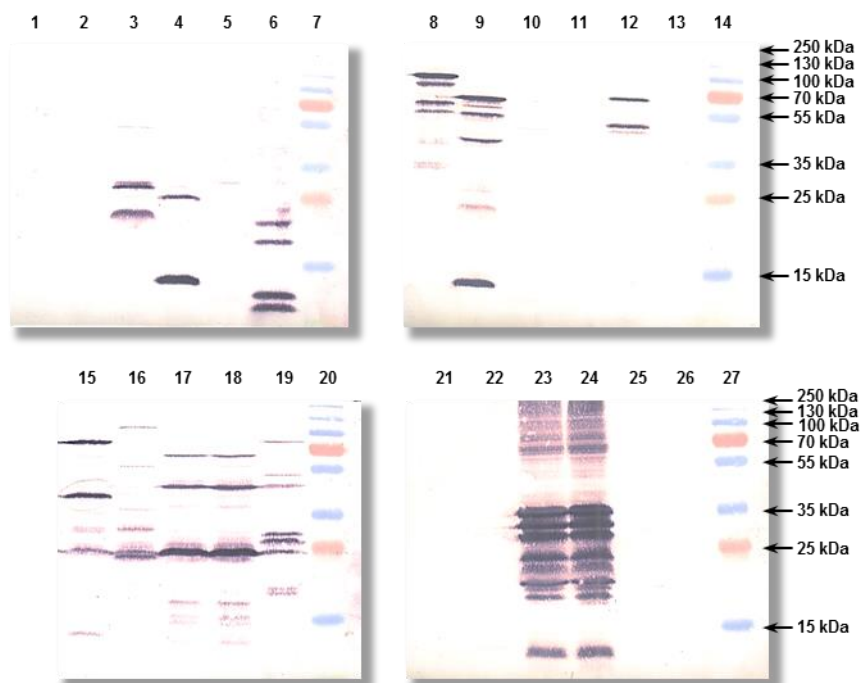
Se han empleado otras variantes para expresar genes de micoplasmas como la expresión en otros micoplasmas (Kent *et al.*, 2012), y en los casos en que pocos codones estén presentes en la secuencia a clonar, la mutagénesis sitio-dirigida (Simionatto *et al.*, 2009), pero todas estas estrategias han brindado, de manera general, resultados pocos satisfactorios y, en el caso particular de esta última, consume mucho tiempo y es cara, creciendo su costo con el número de codones UGA a mutar (Halbedel and Stulke, 2007)

En el futuro se puede pensar en la obtención de los antígenos recombinantes mediante la generación de genes sintéticos (Blackburn *et al.*, 2016, Zampini *et al.*, 2015). Este procedimiento, aunque es más caro, posee dos ventajas significativas. Simplifica extraordinariamente el trabajo y, consecuentemente, el tiempo necesario para obtener los resultados. Por otro lado, además de eliminar el problema de la presencia de los codones opal en la secuencia del gen a expresar, permite muy altos niveles de generación del producto recombinante, ya que en el proceso de preparación de estas construcciones, se optimizan los genes para su expresión en el sistema heterólogo seleccionado (Perez-Casal *et al.*, 2015). Esto incrementa los rendimientos del proceso de obtención de los productos recombinantes. De esta manera podrían obtenerse los productos que en esta oportunidad no fue posible expresar y, además, los rendimientos serían mucho más altos y la calidad de los productos recombinantes, superior.

### 3.3.4 Purificación de los polipéptidos recombinantes de *M. gallisepticum*

Para los subsiguientes pasos se excluyeron de la investigación, aquellas construcciones a partir de las cuales no se habían obtenido productos recombinantes, (pQE30-MGA\_0084-51, pQE30-MGA\_0281-3, pQE30-MGA\_0485-5, pQE30-MGA\_0760-103, pQE30-MGA\_0797-29, pQE30-MGA\_1161-78 y pQE30-MGA\_1161-92). También se excluyeron aquellas de las que se disponía de más de un clon (pQE30-MGA\_0993-11, pQE30-crmC-18, pQE30-mgc2-44), con la excepción de pQE30-MGA\_0416, ya que los dos clones de este gen, poseían patrones de expresión diferentes (Figura 8) A continuación se muestra el listado de los 14

productos recombinantes, correspondientes a 13 genes, con los que se continuó la investigación (Tabla 9).



**Figura 8.** Western-Blot con anticuerpo monoclonal específico a péptidos de poli-histidina (antecedido por SDS-PAGE 12%), de las células totales después de la inducción con IPTG de los cultivos de *E. coli* ISM612 de las construcciones recombinantes seleccionadas. **Línea 1:** pQE30, **Línea 2:** pQE30-0281-3, **Línea 3:** pQE30-0993-1, **Línea 4:** pQE30-0416-5, **Línea 5:** pQE30-0843-2, **Línea 6:** pQE30-0993-11, **Línea 7:** PPM, **Línea 8:** pQE30-0267-8, **Línea 9:** pQE30-0416-3, **Línea 10:** pQE30-0319-5, **Línea 11:** pQE30-0485-5, **Línea 12:** pQE30-0398-13, **Línea 13:** pQE30-0797-29, **Línea 14:** PPM, **Línea 15:** pQE30-gapA-17, **Línea 16:** pQE30-crmA-16, **Línea 17:** pQE30-crmC-17, **Línea 18:** pQE30-crmC-18, **Línea 19:** pQE30-0648-26, **Línea 20:** PPM, **Línea 21:** pQE30-1161-78, **Línea 22:** pQE30-1161-92, **Línea 23:** pQE30-mgc2-32, **Línea 24:** pQE30-mgc2-44, **Línea 25:** pQE30-0084-51, **Línea 26:** pQE30-0760-103, **Línea 27:** PPM.

**Tabla 9.** Construcciones seleccionadas para realizar los experimentos de solubilización y purificación

Construcción	Clon	Construcción	Clon	Construcción	Clon
pQE30-MGA_0267	8	pQE30-MGA_0648	26	pQE30-crmC	17
pQE30-MGA_0319	5	pQE30-MGA_0843	2	pQE30-mgc2	32
pQE30-MGA_0398	13	pQE30-MGA_0993	1	pQE30-oppA	2
pQE30-MGA_0416	3	pQE30-gapA	17	pQE30-vlhA.5.02	6
pQE30-MGA_0416	5	pQE30-crmA	16		

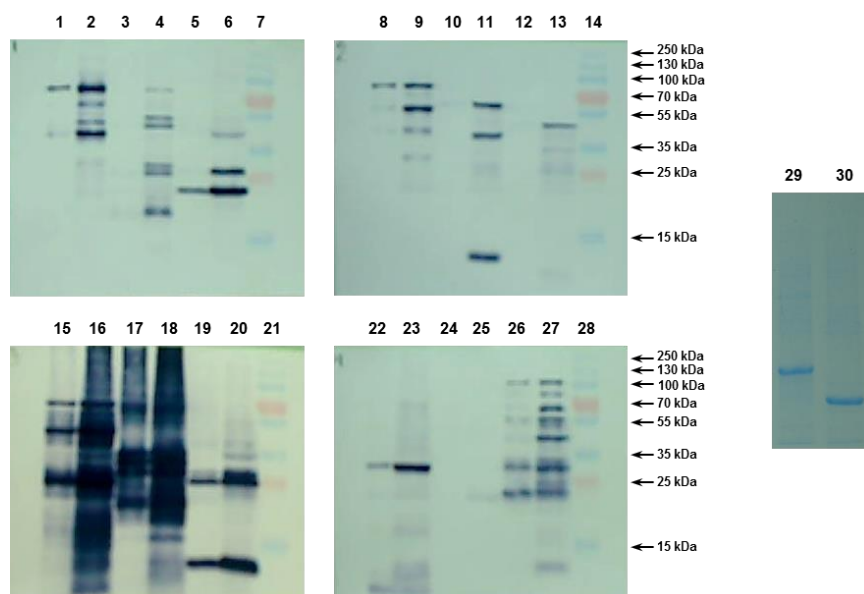
### 3.3.4.1 Extracción y solubilización de los productos recombinantes

Altos niveles de expresión de las proteínas recombinantes usualmente conllevan la acumulación de la proteína en forma de agregados insolubles, frecuentemente denominados cuerpos de inclusión (CI). Debido a que los CI, durante mucho tiempo se han considerado productos de desecho esencialmente formados por proteínas mal plegadas, propensas a

agregarse, se han realizado importantes esfuerzos para eliminar, o al menos parcialmente disminuir, su formación. Sin embargo, en los últimos años el escenario ha cambiado completamente. Recientemente se ha demostrado que las proteínas incluidas en CI no están necesariamente mal plegadas, sin actividad biológica y por ende, pueden ser empleadas con fines biomédicos (Garcia-Fruitos, 2010, Rodríguez-Carmona et al., 2010). No obstante, para cualquier uso terapéutico posterior es recomendable solubilizar y renaturalizar las proteínas *in vitro*.

Debido a los altos niveles de expresión que se obtuvieron para OppA y VlhA.5.02, era de esperar que se generaran en forma insoluble. Por los bajos niveles de expresión del resto, podía suponerse la generación de los productos en forma soluble, Esto no ocurrió así, probablemente porque se trataba de proteínas de membrana las cuales son particularmente insolubles (Roussel et al., 2013, Routledge et al., 2016). Por esta razón, fue necesario someter a pasos de solubilización a todos los productos generados.

Una de las ventajas de la obtención de las proteínas en CI, es que la mayor parte de las proteínas del hospedero se expresa de forma soluble y una simple centrifugación, a continuación de la lisis celular, permite eliminar la mayor parte de los contaminantes sin la necesidad de emplear complicadas y costosas técnicas de purificación (Peternel and Komel, 2010). La solubilización de las proteínas con SDS permitió la obtención de las proteínas puras de manera soluble, sin necesidad de la aplicación de complicados y poco eficientes procesos de renaturalización. Como se puede apreciar en la Figura 9, contrario a lo esperado, para solubilizar los productos obtenidos en ISM612, que se expresaron en muy bajas cantidades, fue necesario emplear una concentración de SDS superior a la utilizada para solubilizar a OppA y VlhA.5.02 que se obtuvieron a altos niveles. Esto quizá se deba al hecho de que estas dos últimas proteínas eran fragmentos que se seleccionaron, entre otras razones, porque probablemente no poseían regiones transmembrana que son altamente hidrofóbicas y, consecuentemente, disminuyen la solubilidad del producto recombinante.



**Figura 9.** Western-Blot (antecedido por SDS-PAGE 12%) con anticuerpo monoclonal específico a péptidos de poli-histidina, para evaluar la solubilización de los productos recombinantes expresados en *E. coli* ISM612 y SDS-PAGE 10% (Tinción con azul de Coomassie) para el caso de la solubilización de OppA y VlhA.5.02. **Línea 1:** pQE30-gapA-17 (Se), **Línea 2:** pQE30-gapA-17 (S), **Línea 3:** pQE30-0648-26 (Se), **Línea 4:** pQE30-0648-26 (S), **Línea 5:** pQE30-0993-1 (Se), **Línea 6:** pQE30-0993-1 (S), **Línea 7:** PPM, **Línea 8:** pQE30-0267-8 (Se), **Línea 9:** pQE30-0267-8 (S), **Línea 10:** pQE30-0416-3 (Se), **Línea 11:** pQE30-0416-3 (S), **Línea 12:** pQE30-0319-5 (Se), **Línea 13:** pQE30-0319-5 (S), **Línea 14:** PPM, **Línea 15:** pQE30-0843-2 (Se), **Línea 16:** pQE30-0843-2 (S), **Línea 17:** pQE30-0398-13 (Se), **Línea 18:** pQE30-0398-13 (S), **Línea 19:** pQE30-crmA-16 (Se), **Línea 20:** pQE30-crmA-16 (S), **Línea 21:** PPM, **Línea 22:** pQE30-crmC-17 (Se), **Línea 23:** pQE30-crmC-17 (S), **Línea 24:** pQE30-mgc2-32 (Se), **Línea 25:** pQE30-mgc2-32 (S), **Línea 26:** pQE30-0416-5 (Se), **Línea 27:** pQE30-0416-5 (S), **Línea 28:** PPM, **Línea 29:** pQE30-oppA (S), **Línea 30:** pQE30-vlhA.5.02 (S). Donde Se se refiere al sedimento producto de la centrifugación de la mezcla de solubilización y S al sobrenadante.

La eliminación de los agentes solubilizantes con el objetivo de renaturalizar los productos recombinantes, es particularmente complicada cuando se emplean agentes caotrópicos como urea o hidrocloreuro de guanidinio. Generalmente conlleva la precipitación de la proteína y bajos recobrados. Una alternativa puede ser el empleo de SDS, pero, a diferencia del resto de los agentes caotrópicos que se emplean en la solubilización de las proteínas recombinantes, el SDS interfiere en la purificación por IMAC (Schlager *et al.*, 2012). Sin embargo, un protocolo descrito recientemente (Schlager *et al.*, 2012, Roussel *et al.*, 2013) soluciona este problema de una manera sencilla. La incubación de la muestra por 30 minutos a 4°C y su posterior centrifugación, permiten la eliminación del SDS del sistema por precipitación. Ni las proteínas del hospedero ni las recombinantes co-precipitan (Schlager *et al.*, 2012). Otras ventajas consisten en la velocidad y la fácil ejecución del procedimiento y el bajo costo del SDS. Los resultados de este trabajo coincidieron con lo reportado, las proteínas recombinantes se solubilizaron con SDS y este se eliminó del sistema sin que esto implicara la re-precipitación indeseada de los productos recombinantes.

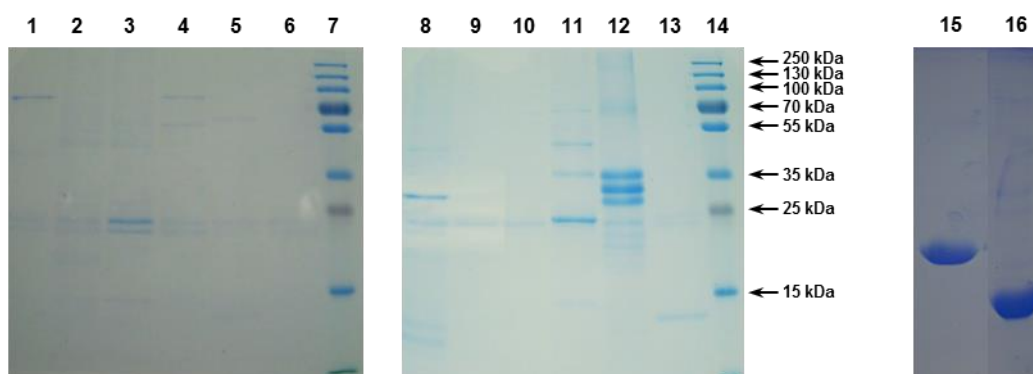


El hecho de obtener las proteínas bajo condiciones nativas es un importante resultado, ya que facilita los siguientes experimentos.

### 3.3.4.2 Purificación por IMAC

La ruptura celular previa a la solubilización con SDS facilitó disminuir considerablemente los niveles de impurezas de las muestras, sobre todo, por la eliminación, mediante centrifugación, de las proteínas solubles citoplasmáticas. La purificación por IMAC permitió incrementar la pureza de los productos, pero, en lo fundamental, aprovechar la propiedad que tienen estos sistemas de concentrar las muestras. Este aspecto era muy importante debido al bajo nivel de producción de las proteínas recombinantes en la cepa ISM612.

Los resultados de la purificación se observan en la Figura 10. Como se aprecia, las proteínas se eluyeron de las columnas, luego del sencillo y rápido proceso cromatográfico que solo incluyó tres soluciones tampón.

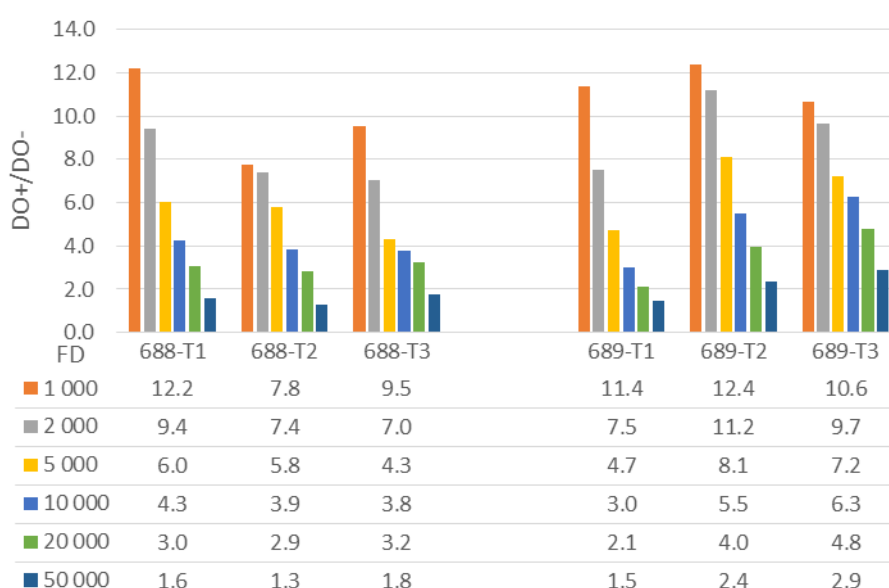


**Figura 10.** SDS PAGE 12% (Expresión en *E. coli* ISM612) y SDS-PAGE 10% (OppA y VlhA.5.02) para evaluar la purificación de los productos recombinantes. **Línea 1:** pQE30-gapA-17, **Línea 2:** pQE30-0648-26, **Línea 3:** pQE30-0993-1, **Línea 4:** pQE30-0267-8, **Línea 5:** pQE30-0416-3, **Línea 6:** pQE30-0319-5, **Línea 7:** PPM, **Línea 8:** pQE30-0843-2, **Línea 9:** pQE30-0398-13, **Línea 10:** pQE30-crmA-16, **Línea 11:** pQE30-crmC-17, **Línea 12:** pQE30-mgc2-32, **Línea 13:** pQE30-0416-5, **Línea 14:** PPM, **Línea 15:** pQE30-oppA, **Línea 16:** pQE30-vlhA.5.

El paso por las columnas, en ausencia total del agente solubilizante (SDS), no provocó la precipitación de los productos recombinantes, fenómeno que suele ocurrir cuando se trabaja con los agentes caotrópicos antes mencionados. Adicionalmente, permitió la concentración de los productos, aspecto este importante para los expresados en la cepa ISM612, algunos de los cuales ya visibles por SDS-PAGE.

### 3.3.5 Generación de anticuerpos específicos contra las proteínas de membrana externa

En los dos conejos inmunizados (688 y 689) se indujo una fuerte respuesta de anticuerpos contra las proteínas de superficie. Los títulos de los seis sueros evaluados (obtenidos de cada uno de los conejos en los tres tiempos de extracción) están por encima de 1:20 000 y los sueros correspondientes a los dos últimos tiempos del conejo 689 tienen títulos de 1:50 000 (Figura 11) Todos los sueros se unieron para homogenizar los análisis posteriores y hacerlos reproducibles. Estos anticuerpos específicos contra las proteínas de membrana de *M. gallisepticum* se utilizaron para confirmar la localización en la superficie de las moléculas en estudio.



**Figura 11.** ELISA para la titulación de los sueros contra proteínas de membrana de *M. gallisepticum*, obtenidos en los conejos 688 y 689 a los tiempos T1 (42 días post-inoculación), T2 (56 días post-inoculación) y T3 (70 días post-inoculación). FD – Factor de dilución

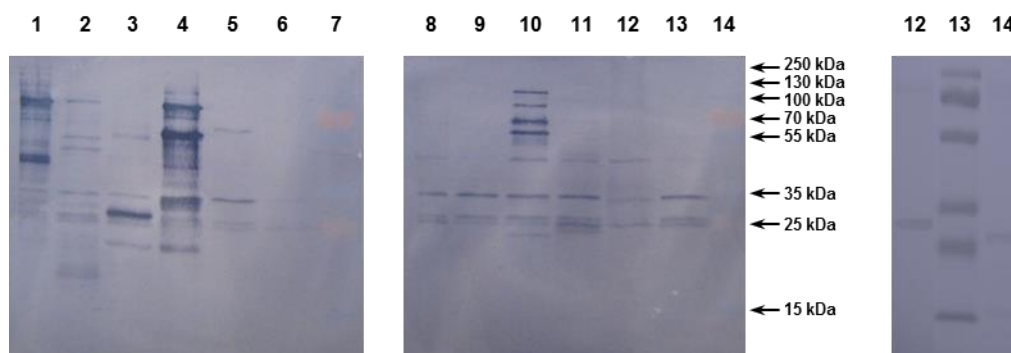
### 3.3.6 Evaluación de la antigenicidad de los polipéptidos recombinantes

La identificación *in silico* de un marco de lectura abierto en la secuencia genómica de un microorganismo, no garantiza la existencia de este gen, y aún, en caso de existir, que se exprese en todas las condiciones (Hyatt et al., 2010, Hyatt et al., 2012). Por esta razón, un primer paso, antes de evaluar los resultados del predictor, era determinar si realmente las proteínas en estudio formaban parte del proteoma de *M. gallisepticum*. Una manera de

lograr este objetivo es la identificación inmunoquímica de las moléculas. Esto permitió, adicionalmente determinar si los productos recombinantes eran antigénicos, si las proteínas nativas eran inmunogénicas y aportó las primeras evidencias experimentales de la localización en la superficie bacteriana, de las proteínas identificadas computacionalmente.

Como se puede apreciar en la Figura 12, el resultado de este ensayo mostró que, de las 13 proteínas recombinantes evaluadas, nueve (GapA, MGA\_0648, MGA\_0993, MGA\_0267, MGA\_0416, CrmA, CrmC, OppA y VlhA.5.02) reaccionaron con el antisuero. Esto demostró que estas construcciones codificaban para fragmentos de proteínas inmunogénicas presentes en el proteoma de *M. gallisepticum*. Las proteínas de superficie bacterianas son las que tienen las mayores posibilidades de interactuar con el sistema inmune y de estimular una respuesta efectiva (de Alvarenga Mudadu *et al.*, 2015), por esta razón se han desarrollado vacunas basadas únicamente en las proteínas de membrana (Madampage *et al.*, 2015) y son el blanco principal en los estudios de VR (Seib *et al.*, 2012). El reconocimiento por parte del antisuero contra la cepa R(low) de *M. gallisepticum* de los productos recombinantes obtenidos, podría ser una evidencia de la posible localización en la superficie bacteriana de las mismas, aunque no es concluyente ya que aún existe una población de anticuerpos que se induce contra moléculas provenientes de otros compartimentos celulares. La presencia de señales únicas en las muestras identificadas como positivas, no presentes en ninguna otra muestra, garantiza la especificidad de esta reacción. Las señales derivadas de reacciones inespecíficas con proteínas contaminantes de *E. coli*, están presentes, como se aprecia en la figura, en todos carriles.

Este resultado implica que los productos recombinantes son antigénicos y que conservan esta propiedad, incluso después del proceso de purificación. Además, quedó demostrado que las proteínas nativas correspondientes a estos productos recombinantes, son inmunogénicas.



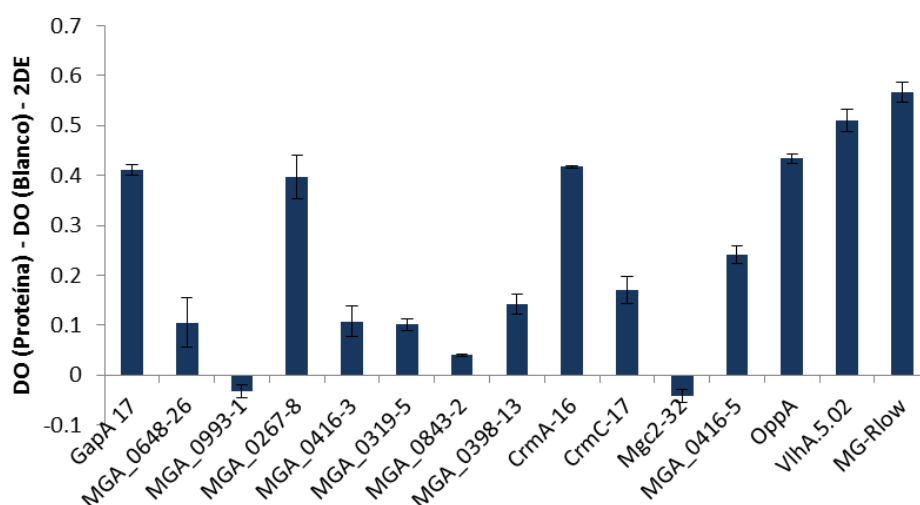
**Figura 12.** Western-Blot con anticuerpo policlonal específico para proteínas totales de la cepa R(low) de *M. gallisepticum* para evaluar la antigenicidad de los productos recombinantes. Línea 1: pQE30-gapA-17 , Línea 2: pQE30-MGA\_0648-26, Línea 3: pQE30-MGA\_0993-1 , Línea 4: pQE30-MGA\_0267-8 , Línea 5: pQE30-MGA\_0416-3, Línea 6: pQE30-MGA\_0319-5 , Línea 7: PPM, Línea 8: pQE30-MGA\_0843-2 , Línea 9: pQE30-MGA\_0398-13 , Línea 10: pQE30-crmA-16 , Línea 11: pQE30-crmC-17 , Línea 12: pQE30-mgc2-32 , Línea 13: pQE30-MGA\_0416-5 , Línea 14: PPM, Línea 15: pQE30-oppA , Línea 16: PPM, Línea 17: pQE30-vlhA.5.02 .

De las proteínas identificadas, cinco corresponden a proteínas de membrana conocidas que se habían incluido como control en el estudio, lo que demuestra la confiabilidad de la metodología empleada. Se conoce que tres de estas proteínas [GapA (Shil *et al.*, 2011a), CrmA (Shil *et al.*, 2011a) y VlhA.5.02] participan activamente en los mecanismos de patogenicidad de la bacteria (Bearson *et al.*, 2003). De estas, la primera, está identificada como la principal citadhesina de *M. gallisepticum* (Shil *et al.*, 2011a, Shil *et al.*, 2011b). La intensidad de la señal de reconocimiento de dos de las nuevas proteínas obtenidas, MGA\_0993 y MGA\_0267 (particularmente esta última), fue de la misma magnitud que la intensidad con que se reconoció GapA, lo cual es un resultado promisorio.

Es importante señalar que la ausencia de señal en los Western-Blot no implica, necesariamente, que los productos recombinantes no son antigénicos. Varios factores pueden explicar este resultado. Los epítomos dominantes de estas moléculas podrían ser conformacionales y los anticuerpos del suero, generados contra las proteínas nativas, ser incapaces de reconocer a los polipéptidos recombinantes, sobre todo por Western-Blot, técnica en la que las proteínas se desnaturalizan luego de calentarse en presencia de  $\beta$ -mercaptoetanol y de estar en presencia de SDS. Por otro lado la causa podría estar relacionada con el fenómeno de variación antigénica de las proteínas de superficie, tan marcado en el género *Mycoplasma* (Adan-Kubo *et al.*, 2012, Hermeyer *et al.*, 2012, Dijkman

et al., 2014, Zimmerman et al., 2014, Xu et al., 2014, Calcutt and Foecking, 2015) y, en particular, en *M. gallisepticum* (Grodio et al., 2012, Tulman et al., 2012, Pflaum et al., 2015). Si la expresión de alguna proteína correspondiente a los productos en estudio, hubiera estado inactivada en el antígeno que se empleó para inmunizar los conejos, no se inducirían anticuerpos contra la misma y entonces sería imposible detectar la expresión del producto recombinante. Esta variación es uno de los principales mecanismos empleados por los micoplasmas para evadir al sistema inmune (Rivera-Tapia, 2002, Citti et al., 2010) y ha incidido negativamente en el diagnóstico serológico de los micoplasmas (Noormohammadi, 2007) y en particular, de *M. gallisepticum* (Bencina, 2002).

Para complementar este resultado y evaluar si las reacciones negativas eran consecuencia de la desnaturalización de las proteínas que se produce en el Western-Blot, se evaluó la reacción en condiciones nativas, por ELISA (Figura 13). Los resultados confirman los obtenidos por Western-Blot, aunque, probablemente debido a la mayor sensibilidad de este último método, todos los productos recombinantes fueron reconocidos por el antisuero contra la cepa R(low), con la excepción de los productos MGA\_0993 y Mgc2. La proteína MGA\_0267 es reconocida incluso con mayor fuerza que GapA y CrmA. Esta proteína podría ser evaluada como candidato para estudios futuros de una vacuna contra esta patógeno.

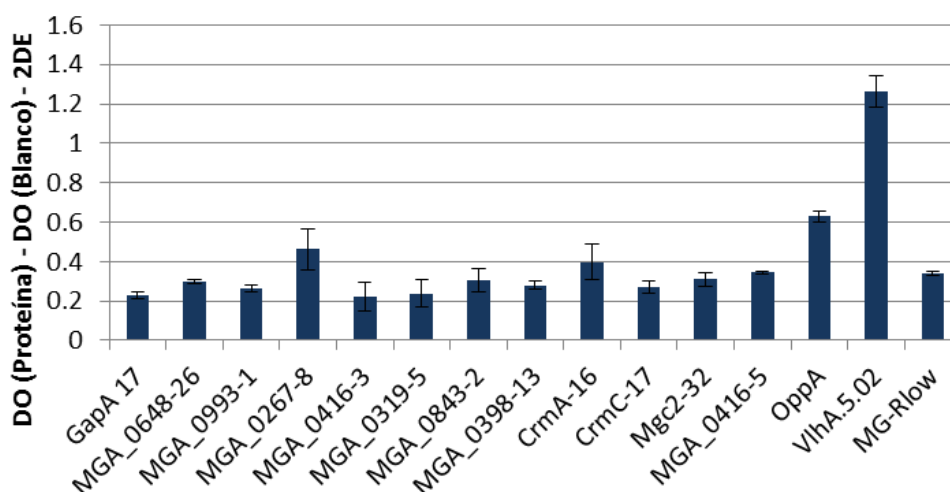


**Figura 13.** ELISA con un anticuerpo contra proteínas completas de la cepa R(low) de *M. gallisepticum* para evaluar antigenicidad de los productos recombinantes. Donde DO es densidad óptica y DE es la desviación estándar del blanco

### 3.3.7 Evaluación preliminar de la localización en la superficie bacteriana de las proteínas de *M. gallisepticum* en estudio

En la Figura 14 se muestran los resultados de la evaluación por ELISA de la capacidad del anticuerpo contra proteínas de membrana de reconocer a las proteínas recombinantes, con el objetivo de confirmar la validez de las predicciones computacionales en cuanto a la localización celular de las proteínas en estudio.

Como se puede observar el antisuero policlonal contra proteínas de membrana reconoció a las catorce proteínas, en algunos casos con una señal incluso superior a la del antígeno completo de *M. gallisepticum*, lo que demuestra que todas las proteínas que fueron evaluadas en esta etapa de la investigación, se localizan en la membrana del patógeno y son inmunogénicas. No obstante, para confirmar el resultado sería recomendable llevar a cabo otros ensayos que pasan por la generación de anticuerpos mono-específicos contra cada una de las proteínas recombinantes obtenidas. En un caso con los anticuerpos obtenidos se evaluaría el tratamiento con tripsina de células completas de la cepa R(low) de *M. gallisepticum* y, en el otro, la identificación con estos mismos anticuerpos, de las proteínas nativas en la fracción de Triton X-114 luego de la extracción de proteínas de membrana de la cepa R(low) de *M. gallisepticum*.



**Figura 14.** ELISA con suero policlonal contra proteínas de membrana de la cepa R(low) de *M. gallisepticum* para evaluar la localización en la superficie bacteriana de las proteínas en estudio.

El hecho de que en este ELISA se reconozcan los 14 productos recombinantes (Figura 14), mientras que en el ELISA con el antisuero contra el antígeno completo esto no suceda (Figura 13), puede explicarse porque en el inmunógeno empleado para obtener este último antisuero estaban presentes moléculas no localizadas en la superficie bacteriana, que diluyen la respuesta inmune contra las proteínas de membrana externa. Esto provoca que los niveles de anticuerpos contra las proteínas de superficie en este suero sean inferiores a los niveles en el suero contra proteínas de membrana y, por ende, que exista la posibilidad de que en el mismo haya proteínas que no se puedan detectar.

En el resultado es significativo el hecho de que casi todas las proteínas generaron una señal superior a la de GapA. La incorporación de esta proteína a la vacuna ts-11, incrementó la inmunogenicidad del preparado vacunal, así como los niveles de protección contra la enfermedad, de los pollos vacunados (Shil et al., 2011b, Shil et al., 2011a). Este comportamiento convierte al resto de las proteínas en estudio en blancos para las investigaciones en el campo del desarrollo de vacunas contra este microorganismo.

En sentido general los resultados de este estudio demostraron, de modo experimental, que la metodología computacional empleada en este trabajo para la identificación de antígenos de superficie bacterianos, permite la predicción de los mismos, al menos en el caso de los micoplasmas. Además, permitió la identificación de varios blancos para candidatos vacunales.

### **3.3.8 Secuenciación de los segmentos clonados**

Para confirmar la identidad a nivel nucleotídico de los insertos presentes en las construcciones que codifican para los genes que demostraron ser antígenos de superficie de *M. gallisepticum* y verificar a este nivel la especificidad de los resultados obtenidos, se secuenciaron las mismas.

Los resultados del alineamiento de las secuencias de los insertos con los genes de origen respectivos se muestran en el anexo 7. Las secuencias de los insertos de las construcciones

MGA\_0993-1 y pQE30-vlhA.5.02 no presentaron ninguna similitud con las secuencias de los genes. Esto puede ser debido, fundamentalmente, o a errores en el proceso de clonaje, o a impurezas presentes en la muestra para secuenciar. Es necesario volver a secuenciar estos clones para confirmar estos resultados. El tamaño de las regiones a secuenciar de las construcciones pQE30-MGA\_0267-8 y pQE30-crmA-16, impidió secuenciar estos insertos completamente, pero hubo coincidencia en más del 75% de los mismos, incluyendo el inicio y el final de los mismos, lo que, unido al resultado de las pruebas inmunoquímicas, ofrece confianza acerca de la identidad de los genes clonados. El resto de los genes clonados posee la secuencia correcta. Esto valida los resultados previos obtenidos.

### **3.3.9 Identificación de potenciales genes de proteínas de superficie en aislamiento cubano de *M. gallisepticum***

Sin dudas, poder identificar potenciales antígenos de superficie en aislamientos cubanos de *M. gallisepticum*, constituye una enorme ventaja al diseñar una vacuna para el control de la micoplasmosis aviar en Cuba. Si se tiene en cuenta que el empleo de cepas endémicas ha probado ser exitoso para el control de numerosas enfermedades (Qian et al., 2015, Ladhani et al., 2016, Savard et al., 2016) y que las diferencias entre cepas vacunales y de campo provoca brotes de enfermedades incluso después de la vacunación (Gildea et al., 2013, Aquino et al., 2014, Amen et al., 2015, Kenny et al., 2016), el empleo de antígenos presentes en las cepas circulantes en nuestro país incrementaría la eficacia de un potencial producto vacunal. Como se puede apreciar en la Tabla 10, a excepción de nueve, la casi totalidad de los genes evaluados, está presente en el aislamiento de campo de *M. gallisepticum* obtenido en la provincia de Mayabeque. Adicionalmente, todos los genes que codifican para los productos recombinantes obtenidos en este trabajo, están presentes en el aislamiento cubano, lo cual podría tener un impacto positivo en estudios futuros para el desarrollo de una vacuna, al contar con estas herramientas moleculares.



**Tabla 10.** Resultados de la amplificación por PCR, en el aislamiento cubano, de los genes identificados *in silico*, como potenciales genes de proteínas de *M. gallisepticum*

Gen	Resultado	Gen	Resultado	Gen	Resultado	Gen	Resultado
<i>crmA</i>	+	MGA_0267	+	MGA_0557	+	MGA_0878	+
<i>crmC</i>	+	MGA_0274	+	MGA_0558	-	MGA_0879	+
<i>gapA</i>	+	MGA_0280	-	MGA_0565	+	MGA_0913	+
MGA_0008	+	MGA_0281	+	MGA_0566	+	MGA_0954	+
MGA_0021	+	MGA_0298	+	MGA_0588	+	MGA_0956	+
MGA_0023	+	MGA_0310	+	MGA_0600	+	MGA_0957	+
MGA_0046	+	MGA_0319	+	MGA_0648	+	MGA_0993	+
MGA_0049	+	MGA_0323	+	MGA_0656	+	MGA_1010	+
MGA_0050	+	MGA_0335	+	MGA_0674	+	MGA_1014	+
MGA_0063	+	MGA_0337	+	MGA_0676	+	MGA_1037	+
MGA_0082	+	MGA_0398	+	MGA_0702	+	MGA_1079	+
MGA_0084	+	MGA_0403	+	MGA_0704	+	MGA_1085	-
MGA_0097	+	MGA_0416	+	MGA_0760	+	MGA_1087	+
MGA_0103	+	MGA_0436	+	MGA_0789	+	MGA_1122	+
MGA_0111	+	MGA_0472	+	MGA_0797	+	MGA_1161	+
MGA_0115	+	MGA_0477	+	MGA_0801	+	MGA_1191	+
MGA_0119	+	MGA_0485	+	MGA_0808	+	MGA_1215	+
MGA_0126	+	MGA_0487	+	MGA_0809	+	MGA_1216	+
MGA_0173	-	MGA_0512	+	MGA_0810	+	MGA_1315	-
MGA_0211	-	MGA_0549	-	MGA_0812	+	MGA_1318	+
MGA_0213	+	MGA_0551	+	MGA_0830	+	MGA_1357	+
MGA_0217	+	MGA_0552	-	MGA_0843	+	<i>mgc2</i>	+
MGA_0226	+	MGA_0554	+	MGA_0849	+	<i>pvpAa</i>	+
MGA_0241	+	MGA_0556	+	MGA_0867	+	<i>vlhA.5.02</i>	+
MGA_0250	-						

genes que no se amplificaron, genes de conocidos factores de virulencia, genes de las proteínas que fueron obtenidas por vía recombinante.

Si se tienen en cuenta los marcados síntomas clínicos de los animales infectados a partir de los cuales se obtuvo el aislamiento, es de esperar que este se corresponda con una cepa virulenta. De confirmarse este carácter patógeno, y se contara con otros aislamientos de diferentes regiones del país, las diferencias observadas ayudarían a reducir el conjunto de potenciales candidatos vacunales contra *M. gallisepticum*. La conservación de los potenciales antígenos vacunales entre diferentes cepas, es deseada, ya que la misma incrementa el diapasón de cepas contra las cuales el producto vacunal final podría inducir protección (Moriel et al., 2010, Palmer et al., 2011, Nesta et al., 2012, Xiang and He, 2013, Talukdar et al., 2014, de Alvarenga Mudadu et al., 2015)

La detección de la presencia de estos 88 genes codificantes para potenciales proteínas de membrana de *M. gallisepticum*, es la primera aplicación práctica de todo este trabajo, y remarca la importancia de la aplicación de la VR al campo de la vacunología, en el contexto

del siglo XXI. En este caso, incluso sin contar con la secuencia genómica de los aislamientos de campo, es posible, basado en estos estudios preliminares, identificar la presencia de potenciales candidatos vacunales en las cepas circulantes en el país.

### **Conclusiones parciales**

1. El SDS es efectivo para solubilizar las proteínas recombinantes, y llevar a cabo la purificación de las mismas en condiciones nativas.
2. Los productos de los genes, MGA\_0648, MGA\_0993, MGA\_0267, MGA\_0416, MGA\_0319, MGA\_0843 y MGA\_0398, codifican para proteínas inmunogénicas de superficie de *M. gallisepticum* y los productos recombinantes obtenidos a partir de estos genes son antigénicos.
3. Por su potencial inmunogénico, la proteína MGA\_0267 es de primera elección para diseñar un candidato vacunal.

## 4 CAPÍTULO 4. DISCUSIÓN GENERAL

El control de la micoplasmosis aviar, a pesar de los importantes avances que se han realizado, se mantiene como una tarea sin resolver ya que los brotes de la enfermedad aún ocurren a frecuencias más altas que las aceptables (Gaunson et al., 2006). El tratamiento con antibióticos ha perdido eficacia debido al aumento de los niveles de resistencia de las cepas circulantes (Li et al., 2010, Gerchman et al., 2011), además del creciente rechazo al empleo de antibióticos en animales destinados al consumo humano por motivos relacionados con un incremento general de la resistencia antimicrobiana. La vacunación es la manera más eficaz y sustentable de prevenir las enfermedades infecciosas (Yamanishi, 2011). Las vacunas contra *M. gallisepticum*, si bien disminuyen los síntomas clínicos y las pérdidas económicas, no tienen la efectividad deseada. Las bacterinas, aunque son seguras, tiene como desventajas un alto costo y la necesidad de vacunar individualmente cada ave, además no impiden la colonización de las cepas de campo y no previenen completamente contra los daños (Jacob et al., 2014). Las vacunas vivas atenuadas, por otra parte, plantean la interrogante de la reversión de la virulencia, como se ha detectado en una vacuna basada en la cepa ts-11 de *M. gallisepticum*, considerada hasta el momento segura (El Gazzar et al., 2011, Armour and Ferguson-Noel, 2015). La obtención entonces de vacunas de nueva generación, efectivas en el control y la erradicación de este patógeno se convierte en una necesidad y una prioridad.

El diseño y desarrollo de la nueva generación de vacunas se ha centrado en las proteínas individuales o el ADN codificante para solo una o un número limitado de proteínas. La cruzada para el descubrimiento de estos antígenos, es la piedra angular en el desarrollo de nuevas vacunas de subunidades (Bagnoli et al., 2011).

El primer intento realizado en esta dirección, fue una vacuna basada en el empleo del virus de la viruela como vector y CrmA como antígeno vacunal (Zhang et al., 2010), que no aportó los resultados esperados y los animales vacunados no mostraron diferencias significativas con el grupo control (Ferguson-Noel et al., 2012). El hecho de que CrmA sea una proteína

de importante participación en la patogenia de *M. gallisepticum* (Papazisi et al., 2002, Indikova et al., 2014), apoya la idea de que el empleo de una sola proteína no es suficiente para lograr una vacuna eficaz contra este patógeno, idea sustentada sobre todo por la frecuente incidencia de la variación antigénica de la arquitectura de superficie de esta bacteria (Grodio et al., 2012, Tulman et al., 2012, Pflaum et al., 2015). Identificar nuevos factores de virulencia de *M. gallisepticum* puede ser la clave para desarrollar una vacuna eficaz contra este microorganismo. La adherencia de este microorganismo a las células epiteliales respiratorias del hospedero es un proceso indispensable para el establecimiento de la infección (Papazisi et al., 2002) y se realiza mediante lipoproteínas localizadas en la superficie de la bacteria. Aunque se han identificado algunas de las proteínas involucradas en la adherencia como GapA, CrmA (Goh et al., 1998, Papazisi et al., 2002, Mudahi-Orenstein et al., 2003) y PvpA (Yogev et al., 1994), los datos que emergen de estudios realizados, indican que algunas citadhesinas o moléculas accesorias relacionadas, están aún por ser identificadas, definidas y caracterizadas (Szczepanek et al., 2010b, Tulman et al., 2012).

La VR nos da la posibilidad de acelerar rápidamente la selección de antígenos candidatos para pruebas de inmunogenicidad y la formulación de vacunas (Seib et al., 2012, Goodswen et al., 2014, Heinson et al., 2015).

Como parte de la búsqueda de moléculas potencialmente involucradas en los mecanismos de patogenidad de *M. gallisepticum*, que podrían ser usadas como dianas en estudios para evaluar candidatos diagnósticos y vacunales eficaces, se empleó una herramienta bioinformática desarrollada sobre la base del enfoque de la VR, que prioriza la identificación de algunos de los principales factores de virulencia bacterianos, las lipoproteínas y/o citadhesinas. La publicación de la secuencia genómica de *M. gallisepticum* (Papazisi et al., 2003) fue la base para el desarrollo de este trabajo.

Cuando se evaluó la capacidad del sistema de identificar PVCs bacterianos y moléculas de superficie de micoplasmas, evaluación que incluyó la comparación, en ambos casos, con los mejores programas existentes para estos fines, se observó que el sistema, si bien en

general tenía un desempeño similar y en algunos casos superior al resto, en cuanto a su especificidad, cedía ante todos los predictores de PVCs y ante LocTree3. Si se tiene en cuenta que el escenario ideal en vacunología sería identificar las dianas vacunales de una manera tan eficiente y específica, que se pueda ir directamente de la identificación de los antígenos a partir del genoma hasta la clínica, sin pasar por el laboratorio (Altindis et al., 2015), pues se vuelve incuestionable la necesidad de implementar nuevas y más selectivas estrategias para reducir el volumen de predicciones positivas del sistema evaluado en este trabajo, mediante el aumento de su especificidad. Esto se podría lograr mediante la adición al algoritmo de predicción de Bact-VAP, de búsqueda de patrones o señales específicas de antígenos protectivos que Altindis y colaboradores implementaron exitosamente (Altindis et al., 2015), así como otros criterios. La incorporación de la predicción de epítomos T es uno de ellos. Si bien la predicción de epítomos B ha mostrado un gran potencial para identificar antígenos de superficie bacterianos (Giacco et al., 2012, Singh et al., 2013), la búsqueda de epítomos de células T es más adecuada para patógenos intracelulares, tales como parásitos, virus (a pesar de que la búsqueda de vacunas virales se ha centrado históricamente en la inducción de anticuerpos neutralizantes) (Lundegaard et al., 2010, Patronov and Doytchinova, 2013, Sun et al., 2015), o incluso bacterias con el potencial de invadir las células del hospedero, como es el caso particular de los micoplasmas (Jurkovic et al., 2012, Buim et al., 2011, Furnkranz et al., 2013). De esta manera, las vacunas contra bacterias intracelulares, virus hipervariables y parásitos complejos, probablemente necesitarán incluir antígenos de tipo tanto B como T (Schussek et al., 2014). Por último, la identificación de proteínas únicas entre las cepas patógenas de cada especie y que no posean similitud con las proteínas del hospedero (Schussek et al., 2014), será también clave en la obtención de un listado final filtrado con solo los candidatos más promisorios.

El algoritmo de Bact-VAP que se evaluó en este trabajo, resultó muy conveniente para su aplicación a micoplasmas, ya que las lipoproteínas y en particular las citadhesinas, cuya predicción está incluida en el esquema de Bact-VAP, tienen un papel vital en la patogenia de estas bacterias (Szczepanek et al., 2010a, Furnkranz et al., 2013, Pflaum et al., 2015), lo que convierte a estas moléculas en uno de los blancos más atractivos en los estudios para

el desarrollo de una vacuna contra estos microorganismos (van der Merwe et al., 2011, Ferguson-Noel et al., 2012, Lee et al., 2014).

Nunca antes se había descrito la utilización de herramientas computacionales con el objetivo de identificar potenciales antígenos protectivos de *M. gallisepticum*. En esta primera aproximación para la aplicación de la VR en el análisis de este microorganismo, se identificaron 291 PVCs, de los cuales 210 eran proteínas no descritas previamente, lo que crea un importante espacio para la búsqueda de efectivos candidatos vacunales contra *M. gallisepticum*. La fracción que representan estas 291 proteínas con respecto al proteoma completo, fue algo superior a la media obtenida cuando se aplican procedimientos de VR a genomas bacterianos. Las limitaciones que se mencionaron anteriormente en cuanto a la especificidad del sistema, pudieron haber sido la causa de este resultado, aunque otra posible explicación puede ser la pequeña talla del genoma de los micoplasmas y las carencias que esto conlleva, carencias que se suplen con una marcada interacción con el entorno, lo que provoca un aumento en la tasa de genes codificantes para mediadores de esta interacción, localizados, en su inmensa mayoría, en la superficie de la bacteria (Browning et al., 2011).

Teniendo en cuenta la importancia de identificar *M. gallisepticum* nuevos potenciales PVCs, se decidió trabajar en la obtención de las primeras 50 proteínas en el listado de priorización que se generó luego del procesamiento del proteoma de la cepa R(Low) de este patógeno, con el algoritmo de predicción evaluado, además de seis conocidos factores de virulencia de este microorganismo. El manejo de un elevado volumen de muestras es característico de las tecnologías de alto flujo vinculadas a la bioinformática y, en particular, a la VR lo que ha obligado a los investigadores a generar variantes que simplifiquen el trabajo y permitan, procesar eficientemente gran cantidad de datos o muestras, mediante la reducción racional del número de variables a analizar y en lo posible, la automatización del proceso (Meng et al., 2016). En este trabajo, para generar los productos recombinantes primero se diseñó el clonaje de modo tal que siempre que fuera posible, se emplearan las mismas enzimas de restricción para el clonaje de todos los genes y, posteriormente, un único programa para

amplificarlos por PCR. Para esto fue necesario trabajar con una baja temperatura de hibridación, lo que trajo por consecuencia la aparición de productos inespecíficos. Finalmente, de 56 genes en estudio, se lograron generar 30 construcciones recombinantes y expresión en 17 de ellas, correspondientes a 13 genes, de los cuales sólo uno no tenía la talla esperada. Todos estos genes fueron experimentalmente identificados como antígenos de superficie de *M. gallisepticum*.

Para evaluar el impacto de la generación de productos inespecíficos en el rendimiento del proceso de clonaje es necesario comparar los resultados con resultados previos en los que se empleó un procedimiento similar. Para *Neisseria meningitidis* se identificaron y estudiaron 600 genes, de los cuales se expresaron 350 y solo los productos de 85 de ellos se identificaron como antígenos de superficie (Pizza et al., 2000). Otro ejemplo es *Streptococcus pneumoniae*, se identificaron 130 genes, se expresaron 108 y finalmente se logró protección con seis de las proteínas obtenidas (Wizemann et al., 2001, Maione et al., 2005). Por último *Porphyromonas gingivalis*, de 120 proteínas en estudio, solo 40 resultaron ser antígenos de superficie (Wizemann et al., 2001). Como se puede apreciar, en todos los casos, el número de moléculas que alcanzan el final del proceso es muy inferior a los genes identificados como positivos por los análisis *in silico* y adicionalmente, si se tiene en cuenta las limitaciones en la expresión de genes de micoplasmas en *E. coli* por el uso de codones particular de estos últimos, pues los resultados obtenidos en este trabajo se encuentran en el rango esperado. Por esta razón, podemos concluir que las reacciones inespecíficas observadas en los PCR de amplificación de los genes para el clonaje, no afectaron de manera significativa los resultados. Una opción ideal para la expresión de genes de micoplasmas en *E coli* y usada cada vez con más frecuencia para estos fines (Perez-Casal et al., 2015), es el empleo de genes sintéticos, los cuales están optimizados para la expresión en el hospedero seleccionado. Con esta variante se garantizan altos niveles de expresión de todos los genes, además de reducir significativamente el tiempo de investigación.

La identificación de siete nuevos antígenos de superficie de *M. gallisepticum* es un resultado prometedor para el desarrollo de una vacuna efectiva contra este microorganismo, cuyo formato debe ser cuidadosamente estudiado. Ciertamente las vacunas por subunidades existentes en el mercado generalmente confieren niveles de protección inferiores a los de los organismos vivos (Meeusen et al., 2007), pero los nuevos sistemas de administración vacunal abren nuevas y atractivas posibilidades para el empleo de antígenos con potencial protectorio. Como ejemplos de tales sistemas se pueden mencionar los plásmidos de ADN (Li and Petrovsky, 2016, Meunier et al., 2016), liposomas (Schwendener, 2014, Perrie et al., 2015), nano- o micropartículas (Caetano et al., 2014, San Roman et al., 2014, Ahmed and Aljaeid, 2016, Yang et al., 2016) y vectores vivos (Lin et al., 2015, Xiang et al., 2015).

Todo lo expuesto nos permite afirmar que el presente trabajo representa un importante aporte al desarrollo de vacunas contra la micoplasmosis aviar, con una serie de proteínas con potencial para ser empleadas como blancos vacunales, así como una lista de otros candidatos, que deben ser evaluados. Además, ofrece una metodología específica y sensible, ya implementada, de predicción de potenciales candidatos de vacunales bacterianos.



## CONCLUSIONES

1. El sistema Bact-VAP es efectivo y útil para predecir PVCs bacterianos y antígenos de superficie de micoplasmas.
2. Los productos de los genes, MGA\_0648, MGA\_0993, MGA\_0267, MGA\_0416, MGA\_0319, MGA\_0843 y MGA\_0398, codifican para proteínas inmunogénicas de superficie de *M. gallisepticum* y los productos recombinantes obtenidos son antigénicos.
3. Por su potencial inmunogénico, la proteína MGA\_0267 es de primera elección para diseñar un candidato vacunal.

## RECOMENDACIONES

1. Confirmar los resultados de localización celular empleando otros métodos de evaluación y determinar el papel de las moléculas obtenidas, en la patogenia de la enfermedad.
2. Con los candidatos más promisorios, hacer estudios, primero de inmunogenicidad, y luego de protección contra este microorganismo.
3. Obtener otras proteínas presentes en el listado de priorización, pero empleando genes sintéticos para la obtención de los nuevos productos.
4. Incorporar al proceso predictivo la determinación de proteínas secretadas, la identificación de epítopos T y la determinación de similitud con proteínas de otras cepas de la misma especie y con proteínas del hospedero. Además, introducir más de un programa para cada tipo de predicción.

## PRODUCCION CIENTÍFICA DEL AUTOR RELACIONADA CON LA TESIS

### Publicaciones:

Agüero, J.A.; y col. Implementación del algoritmo de predicción Freeman-Wimley en una aplicación web para la identificación *in silico* de proteínas de membrana barriles-beta. Vaccimonitor. 24(2): 71-8.

Agüero, J.A. y col. (2015). Molecular cloning and expression of a computationally predicted surface antigen of *Mycoplasma gallisepticum*. Rev. Salud Anim. Vol. 37 No. 1 (ene.-abr.): 21-26.

Agüero, J.A. y col. (2011). Expresión *in vitro* en *Escherichia coli*, de un fragmento del gen *vlha.5.02* de la principal familia de hemaglutininas de *Mycoplasma gallisepticum*. Rev. Salud Anim. 33. 1.

### Eventos:

#### Carteles

3rd International Congress on Pharmacology of Vaccines (IMMUNOPHARMACOLOGY - VACCIPHARMA 2015). Varadero. Cuba. Coautor. Metaserver for *in silico* identification of bacterial targets with diagnostic and/or vaccine value.

2nd International Congress on Pharmacology of Vaccines. (VACCIPHARMA 2012). Santa María Cay. Cuba. Author. Genome-scale prediction of surface-associated proteins in mycoplasmas of veterinary interest.

2nd International Congress on Pharmacology of Vaccines. (VACCIPHARMA 2012). Santa María Cay. Cuba. Author. Implementation of SignalP 3.0 hidden markov models in a web-based application, for *in silico* identification of signal peptide.

Seminario Internacional de Sanidad Agropecuaria, SISA (2011) (La Habana, Cuba). Application of Freeman-Wimley prediction algorithm to computational identification of beta-barrel membrane proteins in mycoplasma predicted proteomes: comparison with another method.

18th International Congress of The International Organization for Mycoplasmology (IOM), (2010) (Italy). *In silico* prediction and obtainment of putative vaccine and diagnostic targets in *Mycoplasma gallisepticum*.

16th International Congress of The International Organization for Mycoplasmology (IOM). (2006). UK. Author. Computational identification of outer membrane proteins in *Mycoplasma gallisepticum* and *Mycoplasma synoviae* predicted proteomes

#### *Conferencias*

3<sup>rd</sup> International Congress on Pharmacology of Vaccines (IMMUNOPHARMACOLOGY - VACCIPHARMA 2015). Varadero. Cuba. Author. Molecular cloning and expression of gene fragments from computationally predicted *Mycoplasma gallisepticum* surface antigenic proteins.

2<sup>nd</sup> International Congress on Pharmacology of Vaccines. (VACCIPHARMA 2012). (2012). Santa María Cay. Cuba. Computational identification of *Mycoplasma gallisepticum* membrane-associated proteins using an in house developed pipeline.

Seminario Internacional de Sanidad Agropecuaria, SISA (2011). (La Habana, Cuba). Computer aided identification of potential antibacterial vaccine targets.

First International Congress on Pharmacology of Vaccines (VacciPharma). (2009). Cuba. *In silico* prioritization of vaccine and diagnostic targets in *Mycoplasma gallisepticum*.

## REFERENCIAS BIBLIOGRÁFICAS

- ABU-AMERO, K. K., ABU-GROUN, E. A., HALABLAB, M. A. & MILES, R. J. 2000. Kinetics and distribution of alcohol oxidising activity in *Acholeplasma* and *Mycoplasma* species. *FEMS Microbiol Lett*, 183, 147-51.
- ADAN-KUBO, J., YOSHII, S. H., KONO, H. & MIYATA, M. 2012. Molecular structure of isolated Mvspl, a variable surface protein of fish pathogen *Mycoplasma mobile*. *J Bacteriol.*, 194, 3050-7.
- AGHCHEH, R. K. & KUBICEK, C. P. 2015. Epigenetics as an emerging tool for improvement of fungal strains used in biotechnology. *Appl Microbiol Biotechnol*, 99, 6167-81.
- AGRAWAL, A. & HUANG, X. 2009. PSIBLAST\_PairwiseStatSig: reordering PSI-BLAST hits using pairwise statistical significance. *Bioinformatics*, 25, 1082-3.
- AGÜERO, J. A., SÁNCHEZ, O., BARRERA, M. & TOLEDO, J. R. 2008. Molecular cloning and expression of a fragment of the gene codifying for the protein Erns of the Classical Swine Fever Virus. *Rev. Salud Anim.*, 30, 85-92.
- AHMED, T. A. & ALJAEID, B. M. 2016. Preparation, characterization, and potential application of chitosan, chitosan derivatives, and chitosan metal nanoparticles in pharmaceutical drug delivery. *Drug Des Devel Ther*, 10, 483-507.
- ALGIRE, M. A., MONTAGUE, M. G., VASHEE, S., LARTIGUE, C. & MERRYMAN, C. 2012. A Type III restriction-modification system in *Mycoplasma mycoides* subsp. capri. *Open Biol*, 2, 120115.
- ALLAM, A. B., REYES, L., ASSAD-GARCIA, N., GLASS, J. I. & BROWN, M. B. 2010. Targeted homologous recombination in *Mycoplasma mycoides* subsp. capri is enhanced by inclusion of heterologous recA. *Appl Environ Microbiol*.
- ALMAGOR, M., KAHANE, I., WIESEL, J. M. & YATZIV, S. 1985. Human ciliated epithelial cells from nasal polyps as an experimental model for *Mycoplasma pneumoniae* infection. *Infect Immun*, 48, 552-5.
- ALMEIDA, R. A., LUTHER, D. A., PATEL, D. & OLIVER, S. P. 2011. Predicted antigenic regions of *Streptococcus uberis* adhesion molecule (SUAM) are involved in adherence to and internalization into mammary epithelial cells. *Vet Microbiol*, 148, 323-8.
- ALTINDIS, E., COZZI, R., DI PALO, B., NECCHI, F., MISHRA, R. P., FONTANA, M. R., SORIANI, M., BAGNOLI, F., MAIONE, D., GRANDI, G. & LIBERATORI, S. 2015. Protectome analysis: a new selective bioinformatics tool for bacterial vaccine candidate discovery. *Mol Cell Proteomics*, 14, 418-29.
- ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol*, 215, 403-10.
- ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25, 3389-402.
- AMEN, O., VEMULA, S. V., ZHAO, J., IBRAHIM, R., HUSSEIN, A., HEWLETT, I. K., MOUSSA, S. & MITTAL, S. K. 2015. Identification and characterization of a highly pathogenic H5N1 avian influenza A virus during an outbreak in vaccinated chickens in Egypt. *Virus Res*, 210, 337-43.
- AQUINO, T. L., BRICE, G. T., HAYES, S., MYERS, C. A., MCDOWELL, J., WHITE, B., GARTEN, R. & JOHNSTON, D. 2014. Influenza outbreak in a vaccinated population--USS Ardent, February 2014. *MMWR Morb Mortal Wkly Rep*, 63, 947-9.
- ARMOUR, N. K. & FERGUSON-NOEL, N. 2015. Evaluation of the egg transmission and pathogenicity of *Mycoplasma gallisepticum* isolates genotyped as ts-11. *Avian Pathol*, 1-24.
- AYALEW, S., CONFER, A. W., HARTSON, S. D. & SHRESTHA, B. 2010. Immunoproteomic analyses of outer membrane proteins of *Mannheimia haemolytica* and identification of potential vaccine candidates. *Proteomics*, 10, 2151-64.
- BAGNOLI, F., BAUDNER, B., MISHRA, R. P., BARTOLINI, E., FIASCHI, L., MARIOTTI, P., NARDI-DEI, V., BOUCHER, P. & RAPPUOLI, R. 2011. Designing the Next Generation of Vaccines for Global Public Health. *OMICS*, 15, 545-66.
- BAGOS, P. G., TSIRIGOS, K. D., LIAKOPOULOS, T. D. & HAMODRAKAS, S. J. 2008. Prediction of lipoprotein signal peptides in Gram-positive bacteria with a Hidden Markov Model. *J Proteome Res*, 7, 5082-93.
- BALASUBRAMANIAN, S., KANNAN, T. R. & BASEMAN, J. B. 2008. The surface-exposed carboxyl region of *Mycoplasma pneumoniae* elongation factor Tu interacts with fibronectin. *Infect Immun*, 76, 3116-23.
- BALISH, M. F., ROSS, S. M., FISSEHA, M. & KRAUSE, D. C. 2003. Deletion analysis identifies key functional domains of the cytoadherence-associated protein HMW2 of *Mycoplasma pneumoniae*. *Mol Microbiol*, 50, 1507-16.

- BARLOW, D. J., EDWARDS, M. S. & THORNTON, J. M. 1986. Continuous and discontinuous protein antigenic determinants. *Nature*, 322, 747-8.
- BEARSON, S. M., COLLIER, S. D., BEARSON, B. L. & BRANTON, S. L. 2003. Induction of a mycoplasma gallisepticum pMGA gene in the chicken tracheal ring organ culture model. *Avian Dis*, 47, 745-9.
- BENCINA, D. 2002. Haemagglutinins of pathogenic avian mycoplasmas. *Avian Pathol*, 31, 535-47.
- BERVEN, F. S., FLIKKA, K., JENSEN, H. B. & EIDHAMMER, I. 2004. BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. *Nucleic Acids Res*, 32, W394-9.
- BIRD, D. W., GRABER, K., KNUTSON, A. & WASHBURN, L. R. 2008. Mutation of two *Mycoplasma arthritidis* surface lipoproteins with divergent functions in cytodherence. *Infect Immun*, 76, 5768-76.
- BLACKBURN, M. C., PETROVA, E., CORREIA, B. E. & MAERKL, S. J. 2016. Integrating gene synthesis and microfluidic protein analysis for rapid protein engineering. *Nucleic Acids Res*, 44, e68.
- BLYTHE, M. J. & FLOWER, D. R. 2005. Benchmarking B cell epitope prediction: underperformance of existing methods. *Protein Sci*, 14, 246-8.
- BOGEMA, D. R., DEUTSCHER, A. T., WOOLLEY, L. K., SEYMOUR, L. M., RAYMOND, B. B., TACCHI, J. L., PADULA, M. P., DIXON, N. E., MINION, F. C., JENKINS, C., WALKER, M. J. & DJORDJEVIC, S. P. 2012. Characterization of Cleavage Events in the Multifunctional Cilium Adhesin Mhp684 (P146) Reveals a Mechanism by Which Mycoplasma hyopneumoniae Regulates Surface Topography. *MBio*, 3, e00282-11.
- BOGEMA, D. R., SCOTT, N. E., PADULA, M. P., TACCHI, J. L., RAYMOND, B. B., JENKINS, C., CORDWELL, S. J., MINION, F. C., WALKER, M. J. & DJORDJEVIC, S. P. 2011. The sequence TTKF | QE defines the site of proteolytic cleavage in Mhp683, a novel glycosaminoglycan and cilium adhesin of Mycoplasma hyopneumoniae. *J Biol Chem*, 286, 41217-29.
- BOGUSLAVSKY, S., MENAKER, D., LYSNYANSKY, I., LIU, T., LEVISOHN, S., ROSENGARTEN, R., GARCIA, M. & YOGEV, D. 2000. Molecular characterization of the Mycoplasma gallisepticum pvpA gene which encodes a putative variable cytoadhesin protein. *Infect Immun*, 68, 3956-64.
- BOLLAND, J. R. & DYBVIG, K. 2012. Mycoplasma pulmonis Vsa proteins and polysaccharide modulate adherence to pulmonary epithelial cells. *FEMS Microbiol Lett*, 331, 25-30.
- BOND, K. B., SRIWANTHANA, B., HODGE, T. W., DE GROOT, A. S., MASTRO, T. D., YOUNG, N. L., PROMADEJ, N., ALTMAN, J. D., LIMPAKARNJANARAT, K. & MCNICHOLL, J. M. 2001. An HLA-directed molecular and bioinformatics approach identifies new HLA-A11 HIV-1 subtype E cytotoxic T lymphocyte epitopes in HIV-1-infected Thais. *AIDS Res Hum Retroviruses*, 17, 703-17.
- BOWMAN, B. N., MCADAM, P. R., VIVONA, S., ZHANG, J. X., LUONG, T., BELEW, R. K., SAHOTA, H., GUINEY, D., VALAFAR, F., FIERER, J. & WOELK, C. H. 2011. Improving reverse vaccinology with a machine learning approach. *Vaccine*, 29, 8156-64.
- BROWNING, G. F., MAREDA, M. S., NOORMOHAMMADI, A. H. & MARKHAM, P. F. 2011. The central role of lipoproteins in the pathogenesis of mycoplasmoses. *Vet Microbiol*, 153, 44-50.
- BRUNO, L., CORTESE, M., RAPPUOLI, R. & MEROLA, M. 2015. Lessons from Reverse Vaccinology for viral vaccine design. *Curr Opin Virol*, 11, 89-97.
- BUIM, M. R., BUZINHANI, M., YAMAGUTI, M., OLIVEIRA, R. C., METTIFOGO, E., UENO, P. M., TIMENETSKY, J., SANTELLI, G. M. & PIANTINO FERREIRA, A. J. 2011. Mycoplasma synoviae cell invasion: Elucidation of the Mycoplasma pathogenesis in chicken. *Comp Immunol Microbiol Infect Dis*, 34, 41-7.
- BUMGARDNER, E., BEY, R. F., KITTICHOTIRAT, W., BUMGARNER, R. E. & LAWRENCE, P. K. 2014. Genome Sequences of Seven Mycoplasma hyosynoviae Strains Isolated from the Joint Tissue of Infected Swine (Sus scrofa). *Genome Announc*, 2, e00552-14.
- BURGOS, R., PICH, O. Q., FERRER-NAVARRO, M., BASEMAN, J. B., QUEROL, E. & PINOL, J. 2006. Mycoplasma genitalium P140 and P110 cytoadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development. *J Bacteriol*, 188, 8627-37.
- BURGOS, R., WOOD, G. E., YOUNG, L., GLASS, J. I. & TOTTEN, P. A. 2012. RecA mediates MgpB and MgpC phase and antigenic variation in Mycoplasma genitalium, but plays a minor role in DNA repair. *Mol Microbiol*, 85, 669-83.
- BURKI, S., GASCHEN, V., STOFFEL, M. H., STOJILJKOVIC, A., FREY, J., KUEHNI-BOGHENBOR, K. & PILO, P. 2015. Invasion and persistence of Mycoplasma bovis in embryonic calf turbinate cells. *Vet Res*, 46, 53.
- BUUS, S., ROCKBERG, J., FORSSTROM, B., NILSSON, P., UHLEN, M. & SCHAFFER-NIELSEN, C. 2012. High-resolution mapping of linear antibody epitopes using ultrahigh-density peptide microarrays. *Mol Cell Proteomics*, 11, 1790-800.

- CAETANO, L. A., ALMEIDA, A. J. & GONCALVES, L. M. 2014. Approaches to tuberculosis mucosal vaccine development using nanoparticles and microparticles: a review. *J Biomed Nanotechnol*, 10, 2295-316.
- CALCUTT, M. J. & FOECKING, M. F. 2015. Comparative analysis of the *Mycoplasma capricolum* subsp. *capricolum* GM508D genome reveals subrogation of phase-variable contingency genes and a novel integrated genetic element. *Pathog Dis*, 76, ftv041.
- CAPRA, J. A., LASKOWSKI, R. A., THORNTON, J. M., SINGH, M. & FUNKHOUSER, T. A. 2009. Predicting Protein Ligand Binding Sites by Combining Evolutionary Sequence Conservation and 3D Structure. *PLoS Comput Biol*, 5, e1000585.
- CARLOS, P., ROUPIE, V., HOLBERT, S., ASCENCIO, F., HUYGEN, K., GOMEZ-ANDURO, G., BRANGER, M., REYES-BECERRIL, M. & ANGULO, C. 2015. In silico epitope analysis of unique and membrane associated proteins from mycobacterium avium subsp. paratuberculosis for immunogenicity and vaccine evaluation. *J Theor Biol*, 384, 1-9.
- CASADEVALL, A. 2003. Antibody-mediated immunity against intracellular pathogens: two-dimensional thinking comes full circle. *Infect Immun*, 71, 4225-8.
- CASADIO, R., MARTELLI, P. L. & PIERLEONI, A. 2008. The prediction of protein subcellular localization from sequence: a shortcut to functional genome annotation. *Brief Funct Genomic Proteomic*, 7, 63-73.
- CECCHINI, K. R., GORTON, T. S. & GEARY, S. J. 2007. Transcriptional responses of *Mycoplasma gallisepticum* strain R in association with eukaryotic cells. *J Bacteriol*, 189, 5803-7.
- CHANG, J. M., SU, E. C., LO, A., CHIU, H. S., SUNG, T. Y. & HSU, W. L. 2008. PSLDoc: Protein subcellular localization prediction based on gapped-dipeptides and probabilistic latent semantic analysis. *Proteins*, 72, 693-710.
- CHAUDHURI, R., ANSARI, F. A., RAGHUNANDANAN, M. V. & RAMACHANDRAN, S. 2011. FungalRV: adhesin prediction and immunoinformatics portal for human fungal pathogens. *BMC Genomics*, 12, 192.
- CHAUDHURI, R. & RAMACHANDRAN, S. 2014. Prediction of virulence factors using bioinformatics approaches. *Methods Mol Biol*, 1184, 389-400.
- CHEN, C. P., KERNYTSKY, A. & ROST, B. 2002. Transmembrane helix predictions revisited. *Protein Sci*, 11, 2774-91.
- CHEN, H., YU, S., SHEN, X., CHEN, D., QIU, X., SONG, C. & DING, C. 2011. The *Mycoplasma gallisepticum* alpha-enolase is cell surface-exposed and mediates adherence by binding to chicken plasminogen. *Microb Pathog*, 51, 285-90.
- CHEN, H., ZHAO, C., SHEN, X., CHEN, D., YU, S. & DING, C. 2010. Construction of mini-Tn4001 transposon vector for *Mycoplasma gallisepticum*. *Sci China Life Sci*, 53, 1340-5.
- CITTI, C. & BLANCHARD, A. 2013. Mycoplasmas and their host: emerging and re-emerging minimal pathogens. *Trends Microbiol*, 21, 196-203.
- CITTI, C., NOUVEL, L. X. & BARANOWSKI, E. 2010. Phase and antigenic variation in mycoplasmas. *Future Microbiol*, 5, 1073-85.
- CITTI, C. & ROSENGARTEN, R. 1997. *Mycoplasma* genetic variation and its implication for pathogenesis. *Wien Klin Wochenschr*, 109, 562-8.
- COLLIER, S., BRANTON, S., EVANS, J., LEIGH, S., WAN, X., COOKSEY, A., OLANREWAJU, H. & PHARR, G. 2015. Proteomic analysis of *Mycoplasma gallisepticum* vaccine strain F. *International Journal of Poultry Science*, 14, 1-12.
- CONFER, A. W. & AYALEW, S. 2013. The OmpA family of proteins: roles in bacterial pathogenesis and immunity. *Vet Microbiol*, 163, 207-22.
- CORREA, A. & OPPEZZO, P. 2015. Overcoming the solubility problem in *E. coli*: available approaches for recombinant protein production. *Methods Mol Biol*, 1258, 27-44.
- COSTA, T. R., FELISBERTO-RODRIGUES, C., MEIR, A., PREVOST, M. S., REDZEJ, A., TROKTER, M. & WAKSMAN, G. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol*, 13, 343-59.
- DAVIS, P. J., KATZNEL, A., RAZIN, S. & ROTTEM, S. 1985. *Spiroplasma* membrane lipids. *J Bacteriol*, 161, 118-22.
- DE ALVARENGA MUDADU, M., CARVALHO, V. & LECLERCQ, S. Y. 2015. Nonclassically Secreted Proteins as Possible Antigens for Vaccine Development: A Reverse Vaccinology Approach. *Appl Biochem Biotechnol*, 175, 3360-70.
- DE GROOT, A. S. 2006. Immunomics: discovering new targets for vaccines and therapeutics. *Drug Discov Today*, 11, 203-9.
- DE GROOT, A. S. & BERZOFISKY, J. A. 2004. From genome to vaccine--new immunoinformatics tools for vaccine design. *Methods*, 34, 425-8.
- DE GROOT, A. S., JESDALE, B. M., SZU, E., SCHAFER, J. R., CHICZ, R. M. & DEOCAMPO, G. 1997. An interactive Web site providing major histocompatibility ligand predictions: application to HIV research. *AIDS Res Hum Retroviruses*, 13, 529-31.

- DE GROOT, A. S. & MARTIN, W. 2009. Reducing risk, improving outcomes: bioengineering less immunogenic protein therapeutics. *Clin Immunol*, 131, 189-201.
- DE GROOT, A. S. & MOISE, L. 2007. New tools, new approaches and new ideas for vaccine development. *Expert Rev Vaccines*, 6, 125-7.
- DE GROOT, A. S., SBAL, H., AUBIN, C. S., MCMURRY, J. & MARTIN, W. 2002. Immuno-informatics: Mining genomes for vaccine components. *Immunol Cell Biol*, 80, 255-69.
- DEUTSCHER, A. T., JENKINS, C., MINION, F. C., SEYMOUR, L. M., PADULA, M. P., DIXON, N. E., WALKER, M. J. & DJORDJEVIC, S. P. 2010. Repeat regions R1 and R2 in the P97 paralogue Mhp271 of *Mycoplasma hyopneumoniae* bind heparin, fibronectin and porcine cilia. *Mol Microbiol*, 78, 444-58.
- DEUTSCHER, A. T., TACCHI, J. L., MINION, F. C., PADULA, M. P., CROSSETT, B., BOGEMA, D. R., JENKINS, C., KUIT, T. A., WALKER, M. J. & DJORDJEVIC, S. P. 2012. *Mycoplasma hyopneumoniae* Surface proteins Mhp385 and Mhp384 bind host cilia and glycosaminoglycans and are endoproteolytically processed by proteases that recognize different cleavage motifs. *J Proteome Res*, 11, 1924-36.
- DICKER, M. & STRASSER, R. 2015. Using glyco-engineering to produce therapeutic proteins. *Expert Opin Biol Ther*, 15, 1501-16.
- DIJKMAN, R., FEBERWEE, A. & LANDMAN, W. J. 2014. Variable lipoprotein haemagglutinin gene (*vlhA*) sequence typing of mainly dutch *Mycoplasma synoviae* isolates: comparison with *vlhA* sequences from genbank and with AFLP analysis. *Avian Pathol*, 1-23.
- DOBSON, L., REMENYI, I. & TUSNADY, G. E. 2015. CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Res*, 43, W408-12.
- DONG, Y., DEMARIA, S., SUN, X., SANTORI, F. R., JESDALE, B. M., DE GROOT, A. S., ROM, W. N. & BUSHKIN, Y. 2004. HLA-A2-restricted CD8+-cytotoxic-T-cell responses to novel epitopes in *Mycobacterium tuberculosis* superoxide dismutase, alanine dehydrogenase, and glutamine synthetase. *Infect Immun*, 72, 2412-5.
- DORDET-FRISONI, E., SAGNE, E., BARANOWSKI, E., BRETON, M., NOUVEL, L. X., BLANCHARD, A., MARENGA, M. S., TARDY, F., SIRAND-PUGNET, P. & CITTI, C. 2014. Chromosomal transfers in mycoplasmas: when minimal genomes go mobile. *MBio*, 5, e01958.
- DOS SANTOS, L. F., CLAVIJO, M. J., SREEVATSAN, S., ROVIRA, A., MOREIRA, M. A. & PIETERS, M. 2015. Genotyping of *Mycoplasma hyorhinis* using multiple-locus variable number tandem repeat analysis. *J Microbiol Methods*, 111, 87-92.
- DOYTCHINOVA, I. A. & FLOWER, D. R. 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8, 4.
- EL-MANZALAWY, Y., DOBBS, D. & HONAVAR, V. 2008. Predicting linear B-cell epitopes using string kernels. *J Mol Recognit*, 21, 243-55.
- EL GAZZAR, M., LAIBINIS, V. A. & FERGUSON-NOEL, N. 2011. Characterization of a ts-11-like *Mycoplasma gallisepticum* isolate from commercial broiler chickens. *Avian Dis*, 55, 569-74.
- EMANUELSSON, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*, 2, 953-71.
- ENSHELL-SEIJFFERS, D., DENISOV, D., GROISMAN, B., SMELYANSKI, L., MEYUHAS, R., GROSS, G., DENISOVA, G. & GERSHONI, J. M. 2003. The mapping and reconstitution of a conformational discontinuous B-cell epitope of HIV-1. *J Mol Biol*, 334, 87-101.
- EVANS, J., LEIGH, S., BRANTON, S., COLLIER, S., PHARR, G. & BEARSON, S. 2005. *Mycoplasma gallisepticum*: Current and developing means to control the avian pathogen. *The Journal of Applied Poultry Research*, 14, 757-763.
- FERGUSON-NOEL, N., COOKSON, K., LAIBINIS, V. A. & KLEVEN, S. H. 2012. The efficacy of three commercial *Mycoplasma gallisepticum* vaccines in laying hens. *Avian Dis*, 56, 272-5.
- FRASER, C. M., GOCAYNE, J. D., WHITE, O., ADAMS, M. D., CLAYTON, R. A., FLEISCHMANN, R. D., BULT, C. J., KERLAVAGE, A. R., SUTTON, G., KELLEY, J. M., FRITCHMAN, R. D., WEIDMAN, J. F., SMALL, K. V., SANDUSKY, M., FUHRMANN, J., NGUYEN, D., UTTERBACK, T. R., SAUDEK, D. M., PHILLIPS, C. A., MERRICK, J. M., TOMB, J. F., DOUGHERTY, B. A., BOTT, K. F., HU, P. C., LUCIER, T. S., PETERSON, S. N., SMITH, H. O., HUTCHISON, C. A., 3RD & VENTER, J. C. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science*, 270, 397-403.
- FREEMAN, T. C., JR. & WIMLEY, W. C. 2010. A highly accurate statistical approach for the prediction of transmembrane beta-barrels. *Bioinformatics*, 26, 1965-74.
- FURNKRANZ, U., SIEBERT-GULLE, K., ROSENGARTEN, R. & SZOSTAK, M. P. 2013. Factors influencing the cell adhesion and invasion capacity of *Mycoplasma gallisepticum*. *Acta Vet Scand*, 55, 63.
- GANAPATHY, K. & BRADBURY, J. M. 1998. Pathogenicity of *Mycoplasma gallisepticum* and *Mycoplasma imitans* in red-legged partridges (*Alectoris rufa*). *Avian Pathol*, 27, 455-63.
- GARCIA-FRUITOS, E. 2010. Inclusion bodies: a new concept. *Microb Cell Fact*, 9, 80.



- GARROW, A. G., AGNEW, A. & WESTHEAD, D. R. 2005a. TMB-Hunt: a web server to screen sequence sets for transmembrane beta-barrel proteins. *Nucleic Acids Res*, 33, W188-92.
- GARROW, A. G., AGNEW, A. & WESTHEAD, D. R. 2005b. TMB-Hunt: an amino acid composition based method to screen proteomes for beta-barrel transmembrane proteins. *BMC Bioinformatics*, 6, 56.
- GAUNSON, J. E., PHILIP, C. J., WHITEHEAD, K. G. & BROWNING, G. F. 2006. Age related differences in the immune response to vaccination and infection with *Mycoplasma gallisepticum*. *Vaccine*, 24, 1687-92.
- GERCHMAN, I., LEVISOHN, S., MIKULA, I., MANSO-SILVAN, L. & LYSNYANSKY, I. 2011. Characterization of in vivo-acquired resistance to macrolides of *Mycoplasma gallisepticum* strains isolated from poultry. *Vet Res*, 42, 90.
- GIACO, L., AMICOSANTE, M., FRAZIANO, M., GHERARDINI, P. F., AUSIELLO, G., HELMER-CITTERICH, M., COLIZZI, V. & CABIBBO, A. 2012. B-Pred, a structure based B-cell epitopes prediction server. *Adv Appl Bioinform Chem*, 5, 11-21.
- GILDEA, S., FITZPATRICK, D. A. & CULLINANE, A. 2013. Epidemiological and virological investigations of equine influenza outbreaks in Ireland (2010-2012). *Influenza Other Respir Viruses*, 7 Suppl 4, 61-72.
- GINGERICH, E. Vaccine and vaccination for *Mycoplasma gallisepticum* (MG). Proc. Cornell Poultry Conference, Ithaca, NY, 2002. 2.
- GIOMBINI, E., ORSINI, M., CARRABINO, D. & TRAMONTANO, A. 2010. An automatic method for identifying surface proteins in bacteria: SLEP. *BMC Bioinformatics*, 11, 39.
- GLAZER, D. S., RADMER, R. J. & ALTMAN, R. B. 2009. Improving structure-based function prediction using molecular dynamics. *Structure*, 17, 919-29.
- GLEW, M. D., BASEGGIO, N., MARKHAM, P. F., BROWNING, G. F. & WALKER, I. D. 1998. Expression of the pMGA genes of *Mycoplasma gallisepticum* is controlled by variation in the GAA trinucleotide repeat lengths within the 5' noncoding regions. *Infect Immun*, 66, 5833-41.
- GLEW, M. D., BROWNING, G. F., MARKHAM, P. F. & WALKER, I. D. 2000. pMGA phenotypic variation in *Mycoplasma gallisepticum* occurs in vivo and is mediated by trinucleotide repeat length variation. *Infect Immun*, 68, 6027-33.
- GOH, M. S., GORTON, T. S., FORSYTH, M. H., TROY, K. E. & GEARY, S. J. 1998. Molecular and biochemical analysis of a 105 kDa *Mycoplasma gallisepticum* cytoadhesin (GapA). *Microbiology*, 144 ( Pt 11), 2971-8.
- GOLDBERG, T., HECHT, M., HAMP, T., KARL, T., YACHDAV, G., AHMED, N., ALTERMANN, U., ANGERER, P., ANSORGE, S., BALASZ, K., BERNHOFER, M., BETZ, A., CIZMADIJA, L., DO, K. T., GERKE, J., GREIL, R., JOERDENS, V., HASTREITER, M., HEMBACH, K., HERZOG, M., KALEMANOV, M., KLUGE, M., MEIER, A., NASIR, H., NEUMAIER, U., PRADÉ, V., REEB, J., SOROKOUMOV, A., TROSHANI, I., VORBERG, S., WALDRAFF, S., ZIERER, J., NIELSEN, H. & ROST, B. 2014. LocTree3 prediction of localization. *Nucleic Acids Res*, 42, W350-5.
- GOODSWEN, S. J., KENNEDY, P. J. & ELLIS, J. T. 2014. Vacceed: a high-throughput in silico vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. *Bioinformatics*, 30, 2381-3.
- GOWTHAMAN, U. & AGREWALA, J. N. 2008. In silico tools for predicting peptides binding to HLA-class II molecules: more confusion than conclusion. *J Proteome Res*, 7, 154-63.
- GRANDI, G. 2003. Rational antibacterial vaccine design through genomic technologies. *Int J Parasitol*, 33, 615-20.
- GRANDI, G. 2004. Bioinformatics, DNA microarrays and proteomics in vaccine discovery: competing or complementary technologies. In *Genomics, Proteomics and Vaccines*. Edited by: Grandi G. John Wiley & Sons Ltd, 23-41.
- GRODIO, J. L., HAWLEY, D. M., OSNAS, E. E., LEY, D. H., DHONDT, K. V., DHONDT, A. A. & SCHAT, K. A. 2012. Pathogenicity and immunogenicity of three *Mycoplasma gallisepticum* isolates in house finches (*Carpodacus mexicanus*). *Vet Microbiol*, 155, 53-61.
- GUDA, C. 2006. pTARGET: a web server for predicting protein subcellular localization. *Nucleic Acids Res*, 34, W210-3.
- HALBEDEL, S. & STULKE, J. 2007. Tools for the genetic analysis of *Mycoplasma*. *Int J Med Microbiol*, 297, 37-44.
- HAMSTEN, C., NEIMAN, M., SCHWENK, J. M., HAMSTEN, M., MARCH, J. B. & PERSSON, A. 2009. Recombinant surface proteomics as a tool to analyze humoral immune responses in bovines infected by *Mycoplasma mycoides* subsp. *mycoides* SC. *Mol Cell Proteomics*, 8, 2544-54.
- HAMSTEN, C., WESTBERG, J., BOLSKE, G., AYLING, R., UHLEN, M. & PERSSON, A. 2008. Expression and immunogenicity of six putative variable surface proteins in *Mycoplasma mycoides* subsp. *mycoides* SC. *Microbiology*, 154, 539-49.

- HAN, L., CUI, J., LIN, H., JI, Z., CAO, Z., LI, Y. & CHEN, Y. 2006. Recent progresses in the application of machine learning approach for predicting protein functional class independent of sequence similarity. *Proteomics*, 6, 4023-37.
- HARRISON, A., PEARL, F., SILLITOE, I., SLIDEL, T., MOTT, R., THORNTON, J. & ORENGO, C. 2003. Recognizing the fold of a protein structure. *Bioinformatics*, 19, 1748-59.
- HASSELBRING, B. M., SHEPPARD, E. S. & KRAUSE, D. C. 2012. P65 Truncation Impacts P30 Dynamics during *Mycoplasma pneumoniae* Gliding. *J Bacteriol*, 194, 3000-7.
- HAWKINS, J. & BODEN, M. 2006. Detecting and sorting targeting peptides with neural networks and support vector machines. *J Bioinform Comput Biol*, 4, 1-18.
- HAYAT, S. & ELOFSSON, A. 2012. BOCTOPUS: Improved topology prediction of transmembrane beta barrel proteins. *Bioinformatics*, 28, 516-22.
- HE, Y., RAPPUOLI, R., DE GROOT, A. S. & CHEN, R. T. 2010a. Emerging vaccine informatics. *J Biomed Biotechnol*, 2010, 218590.
- HE, Y., XIANG, Z. & MOBLEY, H. L. 2010b. Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol*, 2010, 297505.
- HEGDE, S., GABRIEL, C., KRAGL, M. & CHOPRA-DEWASTHALY, R. 2015. Sheep primary cells as in vitro models to investigate *Mycoplasma agalactiae* host cell interactions. *Pathog Dis*, 73, ftv048.
- HEINSON, A. I., WOELK, C. H. & NEWELL, M. L. 2015. The promise of reverse vaccinology. *Int Health*, 7, 85-89.
- HELLMAN, U. 2000. Sample preparation by SDS/PAGE and in-gel digestion. *Proteomics in functional genomics*. Springer.
- HENNIGAN, S. L., DRISKELL, J. D., FERGUSON-NOEL, N., DLUHY, R. A., ZHAO, Y., TRIPP, R. A. & KRAUSE, D. C. 2012. Detection and differentiation of avian mycoplasmas by surface-enhanced Raman spectroscopy based on a silver nanorod array. *Appl Environ Microbiol*, 78, 1930-5.
- HENRICH, E., HEIN, C., DOTSCH, V. & BERNHARD, F. 2015. Membrane protein production in *Escherichia coli* cell-free lysates. *FEBS Lett*, 589, 1713-22.
- HERMEYER, K., BUCHENAU, I., THOMASMEYER, A., BAUM, B., SPERGSE, J., ROSENGARTEN, R. & HEWICKER-TRAUTWEIN, M. 2012. Chronic pneumonia in calves after experimental infection with *Mycoplasma bovis* strain 1067: Characterization of lung pathology, persistence of variable surface protein antigens and local immune response. *Acta Vet Scand*, 54, 9.
- HIMMELREICH, R., HILBERT, H., PLAGENS, H., PIRKL, E., LI, B. C. & HERRMANN, R. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res*, 24, 4420-49.
- HNATOW, L. L., KEELER, C. L., JR., TESSMER, L. L., CZYMMEK, K. & DOHMS, J. E. 1998. Characterization of MGC2, a *Mycoplasma gallisepticum* cytoadhesin with homology to the *Mycoplasma pneumoniae* 30-kilodalton protein P30 and *Mycoplasma genitalium* P32. *Infect Immun*, 66, 3436-42.
- HOELZLE, L. E., ZEDER, M., FELDER, K. M. & HOELZLE, K. 2014. Pathobiology of *Mycoplasma suis*. *Vet J*, 202, 20-5.
- HORTON, P., PARK, K. J., OBAYASHI, T., FUJITA, N., HARADA, H., ADAMS-COLLIER, C. J. & NAKAI, K. 2007. WoLF PSORT: protein localization predictor. *Nucleic Acids Res*, 35, W585-7.
- HU, J. & YAN, C. 2008. A method for discovering transmembrane beta-barrel proteins in Gram-negative bacterial proteomes. *Comput Biol Chem*, 32, 298-301.
- HULO, N., BAIROCH, A., BULLIARD, V., CERUTTI, L., DE CASTRO, E., LANGENDIJK-GENEVAUX, P. S., PAGNI, M. & SIGRIST, C. J. 2006. The PROSITE database. *Nucleic Acids Res*, 34, D227-30.
- HYATT, D., CHEN, G. L., LOCASCIO, P. F., LAND, M. L., LARIMER, F. W. & HAUSER, L. J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11, 119.
- HYATT, D., LOCASCIO, P. F., HAUSER, L. J. & UBERBACHER, E. C. 2012. Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics*, 28, 2223-30.
- IMV 2016. Informe del Instituto de Medicina Veterinaria.
- INABA, H., MARTIN, W., DE GROOT, A. S., QIN, S. & DE GROOT, L. J. 2006. Thyrotropin receptor epitopes and their relation to histocompatibility leukocyte antigen-DR molecules in Graves' disease. *J Clin Endocrinol Metab*, 91, 2286-94.
- INDIKOVA, I., VRONKA, M. & SZOSTAK, M. P. 2014. First identification of proteins involved in motility of *Mycoplasma gallisepticum*. *Vet Res*, 45, 99.
- INNIS, M. A., GELFAND, D. H., SNINSKY, J. J. & WHITE, T. J. 2012. *PCR protocols: a guide to methods and applications*, Academic press.

- ISHMAEL, F. T. & STELLATO, C. 2008. Principles and applications of polymerase chain reaction: basic science for the practicing physician. *Ann Allergy Asthma Immunol*, 101, 437-43.
- IVERSON-CABRAL, S. L., WOOD, G. E. & TOTTEN, P. A. 2015. Analysis of the Mycoplasma genitalium MgpB Adhesin to Predict Membrane Topology, Investigate Antibody Accessibility, Characterize Amino Acid Diversity, and Identify Functional and Immunogenic Epitopes. *PLoS One*, 10, e0138244.
- JACOB, R., BRANTON, S. L., EVANS, J. D., LEIGH, S. A. & PEEBLES, E. D. 2014. Effects of live and killed vaccines against Mycoplasma gallisepticum on the performance characteristics of commercial layer chickens. *Poult Sci*, 93, 1403-1409.
- JACOBS, P. P., RYCKAERT, S., GEYSENS, S., DE VUSSER, K., CALLEWAERT, N. & CONTRERAS, R. 2008. Pichia surface display: display of proteins on the surface of glycoengineered Pichia pastoris strains. *Biotechnol Lett*, 30, 2173-81.
- JAISWAL, V., CHANUMOLU, S. K., GUPTA, A., CHAUHAN, R. S. & ROUT, C. 2013. Jenner-predict server: prediction of protein vaccine candidates (PVCs) in bacteria based on host-pathogen interactions. *BMC Bioinformatics*, 14, 211.
- JIANG, H. X., CHEN, J. R., YAN, H. L., LI, X. N., CHEN, Z. L. & ZENG, Z. L. 2009. Molecular variability of DR-1 and DR-2 within the pvpA gene in Mycoplasma gallisepticum isolates. *Avian Dis*, 53, 124-8.
- JUNCKER, A. S., WILLENBROCK, H., VON HEIJNE, G., BRUNAK, S., NIELSEN, H. & KROGH, A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci*, 12, 1652-62.
- JURKOVIC, D. A., NEWMAN, J. T. & BALISH, M. F. 2012. Conserved terminal organelle morphology and function in Mycoplasma penetrans and Mycoplasma iowae. *J Bacteriol*, 194, 2877-83.
- KANG, I., KIM, D., HAN, K., SEO, H. W., OH, Y., PARK, C., LEE, J., GOTTSCHALK, M. & CHAE, C. 2012. Optimized protocol for multiplex nested polymerase chain reaction to detect and differentiate Haemophilus parasuis, Streptococcus suis, and Mycoplasma hyorhinis in formalin-fixed, paraffin-embedded tissues from pigs with polyserositis. *Can J Vet Res*, 76, 195-200.
- KAUSHIK, D. K. & SEHGAL, D. 2008. Developing antibacterial vaccines in genomics and proteomics era. *Scand J Immunol*, 67, 544-52.
- KEELER, C. L., JR., HNATOW, L. L., WHETZEL, P. L. & DOHMS, J. E. 1996. Cloning and characterization of a putative cytoadhesin gene (mgc1) from Mycoplasma gallisepticum. *Infect Immun*, 64, 1541-7.
- KENNY, L., O'KELLY, E., CONNELL, J., DE GASCUN, C. & HASSAN, J. 2016. Mumps outbreaks in a highly vaccinated population: Investigation of a neutralization titre against the current circulating wildtype genotype G5 mumps virus. *J Clin Virol*, 74, 8-12.
- KENT, B. N., FOECKING, M. F. & CALCUTT, M. J. 2012. Development of a novel plasmid as a shuttle vector for heterologous gene expression in Mycoplasma yeastsii. *J Microbiol Methods*, 91, 121-127.
- KLEE, E. W. & ELLIS, L. B. 2005. Evaluating eukaryotic secreted protein prediction. *BMC Bioinformatics*, 6, 256.
- KLEVEN, S. H. 1998. Mycoplasmas in the etiology of multifactorial respiratory disease. *Poult Sci*, 77, 1146-9.
- KOITA, O. A., DABITAO, D., MAHAMADOU, I., TALL, M., DAO, S., TOUNKARA, A., GUYEYE, H., NOUMSI, C., THIERO, O., KONE, M., RIVERA, D., MCMURRY, J. A., MARTIN, W. & DE GROOT, A. S. 2006. Confirmation of immunogenic consensus sequence HIV-1 T-cell epitopes in Bamako, Mali and Providence, Rhode Island. *Hum Vaccin*, 2, 119-28.
- KRAUSE, D. C. 1996. Mycoplasma pneumoniae cytoadherence: unravelling the tie that binds. *Mol Microbiol*, 20, 247-53.
- KROEMER, J. A., BONNING, B. C. & HARRISON, R. L. 2015. Expression, delivery and function of insecticidal proteins expressed by recombinant baculoviruses. *Viruses*, 7, 422-55.
- KROGH, A., LARSSON, B., VON HEIJNE, G. & SONNHAMMER, E. L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, 305, 567-80.
- KULKARNI-KALE, U., BHOSLE, S. & KOLASKAR, A. S. 2005. CEP: a conformational epitope prediction server. *Nucleic Acids Res*, 33, W168-71.
- KUMAR, A., YOGISHARADHYA, R., RAMAKRISHNAN, M. A., VISWAS, K. N. & SHIVACHANDRA, S. B. 2013. Structural analysis and cross-protective efficacy of recombinant 87 kDa outer membrane protein (Omp87) of Pasteurella multocida serogroup B:2. *Microb Pathog*, 65, 48-56.
- LADHANI, S. N., GIULIANI, M. M., BIOLCHI, A., PIZZA, M., BEEBEEJAUN, K., LUCIDARME, J., FINDLOW, J., RAMSAY, M. E. & BORROW, R. 2016. Effectiveness of Meningococcal B

- Vaccine against Endemic Hypervirulent *Neisseria meningitidis* W Strain, England. *Emerg Infect Dis*, 22, 309-11.
- LARSEN, J. E., LUND, O. & NIELSEN, M. 2006. Improved method for predicting linear B-cell epitopes. *Immunome Res*, 2, 2.
- LAU, Y. L. & FONG, M. Y. 2008. *Toxoplasma gondii*: serological characterization and immunogenicity of recombinant surface antigen 2 (SAG2) expressed in the yeast *Pichia pastoris*. *Exp Parasitol*, 119, 373-8.
- LAURIÉRE, M. 1993. A semidry electroblotting system efficiently transfers both high-and low-molecular-weight proteins separated by SDS-PAGE. *Analytical biochemistry*, 212, 206-211.
- LEE, S. H., LEE, S., CHAE, C. & RYU, D. Y. 2014. A recombinant chimera comprising the R1 and R2 repeat regions of *M. hyopneumoniae* P97 and the N-terminal region of *A. pleuropneumoniae* ApxIII elicits immune responses. *BMC Vet Res*, 10, 43.
- LEE, S. W., BROWNING, G. F. & MARKHAM, P. F. 2008. Development of a replicable oriC plasmid for *Mycoplasma gallisepticum* and *Mycoplasma imitans*, and gene disruption through homologous recombination in *M. gallisepticum*. *Microbiology*, 154, 2571-80.
- LEVERSEN, N. A., DE SOUZA, G. A., MALEN, H., PRASAD, S., JONASSEN, I. & WIKER, H. G. 2009. Evaluation of signal peptide prediction algorithms for identification of mycobacterial signal peptides using sequence data from proteomic methods. *Microbiology*, 155, 2375-83.
- LEVISOHN, S. & KLEVEN, S. H. 2000. Avian mycoplasmosis (*Mycoplasma gallisepticum*). *Rev Sci Tech*, 19, 425-42.
- LEY, D. H., BERKHOFF, J. E. & LEVISOHN, S. 1997a. Molecular epidemiologic investigations of *Mycoplasma gallisepticum* conjunctivitis in songbirds by random amplified polymorphic DNA analyses. *Emerg Infect Dis*, 3, 375-80.
- LEY, D. H., BERKHOFF, J. E. & MCLAREN, J. M. 1996. *Mycoplasma gallisepticum* isolated from house finches (*Carpodacus mexicanus*) with conjunctivitis. *Avian Dis*, 40, 480-3.
- LEY, D. H., MCLAREN, J. M., MILES, A. M., BARNES, H. J., MILLER, S. H. & FRANZ, G. 1997b. Transmissibility of live *Mycoplasma gallisepticum* vaccine strains ts-11 and 6/85 from vaccinated layer pullets to sentinel poultry. *Avian Dis*, 41, 187-94.
- LI, B. B., SHEN, J. Z., CAO, X. Y., WANG, Y., DAI, L., HUANG, S. Y. & WU, C. M. 2010. Mutations in 23S rRNA gene associated with decreased susceptibility to tiamulin and valnemulin in *Mycoplasma gallisepticum*. *FEMS Microbiol Lett*, 308, 144-9.
- LI, L. & PETROVSKY, N. 2016. Molecular mechanisms for enhanced DNA vaccine immunogenicity. *Expert Rev Vaccines*, 15, 313-29.
- LILJEROOS, L., MALITO, E., FERLENGHI, I. & BOTTOMLEY, M. J. 2015. Structural and Computational Biology in the Design of Immunogenic Vaccine Antigens. *J Immunol Res*, 2015, 156241.
- LIN, H. H., RAY, S., TONGCHUSAK, S., REINHERZ, E. L. & BRUSIC, V. 2008a. Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research. *BMC Immunol*, 9, 8.
- LIN, H. H., ZHANG, G. L., TONGCHUSAK, S., REINHERZ, E. L. & BRUSIC, V. 2008b. Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. *BMC Bioinformatics*, 9 Suppl 12, S22.
- LIN, I. Y., VAN, T. T. & SMOOKER, P. M. 2015. Live-Attenuated Bacterial Vectors: Tools for Vaccine and Therapeutic Agent Delivery. *Vaccines (Basel)*, 3, 940-72.
- LIU, T., GARCIA, M., LEVISOHN, S., YOGEV, D. & KLEVEN, S. H. 2001. Molecular variability of the adhesin-encoding gene *pvpA* among *Mycoplasma gallisepticum* strains and its application in diagnosis. *J Clin Microbiol*, 39, 1882-8.
- LUNDEGAARD, C., HOOF, I., LUND, O. & NIELSEN, M. 2010. State of the art and challenges in sequence based T-cell epitope prediction. *Immunome research*, 6, S3.
- MADAMPAGE, C. A., RAWLYK, N., CROCKFORD, G., WANG, Y., WHITE, A. P., BROWNLIE, R., VAN DONKERSGOED, J., DORIN, C. & POTTER, A. 2015. Reverse vaccinology as an approach for developing *Histophilus somni* vaccine candidates. *Biologicals*.
- MAGNUS, M., PAWLOWSKI, M. & BUJNICKI, J. M. 2012. MetaLocGramN: A meta-predictor of protein subcellular localization for Gram-negative bacteria. *Biochim Biophys Acta*, 1824, 425-33.
- MAHALIK, S., SHARMA, A. K. & MUKHERJEE, K. J. 2014. Genome engineering for improved recombinant protein expression in *Escherichia coli*. *Microb Cell Fact*, 13, 177.
- MAIONE, D., MARGARIT, I., RINAUDO, C. D., MASIGNANI, V., MORA, M., SCARSELLI, M., TETTELIN, H., BRETTONI, C., IACOBINI, E. T., ROSINI, R., D'AGOSTINO, N., MIORIN, L., BUCCATO, S., MARIANI, M., GALLI, G., NOGAROTTO, R., NARDI DEI, V., VEGNI, F., FRASER, C., MANCUSO, G., TETI, G., MADOFF, L. C., PAOLETTI, L. C., RAPPUOLI, R., KASPER, D. L., TELFORD, J. L. & GRANDI, G. 2005. Identification of a universal Group B streptococcus vaccine by multiple genome screen. *Science*, 309, 148-50.

- MARITZ-OLIVIER, C., VAN ZYL, W. & STUTZER, C. 2012. A systematic, functional genomics, and reverse vaccinology approach to the identification of vaccine candidates in the cattle tick, *Rhipicephalus microplus*. *Ticks Tick Borne Dis*.
- MARKHAM, P. F., GLEW, M. D., BROWNING, G. F., WHITHEAR, K. G. & WALKER, I. D. 1998. Expression of two members of the pMGA gene family of *Mycoplasma gallisepticum* oscillates and is influenced by pMGA-specific antibodies. *Infect Immun*, 66, 2845-53.
- MARKHAM, P. F., GLEW, M. D., SYKES, J. E., BOWDEN, T. R., POLLOCKS, T. D., BROWNING, G. F., WHITHEAR, K. G. & WALKER, I. D. 1994. The organisation of the multigene family which encodes the major cell surface protein, pMGA, of *Mycoplasma gallisepticum*. *FEBS Lett*, 352, 347-52.
- MCCLEAN, S. 2012. Eight stranded beta -barrel and related outer membrane proteins: role in bacterial pathogenesis. *Protein Pept Lett*, 19, 1013-25.
- MCMURRY, J., SBAI, H., GENNARO, M. L., CARTER, E. J., MARTIN, W. & DE GROOT, A. S. 2005. Analyzing *Mycobacterium tuberculosis* proteomes for candidate vaccine epitopes. *Tuberculosis (Edinb)*, 85, 95-105.
- MEEUSEN, E. N., WALKER, J., PETERS, A., PASTORET, P. P. & JUNGENSEN, G. 2007. Current status of veterinary vaccines. *Clin Microbiol Rev*, 20, 489-510, table of contents.
- MELONI, B. P., TILBROOK, P. A., BOULOS, S., ARTHUR, P. G. & KNUCKEY, N. W. 2006. Erythropoietin preconditioning in neuronal cultures: signaling, protection from in vitro ischemia, and proteomic analysis. *J Neurosci Res*, 83, 584-93.
- MENG, C., ZELEZNIK, O. A., THALLINGER, G. G., KUSTER, B., GHOLAMI, A. M. & CULHANE, A. C. 2016. Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief Bioinform*, bbv108.
- MENLOVE, K. J., CLEMENT, M. & CRANDALL, K. A. 2009. Similarity searching using BLAST. *Methods Mol Biol*, 537, 1-22.
- MEUNIER, M., CHEMALY, M. & DORY, D. 2016. DNA vaccination of poultry: The current status in 2015. *Vaccine*, 34, 202-11.
- MINION, F. C., VANDYK, C. & SMILEY, B. K. 1995. Use of an enhanced *Escherichia coli* opal suppressor strain to screen a *Mycoplasma hyopneumoniae* library. *FEMS Microbiol Lett*, 131, 81-5.
- MOBBS, C. V., MASTAITIS, J. W., ZHANG, M., ISODA, F., CHENG, H. & YEN, K. 2007. Secrets of the lac operon. Glucose hysteresis as a mechanism in dietary restriction, aging and disease. *Interdiscip Top Gerontol*, 35, 39-68.
- MOLLER, S., CRONING, M. D. & APWEILER, R. 2001. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*, 17, 646-53.
- MORA, M., VEGGI, D., SANTINI, L., PIZZA, M. & RAPPUOLI, R. 2003. Reverse vaccinology. *Drug Discov Today*, 8, 459-64.
- MORIEL, D. G., BERTOLDI, I., SPAGNUOLO, A., MARCHI, S., ROSINI, R., NESTA, B., PASTORELLO, I., COREA, V. A., TORRICELLI, G., CARTOCCI, E., SAVINO, S., SCARSELLI, M., DOBRINDT, U., HACKER, J., TETTELIN, H., TALLON, L. J., SULLIVAN, S., WIELER, L. H., EWERS, C., PICKARD, D., DOUGAN, G., FONTANA, M. R., RAPPUOLI, R., PIZZA, M. & SERINO, L. 2010. Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A*, 107, 9072-7.
- MOTT, R., SCHULTZ, J., BORK, P. & PONTING, C. P. 2002. Predicting protein cellular localization using a domain projection method. *Genome Res*, 12, 1168-74.
- MOXON, R. & RAPPUOLI, R. 2002. Bacterial pathogen genomics and vaccines. *Br Med Bull*, 62, 45-58.
- MUCH, P., WINNER, F., STIPKOVITS, L., ROSENGARTEN, R. & CITTI, C. 2002. *Mycoplasma gallisepticum*: Influence of cell invasiveness on the outcome of experimental infection in chickens. *FEMS Immunol Med Microbiol*, 34, 181-6.
- MUDAHI-ORENSTEIN, S., LEVISOHN, S., GEARY, S. J. & YOGEV, D. 2003. Cytoadherence-deficient mutants of *Mycoplasma gallisepticum* generated by transposon mutagenesis. *Infect Immun*, 71, 3812-20.
- MUSATOVOVA, O., HERRERA, C. & BASEMAN, J. B. 2006. Proximal region of the gene encoding cytoadherence-related protein permits molecular typing of *Mycoplasma genitalium* clinical strains by PCR-restriction fragment length polymorphism. *J Clin Microbiol*, 44, 598-603.
- NAIR, R. & ROST, B. 2002. Sequence conserved for subcellular localization. *Protein Sci*, 11, 2836-47.
- NAIR, R. & ROST, B. 2005. Mimicking cellular sorting improves prediction of subcellular localization. *J Mol Biol*, 348, 85-100.
- NAZ, A., AWAN, F. M., OBAID, A., MUHAMMAD, S. A., PARACHA, R. Z., AHMAD, J. & ALI, A. 2015. Identification of putative vaccine candidates against *Helicobacter pylori* exploiting

- exoproteome and secretome: A reverse vaccinology based approach. *Infect Genet Evol*, 32, 280-91.
- NESTA, B., SPRAGGON, G., ALTERI, C., MORIEL, D. G., ROSINI, R., VEGGI, D., SMITH, S., BERTOLDI, I., PASTORELLO, I., FERLENGHI, I., FONTANA, M. R., FRANKEL, G., MOBLEY, H. L., RAPPUOLI, R., PIZZA, M., SERINO, L. & SORIANI, M. 2012. FdeC, a novel broadly conserved *Escherichia coli* adhesin eliciting protection against urinary tract infections. *MBio*, 3.
- NI, B., BAI, F. F., WEI, Y., LIU, M. J., FENG, Z. X., XIONG, Q. Y., HUA, L. Z. & SHAO, G. Q. 2015. Apoptosis induced by lipid-associated membrane proteins from *Mycoplasma hyopneumoniae* in a porcine lung epithelial cell line with the involvement of caspase 3 and the MAPK pathway. *Genet Mol Res*, 14, 11429-11443.
- NIELSEN, M., LUNDEGAARD, C., BLICHER, T., LAMBERTH, K., HARND AHL, M., JUSTESEN, S., RODER, G., PETERS, B., SETTE, A., LUND, O. & BUUS, S. 2007. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One*, 2, e796.
- NOORMOHAMMADI, A. H. 2007. Role of phenotypic diversity in pathogenesis of avian mycoplasmosis. *Avian Pathol*, 36, 439-44.
- OLDSTONE, M. B. 2005. Molecular mimicry, microbial infection, and autoimmune disease: evolution of the concept. *Curr Top Microbiol Immunol*, 296, 1-17.
- OSZ-PAPAI, J., RADU, L., ABDULRAHMAN, W., KOLB-CHEYNEL, I., TROFFER-CHARLIER, N., BIRCK, C. & POTERSZMAN, A. 2015. Insect cells-baculovirus system for the production of difficult to express proteins. *Methods Mol Biol*, 1258, 181-205.
- PALMER, G. H., BROWN, W. C., NOH, S. M. & BRAYTON, K. A. 2011. Genome-wide screening and identification of antigens for rickettsial vaccine development. *FEMS Immunol Med Microbiol*, 64, 115-9.
- PAPAZISI, L., FRASCA, S., JR., GLADD, M., LIAO, X., YOGEV, D. & GEARY, S. J. 2002. GapA and CrmA coexpression is essential for *Mycoplasma gallisepticum* cytoadherence and virulence. *Infect Immun*, 70, 6839-45.
- PAPAZISI, L., GORTON, T. S., KUTISH, G., MARKHAM, P. F., BROWNING, G. F., NGUYEN, D. K., SWARTZELL, S., MADAN, A., MAHAIRAS, G. & GEARY, S. J. 2003. The complete genome sequence of the avian pathogen *Mycoplasma gallisepticum* strain R(low). *Microbiology*, 149, 2307-16.
- PAPAZISI, L., TROY, K. E., GORTON, T. S., LIAO, X. & GEARY, S. J. 2000. Analysis of cytoadherence-deficient, GapA-negative *Mycoplasma gallisepticum* strain R. *Infect Immun*, 68, 6643-9.
- PARKER, K. C., BEDNAREK, M. A. & COLIGAN, J. E. 1994. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol*, 152, 163-75.
- PATRONOV, A. & DOYTCHINOVA, I. 2013. T-cell epitope vaccine design by immunoinformatics. *Open Biol*, 3, 120139.
- PEARSON, W. R. 2000. Flexible sequence similarity searching with the FASTA3 program package. *Methods Mol Biol*, 132, 185-219.
- PEARSON, W. R. & LIPMAN, D. J. 1988. Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A*, 85, 2444-8.
- PEREZ-CASAL, J., PRYSLIAK, T., MAINA, T., WANG, Y., TOWNSEND, H., BERVEROV, E., NKANDO, I., WESONGA, H., LILJANDER, A., JORES, J., NAESSENS, J., GERDTS, V. & POTTER, A. 2015. Analysis of immune responses to recombinant proteins from strains of *Mycoplasma mycoides* subsp. *mycoides*, the causative agent of contagious bovine pleuropneumonia. *Vet Immunol Immunopathol*, 168, 103-10.
- PERRIE, Y., CROFTS, F., DEVITT, A., GRIFFITHS, H. R., KASTNER, E. & NADELLA, V. 2015. Designing liposomal adjuvants for the next generation of vaccines. *Adv Drug Deliv Rev*, 99, 85-96.
- PETERNEL, S. & KOMEL, R. 2010. Isolation of biologically active nanomaterial (inclusion bodies) from bacterial cells. *Microb Cell Fact*, 9, 66.
- PETERSEN, T. N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*, 8, 785-6.
- PFLAUM, K., TULMAN, E. R., BEAUDET, J., LIAO, X. & GEARY, S. J. 2015. Global changes in *Mycoplasma gallisepticum* phase-variable lipoprotein gene (vlhA) expression during in vivo infection of the natural chicken host. *Infect Immun*, 84, 351-5.
- PHARR, G., BRANTON, S., COOKSEY, A., HANSON, L., BURGESS, S., BOTTOMS, C., MCCARTHY, F. & COLLIER, S. 2006. The Recognition of a vlhA Protein from the F-Strain of *Mycoplasma gallisepticum* with Monoclonal Antibody 6F10. *International Journal of Poultry Science*, 5, 789-795.

- PIERLEONI, A., MARTELLI, P. L., FARISELLI, P. & CASADIO, R. 2006. BaCellLo: a balanced subcellular localization predictor. *Bioinformatics*, 22, e408-16.
- PIERLEONI, A., MARTELLI, P. L., FARISELLI, P. & CASADIO, R. 2007. eSLDB: eukaryotic subcellular localization database. *Nucleic Acids Res*, 35, D208-12.
- PILO, P., VILEI, E. M., PETERHANS, E., BONVIN-KLOTZ, L., STOFFEL, M. H., DOBBELAERE, D. & FREY, J. 2005. A metabolic enzyme as a primary virulence factor of *Mycoplasma mycoides* subsp. *mycoides* small colony. *J Bacteriol*, 187, 6824-31.
- PINA-PEDRERO, S., OLVERA, A., PEREZ-SIMO, M. & BENSAID, A. 2012. Genomic and antigenic characterization of monomeric autotransporters of *Haemophilus parasuis*: an ongoing process of reductive evolution. *Microbiology*, 158, 436-47.
- PING, L., ZHANG, H., ZHAI, L., DAMMER, E. B., DUONG, D. M., LI, N., YAN, Z., WU, J. & XU, P. 2013. Quantitative proteomics reveals significant changes in cell shape and an energy shift after IPTG induction via an optimized SILAC approach for *Escherichia coli*. *J Proteome Res*, 12, 5978-88.
- PIZZA, M., SCARLATO, V., MASIGNANI, V., GIULIANI, M. M., ARICO, B., COMANDUCCI, M., JENNINGS, G. T., BALDI, L., BARTOLINI, E., CAPECCHI, B., GALEOTTI, C. L., LUZZI, E., MANETTI, R., MARCHETTI, E., MORA, M., NUTI, S., RATTI, G., SANTINI, L., SAVINO, S., SCARSELLI, M., STORNI, E., ZUO, P., BROEKER, M., HUNDT, E., KNAPP, B., BLAIR, E., MASON, T., TETTELIN, H., HOOD, D. W., JEFFRIES, A. C., SAUNDERS, N. J., GRANOFF, D. M., VENTER, J. C., MOXON, E. R., GRANDI, G. & RAPPUOLI, R. 2000. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science*, 287, 1816-20.
- PLOTKIN, S. A. 2005. Vaccines: past, present and future. *Natural Medicine*, 11, 5-11.
- PLOTNICKY, H., CYBLAT-CHANAL, D., AUBRY, J. P., DEROUET, F., KLINGUER-HAMOUR, C., BECK, A., BONNEFOY, J. Y. & CORVAIA, N. 2003. The immunodominant influenza matrix T cell epitope recognized in human induces influenza protection in HLA-A2/K(b) transgenic mice. *Virology*, 309, 320-9.
- PONOMARENKO, J., BUI, H. H., LI, W., FUSSEDER, N., BOURNE, P. E., SETTE, A. & PETERS, B. 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*, 9, 514.
- POTTER, A. A. & BABIUK, L. A. 2001. New approaches for antigen discovery, production and delivery: vaccines for veterinary and human use. *Curr Drug Targets Infect Disord*, 1, 249-62.
- PRYSLIAK, T., VAN DER MERWE, J., LAWMAN, Z., WILSON, D., TOWNSEND, H., VAN DRUNEN LITTEL-VAN DEN HURK, S. & PEREZ-CASAL, J. 2011. Respiratory disease caused by *Mycoplasma bovis* is enhanced by exposure to bovine herpes virus 1 (BHV-1) but not to bovine viral diarrhoea virus (BVDV) type 2. *Can Vet J*, 52, 1195-202.
- QIAGEN 2003. Sequencing primers for pQE vectors\*. *The QIAexpressionist™. A handbook for high-level expression and purification of 6xHis-tagged proteins*, S, 18, 118.
- QIAN, P., ZHI, X., WANG, B., ZHANG, H., CHEN, H. & LI, X. 2015. Construction and immune efficacy of recombinant pseudorabies virus expressing PrM-E proteins of Japanese encephalitis virus genotype capital I, Ukrainian. *Virol J*, 12, 214.
- RAMMENSEE, H., BACHMANN, J., EMMERICH, N. P., BACHOR, O. A. & STEVANOVIC, S. 1999. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, 50, 213-9.
- RAPPUOLI, R. 2000. Reverse vaccinology. *Curr Opin Microbiol*, 3, 445-50.
- RAPPUOLI, R. 2001. Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine*, 19, 2688-91.
- RAPPUOLI, R. 2004. From Pasteur to genomics: progress and challenges in infectious diseases. *Nat Med*, 10, 1177-85.
- RAZIN, S. 1985. Molecular biology and genetics of mycoplasmas (Mollicutes). *Microbiol Rev*, 49, 419-55.
- RAZIN, S. 1999. Adherence of pathogenic mycoplasmas to host cells. *Biosci Rep*, 19, 367-72.
- RAZIN, S. 2010. Time-line of significant contributions to mycoplasmaology. *Biologicals*, 38, 191-2.
- RAZIN, S. & HAYFLICK, L. 2010. Highlights of mycoplasma research-An historical perspective. *Biologicals*, 38, 183-90.
- RAZIN, S. & JACOBS, E. 1992. Mycoplasma adhesion. *J Gen Microbiol*, 138, 407-22.
- RAZIN, S., YOGEV, D. & NAOT, Y. 1998. Molecular biology and pathogenicity of mycoplasmas. *Microbiol Mol Biol Rev*, 62, 1094-156.
- REEB, J., KLOPPMANN, E., BERNHOFER, M. & ROST, B. 2014. Evaluation of transmembrane helix predictions in 2014. *Proteins*, 83, 473-84.
- REIMER, U. 2009. Prediction of linear B-cell epitopes. *Methods Mol Biol*, 524, 335-44.
- REOLON, L. A., MARTELLO, C. L., SCHRANK, I. S., FERREIRA, H. B., BERRY, K. H., BROWN, M. B., VAUGHN, M., GOWAN, T. A., HASSKAMP, M. A. & TORRES, M. C. 2014. Survey of

- Surface Proteins from the Pathogenic *Mycoplasma hyopneumoniae* Strain 7448 Using a Biotin Cell Surface Labeling Approach. *PLoS One*, e112596.
- RIETHMAN, H. C., BOYER, M. J. & WISE, K. S. 1987. Triton X-114 phase fractionation of an integral membrane surface protein mediating monoclonal antibody killing of *Mycoplasma hyorhinis*. *Infect Immun*, 55, 1094-100.
- RIVERA-TAPIA, J. A. 2002. [Antigenic variation in Mycoplasmas]. *Gac Med Mex*, 138, 289-90.
- ROBIN, X., TURCK, N., HAINARD, A., TIBERTI, N., LISACEK, F., SANCHEZ, J. C. & MULLER, M. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77.
- RODRIGUEZ-CARMONA, E., CANO-GARRIDO, O., SERAS-FRANZOSO, J., VILLAVERDE, A. & GARCIA-FRUITOS, E. 2010. Isolation of cell-free bacterial inclusion bodies. *Microb Cell Fact*, 9, 71.
- ROUSSEL, G., TINTI, E., PERPETE, E. & MICHAUX, C. 2013. Refolding of SDS-denatured proteins using amphipathic cosolvents and osmolytes. *Curr Protoc Protein Sci*, Chapter 28, Unit28.5.
- ROUTLEDGE, S. J., MIKALIUNAITE, L., PATEL, A., CLARE, M., CARTWRIGHT, S. P., BAWA, Z., WILKS, M. D., LOW, F., HARDY, D., ROTHNIE, A. J. & BILL, R. M. 2016. The synthesis of recombinant membrane proteins in yeast for structural studies. *Methods*, 95, 26-37.
- RUBINSTEIN, N. D., MAYROSE, I., MARTZ, E. & PUPKO, T. 2009. Epitepia: a web-server for predicting B-cell epitopes. *BMC Bioinformatics*, 10, 287.
- SACHDEVA, G., KUMAR, K., JAIN, P. & RAMACHANDRAN, S. 2005. SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics*, 21, 483-91.
- SADOVSKAYA, N. S., SUTORMIN, R. A. & GELFAND, M. S. 2006. Recognition of transmembrane segments in proteins: review and consistency-based benchmarking of internet servers. *J. Bioinform. Comput. Biol.*, 4, 1033-1056.
- SAITO, S., FUJISAWA, A., OHKAWA, S., NISHIMURA, N., ABE, T., KODAMA, K., KAMOGAWA, K., AOYAMA, S., IRITANI, Y. & HAYASHI, Y. 1993. Cloning and DNA sequence of a 29 kilodalton polypeptide gene of *Mycoplasma gallisepticum* as a possible protective antigen. *Vaccine*, 11, 1061-6.
- SAN ROMAN, B., GOMEZ, S., IRACHE, J. M. & ESPUELAS, S. 2014. Co-encapsulated CpG oligodeoxynucleotides and ovalbumin in PLGA microparticles; an in vitro and in vivo study. *J Pharm Pharm Sci*, 17, 541-53.
- SANTILLAN, M. & MACKEY, M. C. 2008. Quantitative approaches to the study of bistability in the lac operon of *Escherichia coli*. *J R Soc Interface*, 5 Suppl 1, S29-39.
- SASAKI, Y., ISHIKAWA, J., YAMASHITA, A., OSHIMA, K., KENRI, T., FURUYA, K., YOSHINO, C., HORINO, A., SHIBA, T., SASAKI, T. & HATTORI, M. 2002. The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. *Nucleic Acids Res*, 30, 5293-300.
- SAVARD, C., ALVAREZ, F., PROVOST, C., CHORFI, Y., D'ALLAIRE, S., BENOIT-BIANCAMANO, M. O. & GAGNON, C. A. 2016. Efficacy of Foster's PRRS modified live virus vaccine against a Canadian heterologous virulent field strain of porcine reproductive and respiratory syndrome virus. *Can J Vet Res*, 80, 1-11.
- SCHARF, S. J. 1990. Cloning with PCR. *PCR Protocols: A guide to methods and applications*, 84-91.
- SCHLAGER, B., STRAESSLE, A. & HAFEN, E. 2012. Use of anionic denaturing detergents to purify insoluble proteins after overexpression. *BMC Biotechnol*, 12, 95.
- SCHLEGEL, S., HJELM, A., BAUMGARTEN, T., VIKSTROM, D. & DE GIER, J. W. 2014. Bacterial-based membrane protein production. *Biochim Biophys Acta*, 1843, 1739-49.
- SCHMIDL, S. R., GRONAU, K., HAMES, C., BUSSE, J., BECHER, D., HECKER, M. & STULKE, J. 2009. The Stability of Cytoadherence Proteins in *Mycoplasma pneumoniae* Requires Activity of the Protein Kinase PrkC. *Infect Immun*, 78, 184-92.
- SCHNEIDER, G. & FECHNER, U. 2004. Advances in the prediction of protein targeting signals. *Proteomics*, 4, 1571-80.
- SCHREIBER, A., HUMBERT, M., BENZ, A. & DIETRICH, U. 2005. 3D-Epitope-Explorer (3DEX): localization of conformational epitopes within three-dimensional structures of proteins. *J Comput Chem*, 26, 879-87.
- SCHURWANZ, N., JACOBS, E. & DUMKE, R. 2009. Strategy to create chimeric proteins derived from functional adhesin regions of *Mycoplasma pneumoniae* for vaccine development. *Infect Immun*, 77, 5007-15.
- SCHUSSEK, S., TRIEU, A. & DOOLAN, D. L. 2014. Genome- and proteome-wide screening strategies for antigen discovery and immunogen design. *Biotechnol Adv*, 32, 403-14.
- SCHWENDENER, R. A. 2014. Liposomes as vaccine delivery systems: a review of the recent advances. *Ther Adv Vaccines*, 2, 159-82.



- SCOTT, M. S., CALAFELL, S. J., THOMAS, D. Y. & HALLETT, M. T. 2005. Refining protein subcellular localization. *PLoS Comput Biol*, 1, e66.
- SEIB, K. L., ZHAO, X. & RAPPUOLI, R. 2012. Developing vaccines in the era of genomics: a decade of reverse vaccinology. *Clin Microbiol Infect*, 18, 109-16.
- SERRUTO, D., RAPPUOLI, R. & PIZZA, M. 2004. Meningococcus B: from genome to vaccine. In *Genomics, Proteomics and Vaccines. Edited by: Grandi G. John Wiley & Sons Ltd*, 185-201.
- SEYMOUR, L. M., DEUTSCHER, A. T., JENKINS, C., KUIT, T. A., FALCONER, L., MINION, F. C., CROSSETT, B., PADULA, M., DIXON, N. E., DJORDJEVIC, S. P. & WALKER, M. J. 2010. A processed multidomain Mycoplasma hyopneumoniae adhesin binds fibronectin, plasminogen and swine respiratory cilia. *J Biol Chem*, 285, 33971-8.
- SEYMOUR, L. M., FALCONER, L., DEUTSCHER, A. T., MINION, F. C., PADULA, M. P., DIXON, N. E., DJORDJEVIC, S. P. & WALKER, M. J. 2011a. Mhp107 is a member of the multifunctional adhesin family of Mycoplasma hyopneumoniae. *J Biol Chem*, 286, 10097-104.
- SEYMOUR, L. M., JENKINS, C., DEUTSCHER, A. T., RAYMOND, B. B., PADULA, M. P., TACCHI, J. L., BOGEMA, D. R., EAMENS, G. J., WOOLLEY, L. K., DIXON, N. E., WALKER, M. J. & DJORDJEVIC, S. P. 2011b. Mhp182 (P102) binds fibronectin and contributes to the recruitment of plasmin(ogen) to the Mycoplasma hyopneumoniae cell surface. *Cell Microbiol*, 14, 81-94.
- SHEN, Y. Q. & BURGER, G. 2010. TESTLoc: protein subcellular localization prediction from EST data. *BMC Bioinformatics*, 11, 563.
- SHIL, P. K., KANCI, A., BROWNING, G., MARENDI, M. S., NOORMOHAMMADI, A. H. & MARKHAM, P. F. 2011a. GapA+ Mycoplasma gallisepticum ts-11 has improved vaccine characteristics. *Microbiology*, 157, 1740-9.
- SHIL, P. K., KANCI, A., BROWNING, G. F. & MARKHAM, P. F. 2011b. Development and immunogenicity of recombinant GapA(+) Mycoplasma gallisepticum vaccine strain ts-11 expressing infectious bronchitis virus-S1 glycoprotein and chicken interleukin-6. *Vaccine*, 29, 3197-205.
- SHIMIZU, T. 2015. Pathogenic factors of mycoplasma. *Nihon Saikingaku Zasshi*, 70, 369-74.
- SIMIONATTO, S., MARCHIORO, S. B., GALLI, V., LUERCE, T. D., HARTWIG, D. D., MOREIRA, A. N. & DELLAGOSTIN, O. A. 2009. Efficient site-directed mutagenesis using an overlap extension-PCR method for expressing Mycoplasma hyopneumoniae genes in Escherichia coli. *J Microbiol Methods*, 79, 101-5.
- SINGH, A., UPADHYAY, V. & PANDA, A. K. 2015. Solubilization and refolding of inclusion body proteins. *Methods Mol Biol*, 1258, 283-91.
- SINGH, H., ANSARI, H. R. & RAGHAVA, G. P. 2013. Improved method for linear B-cell epitope prediction using antigen's primary sequence. *PLoS One*, 8, e62216.
- SINGH, R., GUPTA, P., SHARMA, P. K., ADES, E. W., HOLLINGSHEAD, S. K., SINGH, S. & LILLARD, J. W. 2014. Prediction and characterization of helper T-cell epitopes from pneumococcal surface adhesin A. *Immunology*, 141, 514-30.
- SIQUEIRA, F. M., THOMPSON, C. E., VIRGINIO, V. G., GONCHOROSKI, T., REOLON, L., ALMEIDA, L. G., DA FONSECA, M. M., DE SOUZA, R., PROSDOCIMI, F., SCHRANK, I. S., FERREIRA, H. B., DE VASCONCELOS, A. T. & ZAHA, A. 2013. New insights on the biology of swine respiratory tract mycoplasmas from a comparative genome analysis. *BMC Genomics*, 14, 175.
- SMILEY, B. K. & MINION, F. C. 1993. Enhanced readthrough of opal (UGA) stop codons and production of Mycoplasma pneumoniae P1 epitopes in Escherichia coli. *Gene*, 134, 33-40.
- SOARES, S. C., ABREU, V. A., RAMOS, R. T., CERDEIRA, L., SILVA, A., BAUMBACH, J., TROST, E., TAUCH, A., HIRATA, R., JR., MATTOS-GUARALDI, A. L., MIYOSHI, A. & AZEVEDO, V. 2012. PIPS: Pathogenicity Island Prediction Software. *PLoS One*, 7, e30848.
- SPRENGER, J., FINK, J. L. & TEASDALE, R. D. 2006. Evaluation and comparison of mammalian subcellular localization prediction methods. *BMC Bioinformatics*, 7 Suppl 5, S3.
- SPUESENS, E. B., VAN DE KREEKE, N., ESTEVAO, S., HOOGENBOEZEM, T., SLUIJTER, M., HARTWIG, N., VAN ROSSUM, A. M. & VINK, C. 2010. Variation in a surface-exposed region of the Mycoplasma pneumoniae P40 protein as a consequence of homologous DNA recombination between RepMP5 elements. *Microbiology*.
- STRYJEWSKA, A., KIEPURA, K., LIBROWSKI, T. & LOCHYNSKI, S. 2013. Biotechnology and genetic engineering in the new drug development. Part I. DNA technology and recombinant proteins. *Pharmacol Rep*, 65, 1075-85.
- STUDIER, F. W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*, 41, 207-34.
- SUN, J., WU, D., XU, T., WANG, X., XU, X., TAO, L., LI, Y. X. & CAO, Z. W. 2009. SEPPA: a computational server for spatial epitope prediction of protein antigens. *Nucleic Acids Res*, 37, W612-6.

- SUN, P., JU, H., ZHANG, B., GU, Y., LIU, B., HUANG, Y., ZHANG, H. & LI, Y. 2015. Conformational B-cell epitope prediction method based on antigen preprocessing and mimotopes analysis. *Biomed Res Int*, 2015, 257030.
- SUNDQUIST, B. G., CZIFRA, G. & STIPKOVITS, L. 1996. Protective immunity induced in chicken by a single immunization with *Mycoplasma gallisepticum* immunostimulating complexes (ISCOMS). *Vaccine*, 14, 892-7.
- SWEREDOSKI, M. J. & BALDI, P. 2008. PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. *Bioinformatics*, 24, 1459-60.
- SZCZEPANEK, S. M., FRASCA, S., JR., SCHUMACHER, V. L., LIAO, X., PADULA, M., DJORDJEVIC, S. P. & GEARY, S. J. 2010a. Identification of Lipoprotein MslA as a Neoteric Virulence Factor of *Mycoplasma gallisepticum*. *Infect Immun*, 78, 3475-83.
- SZCZEPANEK, S. M., TULMAN, E. R., GORTON, T. S., LIAO, X., LU, Z., ZINSKI, J., AZIZ, F., FRASCA, S., JR., KUTISH, G. F. & GEARY, S. J. 2010b. Comparative Genomic Analyses of Attenuated Strains of *Mycoplasma gallisepticum*. *Infect Immun*, 78, 1760-71.
- T, E. K., BURCHMORE, R. J., HERZYK, P. & DAVIES, R. L. 2012. Predicting the outer membrane proteome of *Pasteurella multocida* based on consensus prediction enhanced by results integration and manual confirmation. *BMC Bioinformatics*, 13, 63.
- TALUKDAR, S., ZUTSHI, S., PRASHANTH, K. S., SAIKIA, K. K. & KUMAR, P. 2014. Identification of Potential Vaccine Candidates Against *Streptococcus pneumoniae* by Reverse Vaccinology Approach. *Appl Biochem Biotechnol*, 172, 3026-41.
- THOMAS, A., LEPRINCE, P., DIZIER, I., BALL, H., GEVAERT, K., VAN DAMME, J., MAINIL, J. & LINDEN, A. 2005. Identification by two-dimensional electrophoresis of a new adhesin expressed by a low-passaged strain of *Mycoplasma bovis*. *Res Microbiol*, 156, 713-8.
- TIWARI, A. K. & SRIVASTAVA, R. 2014. A survey of computational intelligence techniques in protein function prediction. *Int J Proteomics*, 2014, 845479.
- TSIRIGOS, K. D., PETERS, C., SHU, N., KALL, L. & ELOFSSON, A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res*, 43, W401-7.
- TULMAN, E. R., LIAO, X., SZCZEPANEK, S. M., LEY, D. H., KUTISH, G. F. & GEARY, S. J. 2012. Extensive variation in surface lipoprotein gene content and genomic changes associated with virulence during evolution of a novel North American house finch epizootic strain of *Mycoplasma gallisepticum*. *Microbiology*, 158, 2073-88.
- VAN DER MERWE, J., PRYSLIAK, T., GERDTS, V. & PEREZ-CASAL, J. 2011. Protein chimeras containing the *Mycoplasma bovis* GAPDH protein and bovine host-defence peptides retain the properties of the individual components. *Microb Pathog*, 50, 269-77.
- VISHNU, U. S., SANKARASUBRAMANIAN, J., GUNASEKARAN, P. & RAJENDHRAN, J. 2015. Novel Vaccine Candidates against *Brucella melitensis* Identified through Reverse Vaccinology Approach. *Omics*, 19, 722-9.
- VIVONA, S., BERNANTE, F. & FILIPPINI, F. 2006. NERVE: new enhanced reverse vaccinology environment. *BMC Biotechnol*, 6, 35.
- VIVONA, S., GARDY, J. L., RAMACHANDRAN, S., BRINKMAN, F. S., RAGHAVA, G. P., FLOWER, D. R. & FILIPPINI, F. 2008. Computer-aided biotechnology: from immuno-informatics to reverse vaccinology. *Trends Biotechnol*, 26, 190-200.
- WALTERS, M. S. & MOBLEY, H. L. 2010. Bacterial proteomics and identification of potential vaccine targets. *Expert Rev Proteomics*, 7, 181-4.
- WANG, P., SIDNEY, J., DOW, C., MOTHE, B., SETTE, A. & PETERS, B. 2008. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol*, 4, e1000048.
- WANG, Y., WU, W., NEGRE, N. N., WHITE, K. P., LI, C. & SHAH, P. K. 2011. Determinants of Antigenicity and Specificity in immune response for Protein Sequences. *BMC Bioinformatics*, 12, 251.
- WHISSTOCK, J. C. & LESK, A. M. 2003. Prediction of protein function from protein sequence and structure. *Q Rev Biophys*, 36, 307-40.
- WHITHEAR, K. G. 1996. Control of avian mycoplasmoses by vaccination. *Rev Sci Tech*, 15, 1527-53.
- WILTON, J., JENKINS, C., CORDWELL, S. J., FALCONER, L., MINION, F. C., ONEAL, D. C., DJORDJEVIC, M. A., CONNOLLY, A., BARCHIA, I., WALKER, M. J. & DJORDJEVIC, S. P. 2009. Mhp493 (P216) is a proteolytically processed, cilium and heparin binding protein of *Mycoplasma hyopneumoniae*. *Mol Microbiol*, 71, 566-82.
- WIZEMANN, T. M., HEINRICHS, J. H., ADAMOU, J. E., ERWIN, A. L., KUNSCH, C., CHOI, G. H., BARASH, S. C., ROSEN, C. A., MASURE, H. R., TUOMANEN, E., GAYLE, A., BREWAH, Y. A., WALSH, W., BARREN, P., LATHIGRA, R., HANSON, M., LANGERMANN, S., JOHNSON, S. & KOENIG, S. 2001. Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. *Infect Immun*, 69, 1593-8.

- WU, H. N., KAWAGUCHI, C., NAKANE, D. & MIYATA, M. 2012. "Mycoplasmal Antigen Modulation," a Novel Surface Variation Suggested for a Lipoprotein Specifically Localized on *Mycoplasma mobile*. *Curr Microbiol*, 64, 433-40.
- XIANG, K., YING, G., YAN, Z., SHANSHAN, Y., LEI, Z., HONGJUN, L. & MAOSHENG, S. 2015. Progress on adenovirus-vectored universal influenza vaccines. *Hum Vaccin Immunother*, 11, 1209-22.
- XIANG, Z. & HE, Y. 2013. Genome-wide prediction of vaccine targets for human herpes simplex viruses using Vaxign reverse vaccinology. *BMC Bioinformatics*, 14 Suppl 4, S2.
- XIAO, L., PTACEK, T., OSBORNE, J. D., CRABB, D. M., SIMMONS, W. L., LEFKOWITZ, E. J., WAITES, K. B., ATKINSON, T. P. & DYBVIG, K. 2015. Comparative genome analysis of *Mycoplasma pneumoniae*. *BMC Genomics*, 16, 610.
- XU, C. G., HAO, Y. Q., ZHANG, L., HAO, R. X., LIU, X. L. & HUANG, Z. Y. 2014. Molecular cloning and immune response analysis of putative variable lipoproteins from *Mycoplasma mycoides* subsp *capri*. *Genet Mol Res*, 13, 1527-39.
- YAMANISHI, K. 2011. [Perspective of development of vaccine against infectious diseases]. *Nihon Rinsho*, 69, 1535-40.
- YANG, B., SAYERS, S., XIANG, Z. & HE, Y. 2011. Protegen: a web-based protective antigen database and analysis system. *Nucleic Acids Res*, 39, D1073-8.
- YANG, L., LI, W., KIRBERGER, M., LIAO, W. & REN, J. 2016. Design of nanomaterial based systems for novel vaccine development. *Biomater Sci*, 4, 785-802.
- YAO, B., ZHANG, L., LIANG, S. & ZHANG, C. 2012. SVMTriP: A Method to Predict Antigenic Epitopes Using Support Vector Machine to Integrate Tri-Peptide Similarity and Propensity. *PLoS One*, 7, e45152.
- YASSER, E. & HONAVAR, V. 2010. Recent advances in B-cell epitope prediction methods. *Immunome research*, 6, S2.
- YOGEV, D., MENAKER, D., STRUTZBERG, K., LEVISOHN, S., KIRCHHOFF, H., HINZ, K. H. & ROSENGARTEN, R. 1994. A surface epitope undergoing high-frequency phase variation is shared by *Mycoplasma gallisepticum* and *Mycoplasma bovis*. *Infect Immun*, 62, 4962-8.
- YONG, Y., LIU, S., HUA, G., JIA, R., ZHAO, Y., SUN, X., LIAO, M. & JU, X. 2015. Identification and functional characterization of Toll-like receptor 2-1 in geese. *BMC Vet Res*, 11, 108.
- YU, C. S., CHEN, Y. C., LU, C. H. & HWANG, J. K. 2006. Prediction of protein subcellular localization. *Proteins*, 64, 643-51.
- YU, C. S., CHENG, C. W., SU, W. C., CHANG, K. C., HUANG, S. W., HWANG, J. K. & LU, C. H. 2014. CELLO2GO: a web server for protein subCELLular LOcalization prediction with functional gene ontology annotation. *PLoS One*, 9, e99368.
- YU, N. Y., WAGNER, J. R., LAIRD, M. R., MELLI, G., REY, S., LO, R., DAO, P., SAHINALP, S. C., ESTER, M., FOSTER, L. J. & BRINKMAN, F. S. 2010. PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*, 26, 1608-15.
- ZAMPINI, M., STEVENS, P. R., PACHEBAT, J. A., KINGSTON-SMITH, A., MUR, L. A. & HAYES, F. 2015. RapGene: a fast and accurate strategy for synthetic gene assembly in *Escherichia coli*. *Sci Rep*, 5, 11302.
- ZENG, Y., LIU, L., HE, J., LIU, Y., ZHU, C., YOU, X. & WU, Y. 2012. Screening and identification of the mimic epitope of the adhesion protein of *Mycoplasma genitalium*. *Can J Microbiol*, 58, 898-908.
- ZHANG, B., GU, Y., LIU, B., HUANG, Y., ZHANG, H., LI, Y., MUNIKUMAR, M., PRIYADARSHINI, I. V., PRADHAN, D., UMAMAHESWARI, A. & VENGAMMA, B. 2013. Computational approaches to identify common subunit vaccine candidates against bacterial meningitis. *Biomed Res Int*, 5, 155-64.
- ZHANG, G. Z., ZHANG, R., ZHAO, H. L., WANG, X. T., ZHANG, S. P., LI, X. J., QIN, C. Z., LV, C. M., ZHAO, J. X. & ZHOU, J. F. 2010. A safety assessment of a fowlpox-vectored *Mycoplasma gallisepticum* vaccine in chickens. *Poult Sci*, 89, 1301-6.
- ZHANG, J., ZHAO, X., SUN, P., GAO, B. & MA, Z. 2014. Conformational B-cell epitopes prediction from sequences using cost-sensitive ensemble classifiers and spatial clustering. *Biomed Res Int*, 2014, 689219.
- ZHANG, Q. & WISE, K. S. 1997. Localized reversible frameshift mutation in an adhesin gene confers a phase-variable adherence phenotype in mycoplasma. *Mol Microbiol*, 25, 859-69.
- ZHOU, M., BOEKHORST, J., FRANCKE, C. & SIEZEN, R. J. 2008. LocateP: genome-scale subcellular-location predictor for bacterial proteins. *BMC Bioinformatics*, 9, 173.
- ZIMMERMAN, C. U., HERRMANN, R. & ROSENGARTEN, R. 2014. XerC-mediated DNA inversion at the inverted repeats of the UU172-phase-variable element of *Ureaplasma parvum* serovar 3. *Microbiol Res*, 170, 263-9.

- ZIMMERMAN, C. U., ROSENGARTEN, R. & SPERGSEER, J. 2013. Interaction of the putative tyrosine recombinases RipX (UU145), XerC (UU222), and CodV (UU529) of *Ureaplasma parvum* serovar 3 with specific DNA. *FEMS Microbiol Lett*, 340, 55-64.
- ZUCKERT, W. R. 2014. Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochim Biophys Acta*, 1843, 1509-16.

## Anexo 1. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA VALIDACIÓN DEL PREDICTOR DE BARRILES- $\beta$

### 1. Conjunto de validación positivo\*

1E54	2F1T	1QD6	1UYN	2HDF	2JMM	2ODJ	1BXW
1TLW	2QDZ	2QOM	2VQI	3FIP	2R8A	3BRZ	2V9U
2XE1	2ZLE	2VDD	2JQY	2GR8	1TQQ	1ORM	3M3R
1PNZ	2GSK	1UJW	2YSU	2HDI	3CSL	1a0s	3a2s
4aiq	3anz	3bry	4c00	4c69	3csl	3down	3dwo
3dzm	4e1t	3efm	3emo	4epa	2erv	2f1c	1fep
3fhh	3fid	4fqe	4frx	4fsp	4fuv	1fw2	4gey
2grx	2guf	1h6s	3hw9	2iah	3jty	2k0l	1k24
4k3b	4k3c	1kmo	3kvn	2lhf	2lme	4mee	4mid
2mlh	1mm4	1mpr	3njt	2o4v	3o44	1p4t	2por
4q35	3qlb	3qq2	3qra	3rbh	3rfz	3slj	3sy9
3syb	1uun	1uxf	1v36	3v89	3w9t	3wi4	2wjq
1wp1	2x27	2x4m	1xkw	1yc9	2ynk		

### 2. Conjunto de validación negativo\*

2a06_E	2a06_G	2a06_W	3a0b_J	3a0b_L	3a0b_M	3a0b_Z	4a01_A
4a2n_B	3a3y_B	3a3y_G	2a65_A	3a7k_A	4a82_A	1a91_A	3abk_B
2ac6_A	2acz_C	2acz_D	4ac5_H	2akh_X	2akh_Y	2akh_Z	2amk_A
3am6_A	3aou_A	3aqp_A	4aq5_A	4aq5_B	4aq5_C	1ar1_B	3arc_I
4atv_A	2aui_A	4av3_A	4aw6_B	2axt_B	2axt_C	2axt_E	2axt_F
2axt_H	3ayf_A	3aym_A	4ayw_A	2b0x_A	2b2f_A	3b29_A	4b2q_J
3b4r_A	3b5d_A	3b5w_A	2b6s_A	2b6v_A	2b76_C	2b76_D	3b8c_A
3b9w_A	4bbj_A	1bcc_C	1be3_K	4bem_J	4bgn_A	2bmn_A	4bpd_A
4bpm_A	2bs2_C	3bvd_B	4bwz_A	1bzk_A	3bz1_y	3bz1_X	3c02_A
1c17_M	3c1g_A	4c48_C	4c7r_A	4c9j_A	4cad_C	4cbj_A	4cg5_A
4cg5_B	4cg6_C	3chx_B	4chv_A	4cl4_A	3cn5_A	4cof_A	2cpb_A
3cx5_D	2d2c_A	2d2c_B	2d2c_D	2d2c_F	2d2c_G	3d31_C	3ddl_A
3det_A	3dhw_A	3din_D	3din_E	4dji_A	3dl8_C	3dl8_E	1dop_D
3dtu_A	4dve_A	4dx5_A	4dxw_A	4ea3_A	3egw_C	3eh3_A	4eiy_A
1et2_S	2evu_A	4ev6_A	4ezc_A	4f35_B	4f4c_A	2f75_A	2f93_A
2f95_B	2fbw_C	2fbw_D	4fc4_A	4fe1_F	4fe1_I	4fe1_K	4fe1_L

Anexo 1. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA  
VALIDACIÓN DEL PREDICTOR DE BARRILES- $\beta$

4fe1_X	1fft_A	1fft_B	1fft_C	3fh6_G	2fkW_A	4fpd_A	3fwl_A
1fx8_A	2fyn_B	2fyn_C	2g1x_A	4g1u_A	2g2a_A	3g6b_A	4g7v_S
4gby_A	4gd3_S	4gd3_A	2gfp_A	2gFz_A	3gi8_C	4grv_A	4gx0_A
4h33_A	2h8a_A	3h9v_A	3hd6_A	3hd7_B	2he6__	4he8_B	4he8_D
4he8_E	4he8_G	4he8_I	4he8_C	4hkr_A	4hqj_E	4htt_A	4huk_A
4huq_S	4huq_T	4hw9_A	4hwa_A	4hyc_A	4hyj_A	4hzu_S	4i9w_A
2ibz_G	2ibz_I	4iff_A	2iil__	1ijd_B	2ik3__	4ikv_A	4il3_A
4in5_L	4in5_H	2iqp_A	2iub_A	4iu8_A	3ixz_A	3ixz_B	4ixq_A
4ixq_K	3iyz_A	3iz1_A	3j09_A	4j05_A	3j1z_P	1j4n_A	2j58_A
3j5p_B	2j7a_C	4j72_A	4j7c_K	4jBw_F	4jc9_5	4jc9_B	4jc9_6
4jkv_A	2jln_A	2jo1_A	2jp3_A	4jta_B	2jwa_A	3jyc_A	3k07_A
4k0e_A	2k1k_A	4k1c_A	3k3f_A	2k73_A	2k9j_B	2k9p_A	2k9y_A
2ka1_A	2kb7_P	3kg2_A	2klu_A	2knc_A	4knf_A	2kog_A	1kpW_A
3kp9_A	4kpp_A	1kqf_B	1kqf_C	2ks9_A	2ksd_A	2kse_A	2ksf_A
1kyk_A	4ky0_A	2l0j_A	3l1l_A	2l2t_A	2l35_A	2l6w_A	4l6r_A
2l8s_A	2lat_A	1lbn_A	2lck_A	4lds_A	4lep_A	1lgh_A	3llq_A
1lnq_A	2lnl_A	2lom_A	2loq_A	2lor_A	2los_A	2lot_A	2lp1_A
4lp8_A	4lto_A	1lvi_A	4lxj_A	2lzl_A	4lz6_A	2m0q_A	2m20_A
4m48_A	1m56_D	2m59_A	4m58_A	2m67_A	4m64_B	2m7g_A	3m71_A
2maw_A	4mbs_A	2mfr_A	2mgy_A	1mhs_A	3mk7_A	3mk7_B	3mk7_C
3mkt_A	4mlb_A	2mpn_A	3mp7_A	3mp7_B	4mqS_A	4mrn_A	4nab_A
3nd0_A	4njn_A	4npp_A	2nq2_A	2nr9_A	4ntj_A	3nym_A	4nyk_A
3o0r_B	3o0r_C	1o5w_A	4o6m_A	4o6y_A	3o7p_A	4o9p_C	4o9p_D
2oar_A	4od4_A	1oea_A	4ogq_C	1orq_C	3org_A	4or2_A	1oz5_A
4p00_A	4p00_B	4p2z_A	1p49_A	3p5n_A	1p7b_A	4p79_A	1pb2_A
1pb4_C	1pb4_D	4pe5_A	4pe5_B	4pgr_A	4phz_A	3pjs_K	3pjz_A
3pl9_A	4pl0_A	1pw4_A	4px7_A	4pyp_A	2q7m_A	3q7k_A	1q90_M
3qe7_A	3qf4_A	3qf4_B	1qle_C	3qnq_A	3rce_A	2rdd_B	3rgb_C
1rh5_A	1rh5_B	1rh5_C	3rhw_A	1rkl_A	3rko_L	3rko_M	3rko_N
3rko_K	3rko_A	3rko_J	3s0x_A	1s6e_A	3sn6_R	1sr1_A	3tdo_A
3tij_A	4tkr_A	3tx3_A	1u5n_A	3udc_A	3ug9_A	3ukm_A	3uq4_A
3ux4_A	3v2w_A	3v5s_A	3vmq_A	3vou_A	2vpw_C	3vr8_C	3vr8_D
2w1p_A	3waj_A	3wdo_A	2wie_A	3wmf_A	3wmm_M	3wmm_A	3wo7_A
1wrg_A	2wsc_A	2wsc_F	2wsc_G	2wsc_H	2wsc_I	2wsc_J	2wsc_K
2wsc_1	2wsc_2	2wsc_3	2wsw_A	1xio_A	2xq2_A	2xut_A	1y36_A
1y5d_A	2y5y_A	2y69_D	2y69_G	2y69_I	2y69_J	2y69_K	2y69_L
2y69_M	1y8s_A	2yev_A	2yev_B	2yev_C	1yg7_A	2yiu_A	2yiu_B
1yo9_L	2yvx_A	1z8e_A	1zc7_A	2zjs_Y	2zjs_E	1zti_A	3zux_A
1zza_A							

\* Se muestran los PDB-ID [identificadores que poseen estas proteínas en la base de datos Protein DataBase (PDB)]

## Anexo 2. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA VALIDACIÓN DE Bact-VAP

### 1. Proteínas de superficie de micoplasmas

Proteína	ID (GenBank)*	Proteína	ID (GenBank)*	Proteína	ID (GenBank)*
P40	CAC87106.1	vpmaF1	YP_003515746.1	MSC_0957	NP_975924.1
VpmaV	CAL59405.1	vpmaA	YP_003515741.1	MSC_0535	NP_975519.1
VpmaW	CAL59406.1	vpmaE	YP_003515745.1	MSC_0111	NP_975121.1
vpmaX	CAL59407.1	vpmaH	YP_003515969.1	MSC_0916	NP_975885.1
VpmaY	CAL59408.1	vpmaC	YP_003515743.1	MSC_0908	NP_975877.1
VpmaZ	CAL59410.1	vpmaI	YP_003515970.1	MSC_0628	NP_975608.1
PhnD	YP_001256411.1	vpmaD1	YP_003515744.1	MSC_0963	NP_975930.1
MAG_6520	YP_001256791.1	vpmaJ	YP_003515972.1	MSC_0991	NP_975958.1
MAG_1050	YP_001256249.1	vpmaB	YP_003515742.1	MSC_0500	NP_975486.1
MAG_3600	YP_001256502.1	pykF	AAP56829.1	MSC_0752	NP_975728.1
MAG_4720	YP_001256611.1	proteinP102	YP_287505.1	MSC_0708	NP_975685.1
MAG_3240	YP_001256466.1	MHP7448_0064	YP_287464.1	MSC_0606	NP_975586.1
MAG_6170	YP_001256756.1	MHP7448_0088	YP_287488.1	MSC_0770	NP_975746.1
MAG_6490	YP_001256788.1	MHP7448_0138	YP_287535.1	MSC_0587	NP_975568.1
MAG_2400	YP_001256382.1	MHP7448_0217	YP_287614.1	MSC_0882	NP_975852.1
MAG_3590	YP_001256501.1	MHP7448_0333	YP_287728.1	MSC_1066	NP_976025.1
MAG_2430	YP_001256385.1	MHP7448_0352	YP_287746.1	MSC_0707	NP_975684.1
MAG_4740	YP_001256613.1	MHP7448_0353	YP_287747.1	MSC_0406	NP_975396.1
MAG_6480	YP_001256787.1	MHP7448_0355	YP_287749.1	MSC_0292	NP_975291.1
MAG_2870	YP_001256429.1	MHP7448_0356	YP_287750.1	MSC_0104	NP_975114.1
MAG_5910	YP_001256730.1	MHP7448_0366	YP_287760.1	MSC_0437	NP_975423.1
MAG_2510	YP_001256393.1	MHP7448_0367	YP_287761.1	MSC_0033	NP_975044.1
MAG_4460	YP_001256585.1	MHP7448_0368	YP_287762.1	MSC_0107	NP_975117.1
MAG_2950	YP_001256437.1	MHP7448_0373	YP_287767.1	MSC_0793	NP_975767.1
MAG_5030	YP_001256641.1	MHP7448_0378	YP_287772.1	MSC_0870	NP_975840.1
MAG_1450	YP_001256288.1	MHP7448_0467	YP_287857.1	MSC_0631	NP_975611.1
MAG_2840	YP_001256426.1	MHP7448_0468	YP_287858.1	MSC_0690	NP_975667.1
MAG_4380	YP_001256577.1	MHP7448_0474	YP_287864.1	MSC_0115	NP_975125.1
MAG_5040	YP_001256642.1	MHP7448_0480	YP_287869.1	MSC_0135	NP_975140.1
MAG_1510	YP_001256294.1	MHP7448_0513	YP_287902.1	MSC_0624	NP_975604.1
MAG_2220	YP_001256364.1	MHP7448_0538	YP_287927.1	MSC_0598	NP_975579.1
P30	YP_001256489.1	MHP7448_0611	YP_287997.1	MSC_0534	NP_975518.1
P48	YP_001256156.1	MHP7448_0621	YP_288007.1	MSC_0869	NP_975839.1
MAG_2540	YP_001256396.1	MHP7448_0629	YP_288014.1	MSC_0632	NP_975612.1
MAG_5080	YP_001256646.1	MHP7448_0661	YP_288046.1	MSC_0417	NP_975403.1
MAG_2350	YP_001256377.1	MHP7448_0674	YP_288059.1	MSC_0671	NP_975649.1
oppA	YP_001256182.1	mhp107	AAV27416.1	MSC_0361	NP_975356.1
MAG_2340	YP_001256376.1	mhp683	AAV28028.1	MSC_0035	NP_975046.1
MAG_2610	YP_001256403.1	mhp108	AAV27417.1	MSC_0405	NP_975395.1
MAG_1670	YP_001256310.1	Vmm	CAE77023.1	MSC_0457	NP_975443.1
MAG_4640	YP_001256603.1	MSC_0321	NP_975319.1	MSC_0155	NP_975158.1
MAG_2000	YP_001256342.1	MSC_0371	NP_975364.1	MSC_0906	NP_975875.1
MAG_5150	YP_001256653.1	MSC_0622	NP_975602.1	MSC_0065	NP_975076.1
MAG_6200	YP_001256759.1	MSC_0776	NP_975752.1	MSC_0714	NP_975691.1
MAG_1000	YP_001256244.1	MSC_0152	NP_975155.1	MSC_0576	NP_975557.1
MAG_3460	YP_001256488.1	MSC_0471	NP_975457.1	MSC_0858	NP_975828.1
MAG_7130	YP_001256851.1	MSC_0941	NP_975908.1	MSC_0757	NP_975733.1
MAG_2530	YP_001256395.1	MSC_0620	NP_975600.1	MSC_0411	NP_975401.1
MAG_1570	YP_001256300.1	MSC_0472	NP_975458.1	MSC_0389	NP_975380.1
MAG_1980	YP_001256340.1	MSC_0236	NP_975236.1	MSC_0623	NP_975603.1

Anexo 2. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA  
VALIDACIÓN DE Bact-VAP|

Proteína	ID (GenBank)*
MSC_0032	NP_975043.1
MSC_0537	NP_975521.1
MSC_0710	NP_975687.1
MSC_0112	NP_975122.1
MSC_0374	NP_975367.1
MSC_1018	NP_975982.1
MSC_0468	NP_975454.1
MSC_0958	NP_975925.1
MSC_0473	NP_975459.1
MSC_0448	NP_975434.1
MSC_0670	NP_975648.1
MSC_0575	NP_975556.1
MSC_0103	NP_975113.1
MSC_0777	NP_975753.1
MSC_0384	NP_975375.1
MSC_0558	NP_975542.1
MSC_0179	NP_975182.1
MSC_0582	NP_975563.1
MSC_0992	NP_975959.1
MSC_0062	NP_975073.1
MSC_0287	NP_975286.1
MSC_0946	NP_975913.1
GapA	AAP56530.2
MGC2	AAP56529.2
Hlp3	AAP56527.2
CrmA	AAP56531.1
MGA_1033	AAP56576.2
pvpA	ADB96887.1
VlhA.4.10	AAP56553.1
vlhA.5.04	AAP56689.1
VlhA.1.02	AAP56773.1
VlhA.1.05	AAP56776.2
VlhA.2.01	AAP56805.2
VlhA.2.02	AAP56806.2
MgPA	AAC71410.1
P32	AAC71540.1
P65	AAC71436.1
HMW1	AAC71534.1
HMW3	AAC71539.1
P200	AAC71613.1
P97	AAV27763.1
P146	AAZ54025.1

Proteína	ID (GenBank)*
P216	AAZ53862.1
Lppt	AAZ44456.1
P102	AAZ44286.1
p76	AAZ44580.1
P65	AAZ44739.1
MHJ_0112	AAZ44204.1
MMOB4250	AAT27911.1
MvspA	AAT27584.1
MvspB	AAT27808.1
MvspC	AAT27809.1
MvspD	AAT27814.1
MvspE	AAT27815.1
MvspG	AAT27818.1
MvspH	AAT27819.1
MvspI	AAT27820.1
MvspJ	AAT27822.1
MvspK	AAT27823.1
MvspL	AAT27824.1
MvspM	AAT28093.1
MvspO	AAT28095.1
MvspP	AAT28119.1
MYPE1550	BAC43946.1
MYPE2630	BAC44054.1
MYPE6520	BAC44444.1
MYPE6580	BAC44450.1
MYPE6760	BAC44468.1
MYPE6770	BAC44469.1
MYPE7035	BAC44495.1
MYPE7030	BAC44496.1
MYPE7040	BAC44497.1
MYPE7060	BAC44499.1
MYPE7330	BAC44527.1
MYPE7370	BAC44531.1
MYPE8180	BAC44611.1
MYPE2900	BAC44082.1
P1	ABR53878.1
P30	NP_110141.1
MPN309	NP_109997.1
MPN447	Q50365.1
MPN310	NP_109998.1
MPN452	Q50360.1

Proteína	ID (GenBank)*
MPN142	NP_109830.1
MPN567	NP_110256.1
VSAA	CAC13696.1
VSAC	CAC13698.1
VSAE	CAC13699.1
VSAG	CAC13700.1
VSAI	CAC13701.1
rplB	AAP56404.1
vlhA.3.0.1	ADB96891.1
vlhA.5.01b	ADB96873.1
potE	AAP56543.1
vlhA.5.01a	AAP56686.2
vlhA.1.08a	ADB96882.1
MGA_0674	AAP56382.2
MGA_0556	AAP57041.2
crmB	AAP56532.1
MGA_0654	AAP56371.1
plpA	AAP56670.2
vlhA.4.03a	ADB96857.1
rplW	AAP56403.2
vlhA.5.10b	ADB96875.1
groEL	AAP56827.2
valS	AAP56918.2
crmC	AAP56533.2
MGA_0042	AAP56756.2
vlhA.5.01c	ADB96874.1
MGA_0241	AAP56872.2
MS53_0613	YP_278731.1
MS53_0264	YP_278394.1
sgaT-2	YP_278801.1
secD	YP_278297.1
MS53_0349	YP_278472.1
MS53_0235	YP_278365.1
oppC-1	YP_278470.1
oppB	YP_278314.1
MS53_0104	YP_278236.1
lgt	YP_278715.1
ktrB	YP_278638.1
oppC	YP_278315.1
pepF	YP_278473.1
MS53_0190	YP_278320.1

\* - Identificador GenBank



## 2. Proteínas citoplasmáticas de micoplasmas

Proteína	NA (UniProt)*
TRMFO	A5IXT6
LON	A5IYF2
YQGF	B3PM70
PNP	B3QZG1
PTH	B3QZM3
EFTS	B3QZN3
SYH	B3QZS0
OBG	B3QZT1
YBEY	B3QZV1
RBFA	B3QZW6
RNC	B3R075
SYI	B3R0D7
GPDA	B3R0E3
KCY	B3R0E5
EFP	B3R0E7
RNH3	B3R0F3
SYS	B3R0F5
SYN	B3R0I1
SYT	B3R0I8
DNAA	B3R0L6
TSAD	B3R0M3
AMPA	B3R0N4
TRMD	B3R0P2
ACP	B3R0P8
RNC	D3FH24
RNC	E1QCT1
KAD	O52352
LDH	O52354
KAD	P10251
DPO3B	P24117
DNAA	P35889
IF1	P38037
RSMA	P43038
PTGA	P45618
EFTS	P47246
MNME	P47254
SYM	P47267
EFP	P47272
DPO3	P47277
SYH	P47281
RSMI	P47302
RNM5	P47303
KPRS	P47304
PTH	P47329
TILS	P47330
UNG	P47343
TRXB	P47348
KGUA	P47353
G6PI	P47357
SYW	P47372
NADK	P47374
RNJ	P47385
SYR	Q4A900
RLMH	Q4A908
KAD	Q50299
DPO3B	Q50313
SYM	Q50319
DNAA	Q59549
TILS	Q601M6

Proteína	NA (UniProt)
KAD	P47417
SYFA	P47436
SYFB	P47437
GRPE	P47443
UVRC	P47448
SCPA	P47455
MRAZ	P47463
RSMH	P47464
TIG	P47480
MTHFS	P47487
Y248	P47490
DLDH	P47513
APT	P47518
SYP	P47525
SYA	P47534
SMC	P47540
PTAS	P47541
G3P	P47543
KCY	P47572
SYI	P47587
TRML	P47588
SYT	P47615
SYR	P47618
RSMG	P47620
URK	P47622
YBEY	P47628
AMPA	P47631
CH60	P47632
HPRR	P47666
TPIS	P47670
PYRH	P47672
HPRT	P47696
AMPA	P47707
TPIS	P47721
DPO3	P47729
YQGF	P57114
RRF	P61306
SYA	P61704
GRPE	P71499
SSRP	P75043
TSAD	P75055
SYD	P75068
PIP	P75092
SYS	P75107
RSMA	P75113
SYI	P75122
RBGA	P75135
PTMA	P75145
CH10	P75205
ERA	P75210
THII	P75228
ULAC	P75292
SYR	Q7NC67
RSMA	Q7NC69
CH60	Q8CXQ7
RSMA	Q8EU92
COAX	Q8EUB0
KAD	Q8EUD3
COAD	Q8EUG2

Proteína	NA (UniProt)
KAD	P47417
SYFA	P47436
SYFB	P47437
GRPE	P47443
UVRC	P47448
SCPA	P47455
MRAZ	P47463
RSMH	P47464
TIG	P47480
MTHFS	P47487
Y248	P47490
DLDH	P47513
APT	P47518
SYP	P47525
SYA	P47534
SMC	P47540
PTAS	P47541
G3P	P47543
KCY	P47572
SYI	P47587
TRML	P47588
SYT	P47615
SYR	P47618
RSMG	P47620
URK	P47622
YBEY	P47628
AMPA	P47631
CH60	P47632
HPRR	P47666
TPIS	P47670
PYRH	P47672
HPRT	P47696
AMPA	P47707
TPIS	P47721
DPO3	P47729
YQGF	P57114
RRF	P61306
SYA	P61704
GRPE	P71499
SSRP	P75043
TSAD	P75055
SYD	P75068
PIP	P75092
SYS	P75107
RSMA	P75113
SYI	P75122
RBGA	P75135
PTMA	P75145
CH10	P75205
ERA	P75210
THII	P75228
ULAC	P75292
Y050	Q98RF9
DPO3B	Q98RK6
ULAB	Q9EXD8
RECF	Q9KHU6
DNAA	Q9KHU8
KGUA	Q9KX62
RSMG	Q9KX67

Anexo 2. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA  
VALIDACIÓN DE Bact-VAP|

Proteína	NA (UniProt)*
COAX	Q6KH54
RSMG	Q6KH77
EFP	Q6KH93
SYF	Q6KH97
SYFB	Q6KHC8
DNAJ	Q6KHF9
GPDA	Q6KHG2
SYA	Q6KHK5
YQGF	Q6KHK6
RNC	Q6KHN3
YBEY	Q6KHU4
SYH	Q6KI17
RIMP	Q6KIE1
SYC	Q6KIE2
SYR	Q6KIG6
GRPE	Q6KIH8
TIG	Q6KIK5
RSMG	Q6MS67
SYI	Q6MT28
MNME	Q7NAD9
SYE	Q7NAE5
SYC	Q7NAH7
TILS	Q7NAI9
Y600	Q7NAN6
TPIS	Q7NAQ4
SYV	Q7NAR7
NADK	Q7NAU0
CLPB	Q7NAZ3
TSAD	Q7NB15
OBG	Q7NB30
ACPS	Q7NB74
SCPA	Q7NB76
SCPB	Q7NB77
MRAZ	Q7NB78
RSMH	Q7NB79
YBEY	Q7NB85
TIG	Q7NBA6
SYI	Q7NBG0
SYH	Q7NBH9
SYP	Q7NBP5
SYH	Q7NBS9
SYT	Q7NBT5
KCY	Q7NBV1
DNAJ	Q7NBW0
YQGF	Q7NBY7
MNMA	Q7NBZ0
RBFA	Q7NBZ3
RRF	Q7NC19
PYRH	Q7NC20
ybeYPE	A0A063YF95
NPL1_03130PE	A0A063YK50
NPL1_00315PE	A0A063YLP9
dnaAPE	A0A067YP51
pheTPE	A0A077L4V8
ftsZPE	A0A077L641
vacBPE	A0A077L694
plsXPE	A0A077L6F6
tpiAPE	A0A077L6G0
uvrCPE	A0A077L6G5
gidBPE	A0A077L6U3
cysSPE	A0A077L918
thiIPE	A0A077L949
leuSPE	A0A077L9K9
mraZPE	A0A077L9Q4

Proteína	NA (UniProt)
PYRH	Q8EUG9
RRF	Q8EUH0
SYFB	Q8EUJ9
DNAJ	Q8EUM4
GRPE	Q8EUM5
SYI	Q8EUN9
YQGF	Q8EUR7
Y802	Q8EUW6
RSMG	Q8EUW9
SYF	Q8EUX2
MIAB	Q8EUX4
PYRDB	Q8EUY2
TIG	Q8EV76
SYR	Q8EVF2
SECA	Q8EVL6
RLMH	Q8EVP2
SYS	Q8EW01
PLSX	Q8EW41
ERA	Q8EW66
RSMH	Q8EW81
SYD	Q8EWB7
GPDA	Q8EWH5
TILS	Q8EWQ7
RBFA	Q8EWT9
SYV	Q8EX08
SCPB	Q8EX11
SCPA	Q8EX12
KITH	Q8RLE0
TILS	Q98PE3
LDH	Q98PG4
SYT	Q98PH6
DNAJ	Q98PI9
TRXB	Q98PK9
RBFA	Q98PP7
MRAZ	Q98Q74
YBEY	Q98QI3
RNR	Q98QL0
SYFB	Q98QL4
SYH	Q98QM8
PGK	Q98QW4
TIG	Q98R01
SYC	Q98R33
GRPE	Q98R67
RSMG	Q98R82
KPRS	Q98R83
ACPS	Q98R97
COAD	Q98RB3
KCY	Q98RC0
RECU	Q98RC3
cmkPE	A0A084U4Q5
dnaNPE	A0A084U4U4
dnaAPE	A0A084U4U5
rsmAPE	A0A097NSH6
uvrCPE	A0A097NSM8
rncPE	A0A097NSP1
infCPE	A0A097NST2
dnaJPE	A0A097NSW0
hisSPE	A0A097NSZ6
rbgAPE	A0A097NT34
thiSPE	A0A097NTM8
grpEPE	A0A097NTP0
rnc_2PE	A0A097NTT5
dnaAPE	A0A097SS13
dnaNPE	A0A097SS25

Proteína	NA (UniProt)
nadKPE	A0A014KTJ7
rnrPE	A0A014L5K7
hsiOPE	A0A014M1J5
ybeYPE	A0A014M1T9
dnaNPE	A0A014M252
adkPE	A0A014M2T3
xseBPE	A0A014MF44
rbfAPE	A0A014MFC3
lplAPE	A0A014MFH0
nanAPE	A0A014MG42
eraPE	A0A014MH91
tpiAPE	A0A014MHE6
uvrCPE	A0A014MID3
gyrAPE	A0A014NM95
metGPE	A0A014NMA2
coaEPE	A0A014NMQ4
rimMPE	A0A014NNF4
recUPE	A0A014NNG5
gcpPE	A0A014NQN8
dnaAPE	A0A059XKZ6
grpEPE	A0A059XVY2
ftsZPE	A0A059XWC0
rnrPE	A0A059XYX5
ftsZPE	A0A059XZB6
thiSPE	A0A059Y0H8
K668_00010PE	A0A059Y2K0
rsmGPE	A0A059Y398
K668_02525PE	A0A059Y4D9
ackAPE	A0A059Y8Y4
K668_03755PE	A0A059Y9D8
dnaJPE	A0A059Y9K2
mesJPE	A0A060PYG1
dnaNPE	A0A060Q102
gidBPE	A0A060Q2A4
ksgAPE	A0A060Q2D2
argSPE	A0A060Q5H5
glnSPE	A0A060Q5I0
rlmHPE	A0A060Q5W9
NPL1_03135PE	A0A063Y815
recUPE	A0A063Y8X5
thiSPE	A0A063YA65
polCPE	A0A063YA82
cmkPE	A0A063YB05
NPL1_03590PE	A0A063YB61
grpEPE	A0A063YCG6
rncPE	A0A063YCK3
NPL1_02860PE	A0A063YDC4
NPL2_02430PE	A0A063YDE6
tigPE	A0A063YE35
uvrCPE	A0A097ST93
dinPPE	A0A097ST94
recUPE	A0A097ST98
cmkPE	A0A097STA0
proSPE	A0A097STB2
MGM1_4780PE	A0A097STC1
trmEPE	A0A097STC9
polCPE	A0A097STD0
coaEPE	A0A097STF4
dnaEPE	A0A097STG7
rbfAPE	A0A097STH5
nusAPE	A0A097STI9
MGM1_5540PE	A0A097STL4
metGPE	A0A097STM4
thiIPE	A0A097STM5

Anexo 2. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA  
VALIDACIÓN DE Bact-VAP|

Proteína	NA (UniProt)*
ungPE	A0A077LBA3
prfAPE	A0A077LBT0
nusAPE	A0A077LBY2
pheSPE	A0A078MU39
dinBPE	A0A084EMA0
pheTPE	A0A084U2K7
pheSPE	A0A084U2K8
tigPE	A0A084U2N6
deoCPE	A0A084U2S0
metGPE	A0A084U2T0
rsmEPE	A0A084U2T9
P271_129PE	A0A084U2W4
ylqFPE	A0A084U2X0
grpEPE	A0A084U2X5
rlmNPE	A0A084U2Y4
nusAPE	A0A084U2Z5
coaEPE	A0A084U306
thiIPE	A0A084U312
tilSPE	A0A084U318
acpSPE	A0A084U323
P271_193PE	A0A084U324
mhCPE	A0A084U325
ybeYPE	A0A084U3A8
recUPE	A0A084U3L0
tpiAPE	A0A084U3N0
smcPE	A0A084U3N8
pgiPE	A0A084U3T9
pfkAPE	A0A084U3U8
ungPE	A0A084U3W5
trmEPE	A0A084U406
proCPE	A0A084U441
gidBPE	A0A084U457
tyrSPE	A0A084U459
acpPPE	A0A084U479
coaXPE	A0A084U491
rlmHPE	A0A084U4D6
ppnKPE	A0A084U4F4
yqgFPE	A0A084U4H4
tsfPE	A0A084U4M8
coaDPE	A0A084U4N1
gyrBPE	A0A084U4P2
grpEPE	A0A0E3H7T8
smcPE	A0A0E3H7Z8
VY93_01475PE	A0A0E3KNC5
VY93_03190PE	A0A0E3N908
VY93_00010PE	A0A0E3NAM8
tilSPE	A0A0E3NBC5
grpEPE	A0A0F5GZY9
frrPE	A0A0F5H034
infBPE	A0A0F5H099
pheTPE	A0A0F5H0A8
smcPE	A0A0F5H0F2
dnaAPE	A0A0F5H0H0
MMELEA_04950PE	A0A0F5H103
tpiAPE	A0A0F5H110
MMELEA_03640PE	A0A0F5H187
ftsZPE	A0A0F5H1H5
dnaAPE	A0A0F6CJR5
recUPE	A0A0F6CK36
aspSPE	A0A0F6CK74
cspRPE	A0A0F6CKI6
grpEPE	A0A0F6CKK2
rpoDPE	A0A0F6CKP3
eraPE	A0A0F6CKP7

Proteína	NA (UniProt)
smcPE	A0A097SS44
coaDPE	A0A097SS45
MGM1_0220PE	A0A097SS46
rncPE	A0A097SS56
plsXPE	A0A097SS65
vacBPE	A0A097SS67
rlmNPE	A0A097SS68
pheSPE	A0A097SS81
pheTPE	A0A097SS83
adkPE	A0A097SSA5
pfkAPE	A0A097SSD0
argSPE	A0A097SSD8
pthPE	A0A097SSE6
MGM1_1340PE	A0A097SSG4
cysSPE	A0A097SSG5
tsaDPE	A0A097SSJ3
eraPE	A0A097SSJ6
ybeYPE	A0A097SSL2
thrSPE	A0A097SSM9
hisSPE	A0A097SSP0
ackAPE	A0A097SSP7
valSPE	A0A097SST1
rbgAPE	A0A097SST8
pgiPE	A0A097SSW0
ungPE	A0A097SSW1
tpiAPE	A0A097SSW7
tyrSPE	A0A097SSY1
mraWPE	A0A097SSZ0
mraZPE	A0A097SSZ2
tsfPE	A0A097ST16
grpEPE	A0A097ST17
udkPE	A0A097ST19
pyrHPE	A0A097ST34
pflAPE	A0A097ST36
dnaJPE	A0A097ST39
alaSPE	A0A097ST54
aptPE	A0A097ST56
tigPE	A0A097ST58
trpSPE	A0A097ST62
tilSPE	A0A097ST64
gidBPE	A0A097ST77
ybeYPE	A0A0H1S705
lonAPE	A0A0H1S769
gidAPE	A0A0H1S7L0
mraWPE	A0A0H1S8I5
cmkPE	A0A0H1S8V1
rncPE	A0A0H1S958
thiIPE	A0A0H1S9R0
tdkPE	A0A0H1SA91
obgEPE	A0A0H1SA94
pthPE	A0A0H1SAD7
gltXPE	A0A0H1SAW6
lonAPE	A0A0H1SC40
udgPE	A0A0H1SCB2
lonPE	A0A0H1SCJ8
cysSPE	A0A0H1SCL8
mesJPE	A0A0H1SCV7
pheTPE	A0A0H1SCW7
secAPE	A0A0H1SDL1
grpEPE	A0A0H1SED9
cmkPE	A0A0H1SEE6
proSPE	A0A0H1SFG3
dnaJPE	A0A0H1SG45
leuSPE	A0A0H1SHK4

Proteína	NA (UniProt)
ksgAPE	A0A097STM8
scpBPE	A0A097STN0
dnaJPE	A0A0A8E600
grpEPE	A0A0A8E6J4
tsfPE	A0A0A8E7R9
pheTPE	A0A0A8E8G7
MYE_01450PE	A0A0C5B6I9
plsXPE	A0A0C5B755
tilSPE	A0A0C5B7J9
infBPE	A0A0C5BAQ0
dnaJPE	A0A0C5BAX1
trmFOPE	A0A0C5BB00
pheTPE	A0A0C5BBD4
udkPE	A0A0C5BBI1
rbgAPE	A0A0C5BBR6
thiIPE	A0A0C5BBV6
tyrSPE	A0A0C5BBW6
polCPE	A0A0C5BDL0
valSPE	A0A0C5BDX3
rsmIPE	A0A0C5BEE2
smcPE	A0A0C5BEG5
trmEPE	A0A0C5BNK4
rsmGPE	A0A0C5BQ27
MDIS_00005PE	A0A0C5WWR8
polCPE	A0A0C5WXY8
nusAPE	A0A0C5X7D6
pyrHPE	A0A0C6FNQ5
uvrAPE	A0A0C6FPD9
glyAPE	A0A0C6FYC2
gmKPE	A0A0C6FYJ1
ksgAPE	A0A0C6G290
eraPE	A0A0C6G4L2
rnrPE	A0A0D5ZJ37
grpEPE	A0A0D5ZJA8
VO56_01050PE	A0A0D5ZJE6
VO56_01550PE	A0A0D5ZJM0
VO56_02895PE	A0A0D5ZJT0
tilSPE	A0A0D5ZK48
rncPE	A0A0D5ZK90
VO56_02890PE	A0A0D5ZKB2
polCPE	A0A0E3H7M8
ppaPE	B3QZX9
gltXPE	B3QZY4
cysSPE	B3QZY5
valSPE	B3QZY8
adkPE	B3R054
ATP_00044PE	B3R067
plsXPE	B3R076
ATP_00076PE	B3R099
pheTPE	B3R0A2
pheSPE	B3R0A3
polCPE	B3R0A4
prfBPE	B3R0C2
secAPE	B3R0C3
pfkAPE	B3R0C8
tdkPE	B3R0G3
thdFPE	B3R0I7
tpiAPE	B3R0I9
dnaNPE	B3R0L7
polCPE	B3R0M5
proSPE	B3R0N8
rimMPE	B3R0P1
infCPE	B3R0P4
gmKPE	B3R0Q0

Anexo 2. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA  
VALIDACIÓN DE Bact-VAP|

Proteína	NA (UniProt)*
arcA_2PE	A0A0F6CKU3
udkPE	A0A0F6CKU4
rrnPE	A0A0F6CKV3
pfkAPE	A0A0F6CKX2
lplAPE	A0A0F6CKX3
thiIPE	A0A0F6CKZ1
umuCPE	A0A0F6CL60
rmvVPE	A0A0F6CL61
ungPE	A0A0F6CL93
rbgAPE	A0A0F6CLA5
dnaNPE	A0A0F6CLJ8
pheTPE	A0A0F6CLY5
rncPE	A0A0F6CM11
mraZPE	A0A0F6X1T0
grpEPE	A0A0F6X204
AAW50_00005PE	A0A0F6ZP68
AAW50_00590PE	A0A0F6ZPC3
rrhBPE	A0A0H1S250
grpEPE	A0A0H1S2C9
pcmPE	A0A0H1S2E5
ppaPE	A0A0H1S2Y9
smpBPE	A0A0H1S331
secAPE	A0A0H1S3J0
ndkPE	A0A0H1S3N9
mnMEPE	A0A0H1S3T1
hisSPE	A0A0H1S459
dnaJPE	A0A0H1S464
gidAPE	A0A0H1S4A7
mnMAPE	A0A0H1S4Z1
asnCPE	A0A0H1S5C4
thyAPE	A0A0H1S664
ksgAPE	A0A0H1S6C1
mesJPE	A0A0H1S6D6
lplAPE	D3VPX6
uvrCPE	D4XUV5
MALL_0639PE	D4XVE3
MALL_0627PE	D4XVF5
dnaNPE	D4XWH4
MALL_0455PE	D4XWH5
tigPE	D4XWS7
nusAPE	D5E4K5
panKPE	D5E4L7
dnaJPE	D5E4M0
glySPE	D5E583
pheTPE	D5E5B6
pheSPE	D5E5B8
hisRSPE	D5E5H2
grpEPE	D5E5Z9
tigPE	D5E689
dnaAPE	E0TJV6
mraZPE	E0TJZ6
adkPE	E0TK22
grpEPE	E0TKB5
rsmIPE	E0TKD2
uvrCPE	E0TKM1
hisSPE	E0TKN7
rbfAPE	E0TKS8
tilSPE	E0TKX3
MHR_0535PE	E0TKX5
pgiPE	E0TLA0
tigPE	E0TLB8
MHR_0610PE	E0TLB9
ybeYPE	E0TLJ5
EF-TsPE	E3VW08

Proteína	NA (UniProt)
uvrBPE	A0A0H1SK83
MPNE_0660PE	A0A0H3DLK9
recUPE	A0A0H3DLL3
frrPE	A0A0H3DM30
MPNE_0310PE	A0A0H3DMW6
oppF-valSPE	B3PLT3
pheTPE	B3PM49
ppaPE	B3PMC0
adkPE	B3PMM7
grpEPE	B3PN02
rncPE	B3PN90
nusAPE	B3PNC2
mesJPE	B3PNH2
tigPE	B3PNH4
tpmPE	B3QZF5
grpEPE	B3QZK5
dnaJPE	B3QZK7
lplAPE	B3QZK8
recUPE	B3QZK9
rsmAPE	B3QZM6
pyrHPE	B3QZN4
ruvXPE	B3QZP7
alaSPE	B3QZP8
ileSPE	B3QZQ2
tilSPE	B3QZS2
acpSPE	B3QZU4
ATP_00295PE	B3QZU5
tigPE	B3QZU7
lonAPE	B3QZU8
eraPE	B3QZV2
nusAPE	B3QZW3
infBPE	B3QZW5
smpBPE	B3QZX8
glxXPE	E8ZKD5
pyrHPE	E8ZKH4
dnaNPE	F0QPY3
gmKPE	F0QQZ0
fusAPE	F0QR88
dnaJPE	F0QR89
grpEPE	F0V1H5
rsmAPE	F0V2M0
valSPE	F6FFX1
polC2PE	F6FFZ1
pheTPE	F6FHP7
ppnKPE	F6FHR4
mraZPE	F6FI13
rncPE	F9QCF2
hisSPE	F9QCR8
GIG_01448PE	F9QCY4
GIG_01478PE	F9QCZ0
tilSPE	F9QD80
GIG_02025PE	F9QD99
dnaAPE	F9QDA0
grpEPE	F9QDP0
dnaJPE	F9QDW2
polCPE	F9QDW4
tigPE	F9QEF9
MCSF7_02679PE	F9UJ78
tilSPE	F9UJ86
dnaAPE	F9UJC1
tpiAPE	F9UK21
ftsZPE	F9UK47
grpEPE	F9UKD5
ksgAPE	G4ELA7

Proteína	NA (UniProt)
rncPE	C4XDU4
plsXPE	C4XDU5
ybeYPE	C4XE22
MBIO_0151PE	C4XE44
tigPE	C4XE89
valSPE	C4XED6
MBIO_0248PE	C4XEE1
MBIO_0263PE	C4XEF6
tufPE	C4XEI5
ffhPE	C4XEJ3
grpEPE	C4XES8
MBIO_0445PE	C4XEY8
ftsZPE	C4XEZ1
rsmHPE	C4XEZ3
mraZPE	C4XEZ4
tdkPE	C4XF39
recUPE	C4XF73
MBIO_0666PE	C4XFK9
tyrSPE	C4XFL4
MBIO_0680PE	C4XFM3
tilSPE	C4XFM9
polCPE	C4XFQ7
rrhBPE	C4XFT1
dnaAPE	C5J5F8
rsmGPE	C5J5I2
MCJ_000500PE	C5J5K4
ungPE	C5J5Y9
ptsHPE	C5J5Z4
tilSPE	C5J6A6
rncPE	C5J6D4
pheTPE	C5J6V6
uvrCPE	C5J756
argSPE	D1J7J7
alaSPE	J3VGP9
trpSPE	J3VHD3
hrcAPE	J3YGY3
coaEPE	J3YSU7
plsXPE	J3YT82
pheTPE	K7XZ75
MCYN0002PE	L0RU71
MCYN0197PE	L0RUE6
MCYN0307PE	L0RUI7
MCYN0412PE	L0RUN7
MCYN0712PE	L0RV36
MCYN0881PE	L0RV91
MCYN0400PE	L0RVL2
MCYN0705PE	L0RWB9
MCYN0748PE	L0RY99
nusAPE	L5L7I0
aspSPE	L5L7R4
scpBPE	L5L992
xseAPE	L5LA71
prsPE	L5LAX2
acpSPE	L5LBB8
arcAPE	M9WCB0
thrSPE	M9WCN5
ybeYPE	M9WDH7
gmKPE	M9WDV4
rmvVPE	M9WE97
coaDPE	M9WGT3
hisSPE	M9WGX4
eraPE	M9WHA1
mraZPE	M9WHM0
dnaEPE	M9WHW6

Anexo 2. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA  
VALIDACIÓN DE Bact-VAP|

Proteína	NA (UniProt)*
MBOVPG45_0249PE	E4PZG8
ftsYPE	E8ZGG9
rncPE	E8ZGH5
dnaAPE	E8ZGH9
dnaNPE	E8ZGI0
proSPE	E8ZGI7
cysSPE	E8ZGK0
rsmHPE	E8ZHS5
hisSPE	E8ZJV9
obgPE	E8ZJW1
pfkAPE	E8ZK10
ybeYPE	E8ZK19
rnRPE	E8ZK21
mnMPE	E8ZK33
rbgAPE	E8ZK35
purAPE	E8ZK36
pheSPE	E8ZK48
gcpPE	E8ZK50
dnaJPE	E8ZK85
grpEPE	E8ZK86
cmkPE	E8ZK98
HF1_00690PE	E8ZKB1
lysSPE	E8ZKB2
pthPE	E8ZKB5
asnSPE	E8ZKC3
dnaNPE	Q6KIT6
dnaAPE	Q6KIT7
pyrHPE	Q6MT13
MSC_0417PE	Q6MTJ1
pfkPE	Q6MTY5
sunPE	Q6MU26
hsp33PE	Q6MUK0
nusAPE	Q7NBZ6
rmVPE	Q8EUI2
MYPE9380PE	Q8EUI8
polCPE	Q8EUK5
MYPE8650PE	Q8EUQ5
MYPE6080PE	Q8EVF6
ackAPE	Q8EVJ4
nadKPE	Q8EUV9
rncPE	Q8EW40
rnRPE	Q8EW72
thrSPE	Q8EWD0
recUPE	Q8EWG0
MYPE2280PE	Q8EWH7
ungPE	Q8EWJ6
coaEPE	Q8EWK2
MYPE1610PE	Q8EWP2
nusAPE	Q8EWU2
MYPE530PE	Q8EWZ7
MYPE20PE	Q8EX48
ackAPE	Q93T71
MYPU_3320PE	Q98QM9
MYPU_3070PE	Q98QQ4
MYPU_3060PE	Q98QQ5
MYPU_3050PE	Q98QQ6
rncPE	Q98R47
MYPU_1560PE	Q98R54
MYPU_0320PE	Q98RH7
MYPU_0010PE	Q98RK7
BN801_00128PE	R5LI23
tpiAPE	R5LIN8
clpPPE	R5LIS3
cobBPE	R5LIV7

Proteína	NA (UniProt)
cysSPE	G4ELJ8
uvrCPE	G4ELM0
gmKPE	G4ELX2
alaSPE	G4EM02
trmDPE	G4EMQ9
tsaDPE	G4ENG5
GUU_04731PE	G4ENI2
purAPE	G8C2N7
glTPE	G8C2T8
asnSPE	G8C2Y8
recUPE	G8C304
mraWPE	G8C305
pyrHPE	G8C394
gmKPE	H6N5U3
aspSPE	I1A4J9
MCANUFG4_01353PE	I1A6W8
secAPE	I6Z6V7
polCPE	I7B941
valSPE	I7BA91
hisSPE	I7C5B5
MHLP_00005PE	I7CEC4
polCPE	J3T8Y9
smcPE	J3TQX5
rnHCPPE	J3TR84
tsfPE	J3VGN6
gpsAPE	R5LNL1
alaSPE	R5LPE6
BN801_00943PE	R5LQD8
rbfAPE	R5LQS9
BN801_01180PE	R5LR16
dnaJPE	R5LR67
BN801_01280PE	R5LRG9
thrSPE	R5LRJ4
BN801_00106PE	R5M0E4
recFPE	R5M0G5
BN801_00236PE	R5M0U5
trpSPE	R5M0Y6
BN801_00276PE	R5M113
BN801_00425PE	R5M1D2
miaBPE	R5M1H4
BN801_00530PE	R5M1M4
BN801_00421PE	R5M1P2
frrPE	R5M2E1
deoBPE	R5M2G0
murBPE	R5M2Q7
proSPE	R5M2V8
ddlPE	R5M309
BN801_00846PE	R5M321
rsmAPE	R5M391
nusAPE	R5M3E3
thrSPE	R5M3H9
pheTPE	R5M3K0
smcPE	R5M3T4
grpEPE	R5M3U3
mnMPE	R5M4H0
tigPE	R5M4I3
BN817_00149PE	R5M7X4
BN817_00176PE	R5M826
gmKPE	R5M9S6
xerCPE	R5MAK7
ileSPE	R5MAY0
pheSPE	R5MB29
murBPE	R5MB81
grpEPE	R5MBL6

Proteína	NA (UniProt)
grpEPE	N9SR57
tigPE	N9TRQ5
grpEPE	N9TRZ9
tilSPE	N9TSY6
tilSPE	N9UAW1
ungPE	N9V498
rncPE	N9VCH4
oppF-valSPE	N9VD52
rsmEPE	N9VEY5
ftsZPE	Q4A669
DeoBPE	Q60111
uvrCPE	Q6KH99
gyrAPE	Q6KHF0
cmkPE	Q6KHG0
recUPE	Q6KHG4
hsp33PE	Q6KHH2
pgkPE	Q6KHJ1
aspSPE	Q6KI16
vacBPE	Q6KI25
nusAPE	Q6KIE0
ftsYPE	Q6KIF2
ungPE	Q6KIL8
thiIPE	Q6KIL9
ptsIPE	Q6KIM9
dnaEPE	Q6KIN3
rbfAPE	R5MIH1
rnHCPPE	R5MU65
pheTPE	R5MUV0
tyrSPE	R5MUY4
uvrCPE	R5MV93
rimMPE	R5MVF1
leuSPE	R5MVV0
murDPE	R5MVG6
recFPE	R5MWB0
rsmGPE	R5MWB9
rsmIPE	R5Y7H3
dnaJPE	R5Y7I1
rnHBPE	R5Y7J3
BN735_00595PE	R5Y7M7
BN735_00170PE	R5Y7P7
infCPE	R5Y7Y2
cysSPE	R5Y812
recAPE	R5Y826
mfdPE	R5Y835
recFPE	R5Y845
tsaDPE	R5Y875
rncPE	R5Y8D0
BN735_00469PE	R5Y8N2
BN735_00509PE	R5Y8R9
BN735_00001PE	R5YDM8
BN735_00463PE	R5YET1
obgPE	R5YEY1
xerCPE	R5YEZ3
ybeYPE	R5YF23
tilSPE	R5YFR3
gmKPE	R5YNA0
recXPE	R5YNZ1
BN735_00431PE	R5YP79
grpEPE	R5YPM9
mrnCPPE	R5YQ82
rsmAPE	R5YQ96
dnaAPE	R5YQB5
murBPE	R5YQF1
rimMPE	R5YQP3

Anexo 2. CONJUNTO DE PROTEÍNAS EMPLEADAS PARA LA  
VALIDACIÓN DE Bact-VAP|

Proteína	NA (UniProt)*
recXPE	R5LJ27
BN801_00564PE	R5LJE4
eraPE	R5LJU5
rimMPE	R5LJY9
hisSPE	R5LK60
xerCPE	R5LK92
murDPE	R5LK99
rsmIPE	R5LKG4
pheSPE	R5LKL7
infBPE	R5LKS7
BN801_01294PE	R5LLG0
arcAPE	R5LLS8
lonPE	R5LLX6
rlmHPE	R5LN13
BN801_00184PE	R5LN44
thiIPE	R5LN57
sepFPE	R5LNB1
grpEPE	R7HFAQ6
tyrSPE	R7HFX7
glxSPE	R7HFX6
alaSPE	R7HGX5
gpsAPE	R7HGX3
BN672_01183PE	R7HGX5
tigPE	R7HGD4
xerCPE	R7HGL4
ldhPE	R7HGX8
thyAPE	R7HH09
ackAPE	R7HH16
rimMPE	R7HH48
BN672_00980PE	R7HHH5
BN672_00985PE	R7HHH9
BN672_00472PE	R7HI07
pgkPE	R7HIH6
recFPE	R7HIK3
rsmGPE	R7HIL2
murAPE	R7HIM8
dtdPE	R7HJE1
smcPE	R7HK27
BN672_01186PE	R7HK68
acpSPE	R7NC37
murBPE	R7NEL5
BN778_00394PE	R7NEQ7

Proteína	NA (UniProt)
sigAPE	R5MBW9
BN817_00957PE	R5MBZ9
rsmAPE	R5MC43
mfdPE	R5MC59
dnaAPE	R5MC73
mnmEPE	R5MC79
cysSPE	R5MCK1
mrmCPE	R5MCM1
htpGPE	R5MCT4
thiSPE	R5MCW6
bukPE	R5MCX9
BN817_00203PE	R5MDX8
lysSPE	R5MEN3
valSPE	R5MF78
smcPE	R5MGA6
recUPE	R5MG10
xseBPE	R5MGX2
BN778_01100PE	R7NF20
coaDPE	R7NF25
murCPE	R7NF75
BN778_00567PE	R7NF82
ddiPE	R7NFD4
xerCPE	R7NFE6
rimMPE	R7NFP3
xerCPE	R7NFT3
BN778_00027PE	R7NFU5
murDPE	R7NG37
mnrPE	R7NGG6
ungPE	R7NGP2
BN778_00822PE	R7NGR1
BN778_01031PE	R7NGT8
pgkPE	R7NGV9
BN778_00543PE	R7NGW8
dapFPE	R7NH26
polCPE	R7NH86
BN778_00674PE	R7NHM0
BN778_00699PE	R7NHP4
recXPE	R7NI19
BN778_01089PE	R7NIX3
glmSPE	R7NIZ9
pheSPE	R7NJ15
pheTPE	R7NJ21

Proteína	NA (UniProt)
BN735_00142PE	R5Z234
xerCPE	R5Z3A7
polCPE	R5Z3I3
queAPE	R5Z3U6
glmSPE	R5Z3Y7
argSPE	R5Z4D8
ddiPE	R5Z4K4
cmkPE	R5Z4N4
thiSPE	R7HDJ2
rsmAPE	R7HEH2
argSPE	R7HEK9
BN672_00596PE	R7HF18
mraZPE	R7HF26
murBPE	R7HF29
pnpPE	R7HF36
tyrSPE	R7HF93
ddiPE	R7HFM0
lysSPE	R7NJ64
nadKPE	R7NJC5
xseAPE	R7NK28
smcPE	S5G7M2
dnaAPE	S6G3H3
PRV_01605PE	U5NBY1
enoPE	U5NCV0
nusAPE	U5ND05
PRV_00005PE	U5NEY2
PRV_01585PE	U5NFX5
tsfPE	U5NG50
OVS_00070PE	V9R7L8
tsfPE	V9R7X3
lysSPE	V9R7Z7
OVS_02725PE	V9RAU1
OVS_00005PE	V9RB24
pheTPE	W5US70
dnaAPE	W5USA7
scpAPE	W5UT97
ybeYPE	W5UTN3
thiSPE	W5UTP5
MYB_00970PE	W5UTQ5
uvrCPE	W5UTV1
mnrPE	W5UUI1

\* - Número de Acceso UniProt

Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |

**Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP, ORGANIZADAS SEGÚN  
SU PROBABILIDAD DE SER ANTÍGENOS DE SUPERFICIE (Listado de  
Priorización)**

Nº	Gen	Proteína	Talla	Lipoproteína	Hélice $\alpha$	Barril $\beta$	Otras	Citadhesina	Antigénica
1	gapA	cytadhesin protein GapA	1122	-	x	-	-	x	x
2	crmA	cytadherence related molecule A(CrmA)	1062	-	x	-	-	x	x
3	MGA_1216	unique hypothetical protein	96	-	x	-	-	x	x
4	MGA_0648	conserved lipoprotein	750	x	-	-	-	x	x
5	MGA_0674	mycoplasma specific lipoprotein A	746	x	-	-	-	x	x
6	MGA_0913	unique hypothetical lipoprotein	271	x	-	-	-	x	x
7	vlhA.4.01	VlhA.4.01 variable lipoprotein family protein	643	x	-	-	-	x	x
8	vlhA.4.02	VlhA.4.02 variable lipoprotein family protein	751	x	-	-	-	x	x
9	vlhA.4.04	VlhA.4.04 variable lipoprotein family protein	679	x	-	-	-	x	x
10	vlhA.4.06	VlhA.4.06 variable lipoprotein family protein	698	x	-	-	-	x	x
11	vlhA.4.07.3	VlhA.4.07.3 variable lipoprotein family protein	679	x	-	-	-	x	x
12	vlhA.4.07.5	VlhA.4.07.5 variable lipoprotein family protein	698	x	-	-	-	x	x
13	vlhA.4.10	VlhA.4.10 variable lipoprotein family protein	795	x	-	-	-	x	x
14	vlhA.4.11	VlhA.4.11 variable lipoprotein family protein	690	x	-	-	-	x	x
15	vlhA.4.12	VlhA.4.12 variable lipoprotein family protein	701	x	-	-	-	x	x
16	MGA_0993	conserved hypothetical lipoprotein	622	x	-	-	-	x	x
17	MGA_1079	conserved hypothetical lipoprotein	1575	x	-	-	-	x	x
18	vlhA.5.02	VlhA.5.02 variable lipoprotein family protein	610	x	-	-	-	x	x
19	vlhA.5.03	VlhA.5.03 variable lipoprotein family protein	728	x	-	-	-	x	x
20	vlhA.5.04	VlhA.5.04 variable lipoprotein family protein	740	x	-	-	-	x	x
21	vlhA.5.05	VlhA.5.05 variable lipoprotein family protein	644	x	-	-	-	x	x
22	vlhA.5.08	VlhA.5.08 variable lipoprotein family protein	661	x	-	-	-	x	x
23	vlhA.5.09	VlhA.5.09 variable lipoprotein family protein	701	x	-	-	-	x	x
24	vlhA.5.10a	VlhA.5.10 variable lipoprotein family proteinN-terminal domain protein	642	x	-	-	-	x	x
25	vlhA.5.11	VlhA.5.11 variable lipoprotein family protein	711	x	-	-	-	x	x
26	vlhA.5.12	VlhA.5.12 variable lipoprotein family protein	678	x	-	-	-	x	x
27	vlhA.1.02	VlhA.1.02 variable lipoprotein family protein	666	x	-	-	-	x	x
28	vlhA.1.03	VlhA.1.03 variable lipoprotein family protein	682	x	-	-	-	x	x
29	vlhA.1.05	VlhA.1.05 variable lipoprotein family protein	730	x	-	-	-	x	x
30	vlhA.1.07	VlhA.1.07 variable lipoprotein family protein	728	x	-	-	-	x	x
31	vlhA.2.02	VlhA.2.02 variable lipoprotein family protein	582	x	-	-	-	x	x
32	MGA_0226	conserved hypothetical lipoprotein	897	x	-	-	-	x	x
33	MGA_0267	unique hypothetical lipoprotein	623	x	-	-	-	x	x
34	vlhA.3.03	VlhA.3.03 variable lipoprotein family protein	645	x	-	-	-	x	x

**Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |**

Nº	Gen	Proteína	Talla	Lipoproteína	Hélice α	Barril β	Otras	Citadhesina	Antigénica
35	vlhA.3.04	VlhA.3.04 variable lipoprotein family protein	734	x	-	-	-	x	x
36	vlhA.3.05	VlhA.3.05 variable lipoprotein family protein	708	x	-	-	-	x	x
37	vlhA.3.07	VlhA.3.07 variable lipoprotein family protein	656	x	-	-	-	x	x
38	vlhA.3.08	VlhA.3.08 variable lipoprotein family protein	692	x	-	-	-	x	x
39	MGA_0416	conserved hypothetical lipoprotein	519	x	-	-	-	x	x
40	MGA_0565	conserved hypothetical lipoprotein	792	x	-	-	-	x	x
41	vlhA.4.05	VlhA.4.05 variable lipoprotein family protein	673	-	-	x	-	x	x
42	vlhA.4.07	VlhA.4.07 variable lipoprotein family protein	667	-	-	x	-	x	x
43	vlhA.4.07.1	VlhA.4.07.1 variable lipoprotein family protein	684	-	-	x	-	x	x
44	vlhA.4.07.4	VlhA.4.07.4 variable lipoprotein family protein	673	-	-	x	-	x	x
45	vlhA.4.07.6	VlhA.4.07.6 variable lipoprotein family protein	667	-	-	x	-	x	x
46	vlhA.4.08	VlhA.4.08 variable lipoprotein family protein	688	-	-	x	-	x	x
47	MGA_1161	conserved hypothetical protein	584	-	-	x	-	x	x
48	vlhA.5.13	VlhA.5.13 variable lipoprotein family protein	616	-	-	x	-	x	x
49	vlhA.1.04	VlhA.1.04 variable lipoprotein family protein	697	-	-	x	-	x	x
50	MGA_0319	conserved hypothetical lipoprotein	1191	-	-	x	-	x	x
51	vlhA.3.06	VlhA.3.06 variable lipoprotein family protein	688	-	-	x	-	x	x
52	vlhA.3.09	VlhA.3.09 variable lipoprotein family protein	707	-	-	x	-	x	x
53	MGA_0812	unique hypothetical protein	119	-	-	-	x	x	x
54	vlhA.4.03a	VlhA.4.03 variable lipoprotein family protein N-terminal domain protein	197	-	-	-	x	x	x
55	vlhA.4.07.2a	VlhA.4.07.2 variable lipoprotein family protein N-terminal domain protein	191	-	-	-	x	x	x
56	vlhA.4.09	VlhA.4.09 variable lipoprotein family protein	710	-	-	-	x	x	x
57	vlhA.5.01a	VlhA.5.01 variable lipoprotein family domain protein	212	-	-	-	x	x	x
58	vlhA.5.06	VlhA.5.06 variable lipoprotein family protein	703	-	-	-	x	x	x
59	vlhA.5.07	VlhA.5.07 variable lipoprotein family protein	681	-	-	-	x	x	x
60	vlhA.1.01	VlhA.1.01 variable lipoprotein family protein	686	-	-	-	x	x	x
61	vlhA.1.06	VlhA.1.06 variable lipoprotein family protein	754	-	-	-	x	x	x
62	vlhA.3.0.1	VlhA.3.0.1 variable lipoprotein family domain protein	156	-	-	-	x	x	x
63	MGA_0398	macrophage-activating lipoprotein-like protein	584	-	-	-	x	x	x
64	MGA_0485	conserved hypothetical domain protein	341	-	-	-	x	x	x
65	MGA_1318	hypothetical membrane domain protein	113	-	x	-	-	x	-
66	MGA_0656	unique hypothetical lipoprotein	227	-	-	-	-	x	x
67	MGA_0801	Subtilisin-like serine protease domain protein	381	-	-	-	-	x	x
68	vlhA.4.03b	VlhA.4.03 variable lipoprotein family protein C-terminal domain protein	506	-	-	-	-	x	x
69	vlhA.4.07.2b	VlhA.4.07.2 variable lipoprotein family protein C-terminal domain protein	506	-	-	-	-	x	x
70	vlhA.5.01b	VlhA.5.01 variable lipoprotein family domain protein	309	-	-	-	-	x	x



**Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |**

Nº	Gen	Proteína	Talla	Lipoproteína	Hélice α	Barril β	Otras	Citadhesina	Antigénica
71	vlhA.5.10b	VlhA.5.10 variable lipoprotein family protein C-terminal domain protein	77	-	-	-	-	x	x
72	MGA_1351	VlhA variable lipoprotein family protein C-terminal domain protein	86	-	-	-	-	x	x
73	vlhA.1.08b	VlhA1.08 variable lipoprotein family protein C-terminal domain protein	494	-	-	-	-	x	x
74	MGA_0126	unique hypothetical protein	372	-	-	-	-	x	x
75	vlhA.3.0.1	VlhA.3.0.1 variable lipoprotein family domain protein	536	-	-	-	-	x	x
76	vlhA.3.02	VlhA.3.02 variable lipoprotein family protein	646	-	-	-	-	x	x
77	MGA_0702	unique hypothetical protein	64	-	-	-	-	x	-
78	MGA_0625a	ABC-type multidrug/protein/lipid(MdlB-like) transport system component domain protein	305	-	x	-	-	-	x
79	MGA_0626	ABC-type multidrug/protein/lipid(MdlB-like) transport system component	633	-	x	-	-	-	x
80	yidC/oxaA	putative inner membrane protein translocase component YidC	383	-	x	-	-	-	x
81	mnuA	membrane nuclease MnuA	389	-	x	-	-	-	x
82	glpO	glycerol 3-Poxidase	396	-	x	-	-	-	x
83	MGA_0654	ABC transporter permease	1361	-	x	-	-	-	x
84	ftsH	cell division protease ftsH-like protein	765	-	x	-	-	-	x
85	nrdF	Ribonucleoside-diphosphate reductase subunit beta	339	-	x	-	-	-	x
86	secY	preprotein translocase subunit secY	498	-	x	-	-	-	x
87	dnaJ	DnaJ-like molecular chaperone	376	-	x	-	-	-	x
88	MGA_0760	conserved hypothetical membrane protein	352	-	x	-	-	-	x
89	ftsK	DNA translocase ftsK	651	-	x	-	-	-	x
90	cdsA	Phosphatidatecytidyl transferase	372	-	x	-	-	-	x
91	MGA_0789	predicted protease	271	-	x	-	-	-	x
92	MGA_0797	unique hypothetical membrane protein	493	-	x	-	-	-	x
93	MGA_0805	putative ABC-type transport system protein	610	-	x	-	-	-	x
94	MGA_0843	conserved hypothetical protein	219	-	x	-	-	-	x
95	MGA_0849	unique hypothetical protein	676	-	x	-	-	-	x
96	ptsG_1	PTS system glucose-specific II ABC component	891	-	x	-	-	-	x
97	MGA_0867	unique hypothetical protein	182	-	x	-	-	-	x
98	MGA_0879	conserved hypothetical membrane protein	710	-	x	-	-	-	x
99	mgc2	cytadhesin protein MGC2	294	-	x	-	-	-	x
100	crmB	cytadherence related molecule B(CrmB)	921	-	x	-	-	-	x
101	crmC	cytadherence related molecule C(CrmC)	853	-	x	-	-	-	x
102	pyrG	CTP synthase	540	-	x	-	-	-	x
103	MGA_0954	conserved hypothetical membrane protein	976	-	x	-	-	-	x
104	MGA_0957	conserved hypothetical protein	175	-	x	-	-	-	x
105	MGA_1014	unique hypothetical domain protein	451	-	x	-	-	-	x
106	MGA_1037	conserved hypothetical membrane protein	582	-	x	-	-	-	x
107	MGA_1061	cation-transportingP-type ATPase	929	-	x	-	-	-	x
108	MGA_1119	cytadherence related molecule (crmB)-like protein	643	-	x	-	-	-	x

Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |

Nº	Gen	Proteína	Talla	Lipoproteína	Hélice $\alpha$	Barril $\beta$	Otras	Citadhesina	Antigénica
		C-terminal domain protein							
109	dnaJ_4	DnaJ-like molecular chaperone	376	-	x	-	-	-	x
110	MGA_1153	Papain family protease	564	-	x	-	-	-	x
111	atpB	ATP synthase F0a chain	297	-	x	-	-	-	x
112	MGA_1287	ABC-type multidrug transport system permease/ATP-binding protein	600	-	x	-	-	-	x
113	MGA_1302	putative ABC-type transport system permease component	649	-	x	-	-	-	x
114	MGA_1315	conserved hypothetical protein	239	-	x	-	-	-	x
115	MGA_0008	conserved hypothetical membrane protein	242	-	x	-	-	-	x
116	MGA_0022	putative ABC-type bacteriocin/I antibiotic exporter	665	-	x	-	-	-	x
117	MGA_0023	unique hypothetical protein	181	-	x	-	-	-	x
118	plsC	1-acyl-sn-glycerol-3-phosphate acyl transferase	265	-	x	-	-	-	x
119	MGA_0047	LemA-family protein	224	-	x	-	-	-	x
120	MGA_0063	unique hypothetical protein	803	-	x	-	-	-	x
121	MGA_0084	unique hypothetical protein	354	-	x	-	-	-	x
122	MGA_0111	conserved hypothetical membrane protein	200	-	x	-	-	-	x
123	MGA_0122	putative preprotein translocase subunit SecE	77	-	x	-	-	-	x
124	potB	ABC-type spermidine/putrescine transport system permease component I	309	-	x	-	-	-	x
125	cvfA	2',3'-cyclic-nucleotide 2'-phosphodiesterase	593	-	x	-	-	-	x
126	MGA_0173	TlyC-like protein	445	-	x	-	-	-	x
127	MGA_0211	unique hypothetical membrane protein	459	-	x	-	-	-	x
128	MGA_0213	conserved hypothetical membrane protein	218	-	x	-	-	-	x
129	MGA_0217	conserved hypothetical protein	182	-	x	-	-	-	x
130	dppB/oppBb_1	ABC-type dipeptide/oligopeptide/nickel transport system permease protein (DppB/OppB) domain protein	158	-	x	-	-	-	x
131	dppC/oppC_2	ABC-type dipeptide/oligopeptide/nickel transport system permease protein (DppC/OppC)	358	-	x	-	-	-	x
132	dppB/oppB_2	ABC-type dipeptide/oligopeptide/nickel transport system permease protein (DppB/OppB)	350	-	x	-	-	-	x
133	MGA_0241	conserved hypothetical protein	1863	-	x	-	-	-	x
134	MGA_0250	unique hypothetical domain protein	375	-	x	-	-	-	x
135	pvpAa	phase-variant protein A (pvpA) domain protein	175	-	x	-	-	-	x
136	MGA_0281	ion channel-like protein	399	-	x	-	-	-	x
137	mscL	large-conductance mechanosensitive channel MscL	220	-	x	-	-	-	x
138	MGA_0298	unique hypothetical protein	328	-	x	-	-	-	x
139	MGA_0310	putative permease	480	-	x	-	-	-	x
140	MGA_0329	sialidase	938	-	x	-	-	-	x
141	MGA_0335	conserved hypothetical membrane protein	455	-	x	-	-	-	x
142	tktA_2	transketolase	643	-	x	-	-	-	x
143	MGA_0365	ABC-type transport system permease	310	-	x	-	-	-	x
144	MGA_0368	ABC-transport system	550	-	x	-	-	-	x

Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |

Nº	Gen	Proteína	Talla	Lipoproteína	Hélice $\alpha$	Barril $\beta$	Otras	Citadhesina	Antigénica
		permease							
145	MGA_0403	unique hypothetical protein	501	-	x	-	-	-	x
146	gatC	Aspartyl/glutamyl-tRNA amido transferase subunit C- like protein	458	-	x	-	-	-	x
147	MGA_0459	serine/threonine protein kinase	420	-	x	-	-	-	x
148	MGA_0468	major facilitator superfamily transporter	525	-	x	-	-	-	x
149	MGA_0472	unique hypothetical protein	186	-	x	-	-	-	x
150	secE	preprotein translocase subunit SecE	128	-	x	-	-	-	x
151	MGA_0487	conserved hypothetical domain protein	416	-	x	-	-	-	x
152	fruA	PTS system fructose-specific enzyme EIIBC component	677	-	x	-	-	-	x
153	MGA_0512	hypothetical protein	148	-	x	-	-	-	x
154	MGA_0477	conserved hypothetical protein	552	-	x	-	-	-	x
155	MGA_0551	unique hypothetical protein	281	-	x	-	-	-	x
156	MGA_0552	unique hypothetical protein	1095	-	x	-	-	-	x
157	MGA_0557	unique hypothetical protein	802	-	x	-	-	-	x
158	MGA_0558	unique hypothetical protein	280	-	x	-	-	-	x
159	MGA_0600	unique hypothetical protein	299	-	x	-	-	-	x
160	MGA_0676	nuclease-like lipoprotein	276	x	-	-	-	-	x
161	MGA_0808	conserved hypothetical lipoprotein	337	x	-	-	-	-	x
162	MGA_0878	unique hypothetical lipoprotein	401	x	-	-	-	-	x
163	MGA_0274	conserved hypothetical lipoprotein	713	x	-	-	-	-	x
164	MGA_0554	conserved hypothetical lipoprotein	1198	x	-	-	-	-	x
165	MGA_1010	conserved hypothetical protein	1003	-	-	-	x	-	x
166	MGA_1087	conserved hypothetical protein	211	-	-	-	x	-	x
167	MGA_1191	unique hypothetical protein	888	-	-	-	x	-	x
168	MGA_1215	unique hypothetical protein	206	-	-	-	x	-	x
169	MGA_0556	major surface antigen P75	750	-	-	-	x	-	x
170	MGA_0588	conserved hypothetical protein	701	-	-	-	x	-	x
171	MGA_0704	conserved hypothetical membrane protein	293	-	x	-	-	-	x
172	MGA_0830	conserved hypothetical protein	143	-	x	-	-	-	x
173	MGA_0956	conserved hypothetical membrane protein	260	-	x	-	-	-	x
174	potE	putative aminoacid permease	464	-	x	-	-	-	x
175	lspA	Lipoprotein signal peptidase	304	-	x	-	-	-	x
176	hatC	ABC-type high affinity transport system permease HatC	539	-	x	-	-	-	x
177	MGA_1085	conserved hypothetical protein	545	-	x	-	-	-	x
178	MGA_1122	conserved hypothetical membrane protein	228	-	x	-	-	-	x
179	MGA_0021	unique hypothetical protein	141	-	x	-	-	-	x
180	MGA_0046	conserved hypothetical membrane protein	289	-	x	-	-	-	x
181	MGA_0049	hypothetical membrane domain protein	125	-	x	-	-	-	x
182	MGA_0050	hypothetical membrane domain protein	97	-	x	-	-	-	x
183	MGA_0082	unique hypothetical protein	231	-	x	-	-	-	x
184	MGA_0097	unique hypothetical protein	153	-	x	-	-	-	x
185	MGA_0103	conserved hypothetical protein	246	-	x	-	-	-	x
186	MGA_0115	permease	371	-	x	-	-	-	x
187	MGA_0119	conserved hypothetical	210	-	x	-	-	-	x

Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |

Nº	Gen	Proteína	Talla	Lipoproteína	Hélice $\alpha$	Barril $\beta$	Otras	Citadhesina	Antigénica
		protein							
188	potC	ABC-type spermidine/putrescine transport system permease component II	281	-	x	-	-	-	x
189	trkG	Cation transport protein	588	-	x	-	-	-	x
190	dppC/oppC_1	ABC-type dipeptide/oligopeptide/nick el transport system permease protein (DppC/OppC)	379	-	x	-	-	-	x
191	dppB/oppBa_1	ABC-type dipeptide/oligopeptide/nick el transport system permease protein (DppB/OppB) domain protein	212	-	x	-	-	-	x
192	MGA_0280	conserved hypothetical membrane protein	387	-	x	-	-	-	x
193	MGA_0323	conserved hypothetical membrane protein	338	-	x	-	-	-	x
194	MGA_0333	conserved hypothetical membrane protein	300	-	x	-	-	-	x
195	MGA_0436	conserved hypothetical membrane protein	428	-	x	-	-	-	x
196	MGA_0549	unique hypothetical protein	188	-	x	-	-	-	x
197	MGA_1357	conserved hypothetical proteindomain protein	93	-	x	-	-	-	x
198	MGA_0566	unique hypothetical protein	125	-	x	-	-	-	x
199	lgt	Prolipo protein diacylglycerol transferase	369	-	x	-	-	-	x
200	MGA_0809	conserved hypothetical lipoprotein	332	x	-	-	-	-	x
201	MGA_0810	conserved hypothetical protein	293	-	-	-	x	-	x
202	MGA_0579	conserved hypothetical lipoprotein	791	x	-	-	-	-	x
203	MGA_0829a	conserved hypothetical lipo domain protein	481	-	-	-	x	-	x
204	MGA_0493	conserved hypothetical protein	611	-	-	-	x	-	x
205	MGA_0754	hypothetical protein	1128	-	x	-	-	-	x
206	MGA_0800	unique hypothetical membrane protein	483	-	x	-	-	-	x
207	MGA_0816	unique hypothetical membrane protein	492	-	x	-	-	-	x
208	MGA_1335	hypothetical protein	88	-	x	-	-	-	x
209	MGA_0889	conserved hypothetical domain protein	539	-	x	-	-	-	x
210	MGA_0907	unique hypothetical membrane protein	488	-	x	-	-	-	x
211	MGA_1011	conserved hypothetical protein	1168	-	x	-	-	-	x
212	MGA_1071	conserved hypothetical protein	512	-	x	-	-	-	x
213	MGA_1103	predicted integral membrane methylase-domain protein	643	-	x	-	-	-	x
214	MGA_1349	hypothetical domain protein	62	-	x	-	-	-	x
215	MGA_1271	conserved hypothetical protein	121	-	x	-	-	-	x
216	MGA_0004	conserved hypothetical protein	662	-	x	-	-	-	-
217	MGA_0048	conserved hypothetical protein	865	-	x	-	-	-	-
218	MGA_1317	hypothetical membrane domain protein	49	-	x	-	-	-	-
219	MGA_0112	unique hypothetical protein	93	-	x	-	-	-	-
220	MGA_0137	unique hypothetical protein	494	-	x	-	-	-	-
221	MGA_0244	unique hypothetical protein	292	-	x	-	-	-	-
222	MGA_0516	unique hypothetical protein	480	-	x	-	-	-	-
223	pstS	ABC-type phosphate transport system periplasmic phosphate binding protein	385	x	-	-	-	-	-
224	MGA_0811	unique hypothetical lipoprotein	308	x	-	-	-	-	-

**Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |**

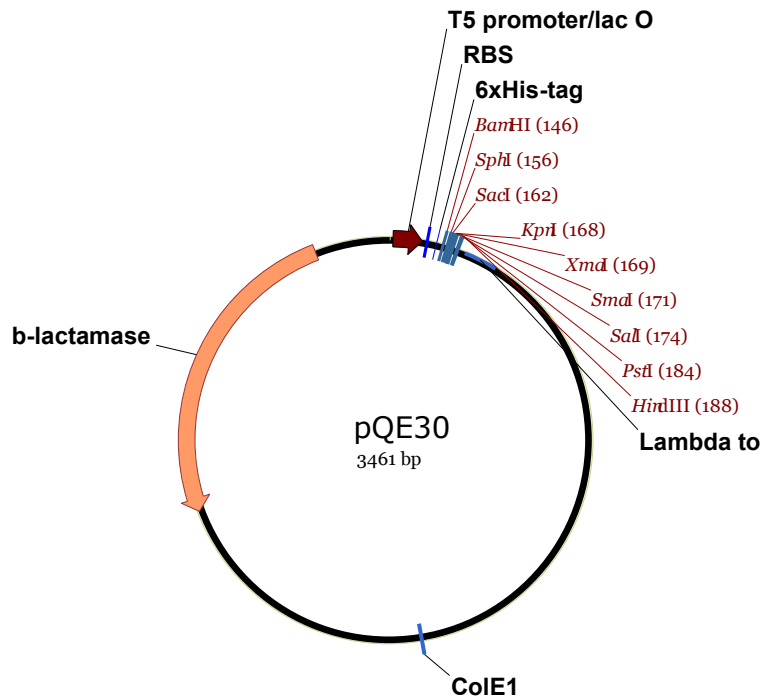
Nº	Gen	Proteína	Talla	Lipoproteína	Hélice $\alpha$	Barril $\beta$	Otras	Citadhesina	Antigénica
225	MGA_0887	unique hypothetical lipoprotein	396	x	-	-	-	-	-
226	MGA_1073	conserved hypothetical lipoprotein	607	x	-	-	-	-	-
227	vlhA.1.08a	VlhA1.08 variable lipoprotein family proteinN-terminal domain protein	98	x	-	-	-	-	-
228	vlhA.2.01	VlhA.2.01 variable lipoprotein family protein	607	x	-	-	-	-	-
229	MGA_1360	hypothetical lipoprotein	133	x	-	-	-	-	-
230	MGA_0495	conserved hypothetical lipoprotein	679	x	-	-	-	-	-
231	MGA_0798	Subtilisin-like serine protease	647	-	-	x	-	-	-
232	MGA_0906	Subtilisin-like serine protease	649	-	-	x	-	-	-
233	oppA	putative ABC-type oligo peptide transport solute binding protein OppA	1034	-	-	x	-	-	-
234	MGA_0321	conserved hypothetical lipoprotein	1150	-	-	x	-	-	-
235	MGA_0802	Subtilisin-like serine protease domain protein	249	-	-	-	x	-	-
236	hatA	ABC-type high affinity transport system protein HatA	366	-	-	-	x	-	-
237	MGA_1031	conserved hypothetical protein	199	-	-	-	x	-	-
238	MGA_1107	conserved hypothetical RmuC-domain protein	499	-	-	-	x	-	-
239	rpml	50S ribosomal protein L35	64	-	-	-	x	-	-
240	MGA_0131	putative ABC-type spermidine/putrescine-binding periplasmic transport protein potD	479	-	-	-	x	-	-
241	MGA_0332	conserved hypothetical lipoprotein	273	-	-	-	x	-	-
242	MGA_0518	Subtilisin-like serine protease domain protein	276	-	-	-	x	-	-
243	MGA_0562	conserved hypothetical protein	674	-	-	-	x	-	-
244	MGA_0576	unique hypothetical protein	691	-	-	-	x	-	-
245	MGA_0584	conserved hypothetical domain protein	243	-	-	-	x	-	-
246	MGA_0586	conserved hypothetical protein	705	-	-	-	x	-	-
247	MGA_1322d	ABC transporter component domain protein	64	-	x	-	-	-	-
248	glpF	glycerol uptake facilitator protein GlpF	243	-	x	-	-	-	-
249	MGA_0662	unique hypothetical protein	296	-	x	-	-	-	-
250	ugpA	ABC-type sn-glycerol-3-phosphate transport system permease protein ugpA	329	-	x	-	-	-	-
251	ugpE	ABC-type sn-glycerol-3-phosphate transport system permease protein ugpE	325	-	x	-	-	-	-
252	pstC/pstA	ABC-type phosphate transport system permease pstC/pstA	652	-	x	-	-	-	-
253	MGA_0817a	unique hypothetical domain protein	48	-	x	-	-	-	-
254	MGA_0817b	unique hypothetical membrane domain protein	443	-	x	-	-	-	-
255	MGA_0871	unique hypothetical protein	69	-	x	-	-	-	-
256	MGA_0875	unique hypothetical protein	63	-	x	-	-	-	-
257	MGA_0892	unique hypothetical protein	145	-	x	-	-	-	-
258	MGA_0901	unique hypothetical protein	213	-	x	-	-	-	-
259	MGA_0908	unique hypothetical membrane domain protein	231	-	x	-	-	-	-
260	MGA_0911	unique hypothetical membrane domain protein	269	-	x	-	-	-	-
261	MGA_1046	unique hypothetical membrane protein	491	-	x	-	-	-	-
262	pgsA	CDP-diacylglycerol-glycerol-3-phosphate3-phosphatidyl	232	-	x	-	-	-	-

**Anexo 3. PROTEÍNAS POSITIVAS POR Bact-VAP,  
ORGANIZADAS SEGÚN SU PROBABILIDAD DE SER  
ANTÍGENOS DE SUPERFICIE (Listado de Priorización) |**

Nº	Gen	Proteína	Talla	Lipoproteína	Hélice $\alpha$	Barril $\beta$	Otras	Citadhesina	Antigénica
		transferase							
263	MGA_1077	ABC-type sugar transport system permease protein	319	-	x	-	-	-	-
264	MGA_1078	ABC-type sugar transport system permease protein	279	-	x	-	-	-	-
265	MGA_1089	ABC transporter permease domain protein	516	-	x	-	-	-	-
266	MGA_1113	unique hypothetical domain protein	111	-	x	-	-	-	-
267	MGA_1140	putative ABC-type multidrug transport system permease component	368	-	x	-	-	-	-
268	atpI	putative ATP synthase protein I	169	-	x	-	-	-	-
269	atpE	ATP synthase C chain	96	-	x	-	-	-	-
270	atpF	ATP synthase B chain	198	-	x	-	-	-	-
271	MGA_1188	Aminoacid permease	586	-	x	-	-	-	-
272	MGA_1260	Ribonuclease HIII-domain protein	148	-	x	-	-	-	-
273	MGA_1350b	hypothetical domain protein	73	-	x	-	-	-	-
274	MGA_1350a	hypothetical domain protein	55	-	x	-	-	-	-
275	MGA_1284	ABC transporterATP-binding proteinN-terminal domain protein	312	-	x	-	-	-	-
276	MGA_0002	conserved hypothetical protein	619	-	x	-	-	-	-
277	MGA_0287	aminoacid permease	584	-	x	-	-	-	-
278	ptsG_2	putative PTS systemglucose-specific enzyme IIB PtsG	125	-	x	-	-	-	-
279	MGA_0313	conserved hypothetical protein, DUF1600 family	268	-	x	-	-	-	-
280	MGA_0315	conserved hypothetical protein, DUF1600 family	265	-	x	-	-	-	-
281	MGA_0316	conserved hypothetical protein, DUF1600 family	258	-	x	-	-	-	-
282	MGA_0324	unique hypothetical protein	85	-	x	-	-	-	-
283	MGA_0337	conserved hypothetical membrane protein	375	-	x	-	-	-	-
284	MGA_0340	conserved hypothetical protein	73	-	x	-	-	-	-
285	cbiQ	ABC-type cobalt transport permease protein	389	-	x	-	-	-	-
286	MGA_0409	conserved hypothetical protein	582	-	x	-	-	-	-
287	MGA_0427	putative esterase/lipase	314	-	x	-	-	-	-
288	MGA_1354	hypothetical protein	50	-	x	-	-	-	-
289	coaE	Dephospho-CoAkinase	200	-	-	-	x	-	-
290	MGA_1083	HAD super family hydrolase Cof	282	-	-	-	x	-	-
291	MGA_1117	cytadherence-related molecule (crmB)-like protein N-terminal domain protein	342	-	-	-	x	-	-

\* Nota: En rojo se señalan las proteínas seleccionadas para la evaluación

## Anexo 4. MAPA pQE30



Mapa del plásmido pQE30

Posiciones de los elementos en bases

Talla del vector	3461
Inicio de la numeración en <i>Xho</i> I (CTCGAG)	1-6
Promotor T5/Operador lac	7-87
Inicio de la transcripción de T5	61
Secuencia codificante para el marcador de 6xHis	127-144
Sitio de clonaje múltiple	145-192
Región de terminación de la transcripción Lambda T <sub>0</sub>	208-302
Región de terminación de la transcripción <i>rnnB</i> T1	1064-1162
Origen de replicación ColE1	1638
Secuencia codificante para la β-lactamasa	3256-2396

## Anexo 5. PREPARACIÓN DE MEDIOS Y SOLUCIONES

### Cultivo de *Mycoplasma gallisepticum*

- Rojo Fenol (1%)

Rojo Fenol      250mg

NaOH (1N)      7,5mL

H<sub>2</sub>O<sub>dd</sub>\*      17,5mL

\* dd-bidestilada

- Medio Hayflick Caldo

PPLO caldo      5,25g

Extracto de Levadura      25mg

H<sub>2</sub>O<sub>dd</sub>      200mL

Ajustar a pH 7,8

Autoclavear a 115°C durante 15 minutos

Adicionar estérilmente

Suero equino\*      50mL

Ampicilina (100mg/mL)      0,25mL

Rojo Fenol (1%)      0,5mL

\* Suero equino filtrado e inactivado a 56°C durante 30min

- Medio Hayflick Agar

PPLO agar      8,75g

Extracto de Levadura      25mg

H<sub>2</sub>O<sub>dd</sub>      200mL

Ajustar a pH 7,8

Autoclavear a 115°C durante 15 minutos

Adicionar estérilmente



Suero equino <sup>1</sup>	50mL
Ampicillina (100mg/mL)	0,25mL

<sup>1</sup>Suero equino filtrado e inactivado a 56°C durante 30min

### **Cultivo de *E. coli***

- Medio Luria Bertani (LB)

Volumen final: 1L

Reactivos	Cantidades
NaCl	10g
Triptona	10g
Extracto de levadura	5g
H <sub>2</sub> O <sub>dd</sub>	1L

Para el medio sólido se adicionan 15g de agar técnico

Autoclavear a 115°C durante 15 minutos

- Medio Super Broth (SB)

Volumen final: 1L

Reactivos	Cantidades
NaCl	5g
Triptona	35g
Extracto de levadura	20g
NaOH 2N	2,5 mL
H <sub>2</sub> O <sub>dd</sub>	1L

Para el medio sólido se adicionan 15g de agar técnico

Autoclavear a 115°C durante 15 minutos

### **Extracción de ADN cromosomal de *M. gallisepticum***

- Tampón de lisis

Volumen final: 250mL

Reactivos	Cantidades	Concentración Final
SDS	2,5g	1%
Tris-HCl (pH 8, 1M)	12,5mL	50mM
EDTA	0,73g	10mM
H <sub>2</sub> O <sub>dd</sub>	250mL	-

- Proteinasa K (5mL)

Reactivos	Cantidades	Concentración Final
Proteinasa K	500mg	100mg/mL
H <sub>2</sub> O <sub>dd</sub>	5mL	-

Acetato de Potasio (KAc) (5mL)

Reactivos	Cantidades	Concentración Final
KAc	6,65g	5M
H <sub>2</sub> O <sub>dd</sub>	5mL	-

### Extracción de ADN plasmídico de *E. coli* por lisis alcalina

- Tampón de resuspensión

Volumen final: 100mL

Reactivos	Cantidades	Concentración
Tris-HCl	0,394g	25mM
EDTA	0,29g	10mM
H <sub>2</sub> O <sub>dd</sub>	100mL	-

Ajustar a pH 8

Añadir 10mL de ARNsaA (10mg/mL)

- Tampón de lisis

Volumen final: 100mL

Reactivos	Cantidades	Concentración
NaOH (400mM)	50mL	200mM
SDS (2%)	50mL	1%

- Tampón de neutralización

Volumen final: 100mL

Reactivos	Cantidades	Concentración
KAc	29,45g	3M
H <sub>2</sub> O <sub>dd</sub>	100mL	-
Ajustar a pH 5.5		

### Electroforesis de ADN

- Tampón de corrida TAE 10x

Volumen final: 200mL

Reactivos	Cantidades
SDS	37,22g
Ácido acético glacial	11,42mL
Tris	24,22g
EDTA	7,44g
H <sub>2</sub> O <sub>dd</sub>	188,6mL

Ajustar a pH 7,4 ÷ 7,8

- Preparación de geles de agarosa 1%

Volumen final: 100mL

Reactivos	Cantidades
Agarosa	1g
TE 10x	10mL
Bromuro de etidio (10mg/mL)	5µL
H <sub>2</sub> O <sub>dd</sub>	90mL

La agarosa se funde y se prepara el gel en su molde

La electroforesis se corre en TE 1x

### Electroforesis de Proteínas y Transferencia

- Tampón de muestra 5x

Volumen final: 10mL

Reactivo	Cantidad	Concentración final
Tris-HCl 1M, pH 6,8	1,54mL	154mM
$\beta$ -mercaptoetanol	2,5mL	3,5M
SDS	1g	10%
Glicerol	5mL	50%
Bromofenol azul	5mg	0,05%
H <sub>2</sub> O <sub>dd</sub>	0,96mL	

- Solución de tinción de geles con Coomasie

Volumen final: 100mL

Reactivo	Cantidad	Concentración final
Coomasie Brilliant Blue R-250	0,1g	0,1%
HAc	10mL	10%
EtOH	40mL	40%
H <sub>2</sub> O <sub>dd</sub>	50mL	

- Preparación de los geles (para dos geles)

	Gel separador	Gel concentrador
% de Acrilamida	12%	5%
H <sub>2</sub> O <sub>dd</sub>	1,73mL	1,2mL
Acrilamida 40% - Bis-Acrilamida 1%	1,2mL	0,25mL
S/Tris 1,5M, pH 8,8	1mL	--
C/Tris 1M, pH 6,8	--	0,5mL
Persulfato de amonio 10%	64 $\mu$ L	40 $\mu$ L
TEMED	5 $\mu$ L	5 $\mu$ L

- Reactivos del tampones de los geles separador y concentrador

gel separador 8x

gel concentrador

pH 9,18 (1000mL)			pH 6,14 (100mL)		
Reactivo	Cantidad	Concentración	Reactivo	Cantidad	Concentración
Tris	205,6g	1,7M	Tris	2,62g	0,22M

- Tampón Tris-Ácido Bórico 10x (pH 8,64)

Volumen final: 1000mL

Reactivo	Cantidad
Tris	4,92g
Ácido Bórico	2,47g
H <sub>2</sub> O <sub>dd</sub>	1000mL

- Tampón de la cámara concentradora 1x

Volumen final: 100mL

Reactivo	Cantidad
SDS 10%	1mL
Tris-Ácido Bórico 10x	10mL
H <sub>2</sub> O <sub>dd</sub>	89mL

- Tampón de la cámara separadora 1x

Volumen final: 100mL

Reactivo	Cantidad
Tampón gel separador 8x	25mL
H <sub>2</sub> O <sub>dd</sub>	75mL

- Tampón de transferencia

Volumen final: 500mL

Reactivo	Cantidad	Concentración final
Tris	1,41g	20mM
Glicina	6,63g	150mM
SDS 20%	50μl	0,002%
Metanol	100mL	20%
H <sub>2</sub> O <sub>dd</sub>	420mL	

### **Purificación por Cromatografía de interacción con Quelatos Metálicos (IMAC, por su siglas en inglés)**

- Solución madre Imidazol

Reactivo	Cantidad	Concentración
----------	----------	---------------

Imidazol	6,8mg	1M
H <sub>2</sub> O <sub>dd</sub>	100mL	

- Solución madre NaCl

Reactivo	Cantidad	Concentración
NaCl	35,1mg	3M
H <sub>2</sub> O <sub>dd</sub>	200mL	

- Tampón fosfato salino (PBS) 500mM (10X)

Volumen final: 1L

Reactivo	Cantidad
Na <sub>2</sub> HPO <sub>4</sub>	11,5g
KH <sub>2</sub> PO <sub>4</sub>	2g
NaCl	80g
KCl	2g
H <sub>2</sub> O <sub>dd</sub>	1L

Ajustar a pH 7,2 ÷ 7,4

- Tampón de equilibrio

Reactivo	Cantidad	Concentración
PBS 10X	100mL	50mM
Imidazol 1M	10mL	10mM
NaCl 3M	100mL	300mM
H <sub>2</sub> O <sub>dd</sub>	790mL	

- Tampón de lavado

Reactivo	Cantidad	Concentración
PBS 10X	100mL	50mM
Imidazol 1M	20mL	20mM
NaCl 3M	100mL	300mM
H <sub>2</sub> O <sub>dd</sub>	780mL	

- Tampón de elución

Reactivo	Cantidad	Concentración
PBS 10X	100mL	50mM
Imidazol 1M	250mL	250mM
NaCl 3M	100mL	300mM
H <sub>2</sub> O <sub>dd</sub>	550mL	

### Cuantificación de proteínas totales solubles (Método de Bradford)

- Reactivo de Bradford

Volumen final: 1L

Reactivo	Cantidad
Coomasie Brilliant Blue G (CBB-G)	40mg
Etanol	40mL
H <sub>3</sub> PO <sub>4</sub> 85%	100mL
SDS	30mg
H <sub>2</sub> O <sub>dd</sub>	860mL

Disolver 40mg de CBB-G en 40mL de EtOH y agitar 30 minutos moderadamente.

Añadir 100mL de H<sub>3</sub>PO<sub>4</sub> (85%), Agitar 5 minutos.

Adicionar 760mL de H<sub>2</sub>O<sub>dd</sub>.

Disolver 30mg de SDS en 100mL de H<sub>2</sub>O<sub>dd</sub> y añadir a la solución de CBB-G y H<sub>3</sub>PO<sub>4</sub>.

Llevar el volumen de la solución a 1L con H<sub>2</sub>O<sub>dd</sub>.

Guardar en frasco ámbar.

### Técnicas de inmuno-detección

#### ELISA

- Tampón de cubrimiento (BC)

Para volumen final 100mL

Reactivo	Cantidad
Na <sub>2</sub> CO <sub>3</sub>	159mg
NaHCO <sub>3</sub>	293mg
H <sub>2</sub> O <sub>dd</sub>	100mL

Ajustar a pH 9,6

- Tampón fosfato salino (PBS)

Volumen final: 1L

Reactivo	Cantidad
Na <sub>2</sub> HPO <sub>4</sub>	1,15g
KH <sub>2</sub> PO <sub>4</sub>	0,2g
NaCl	8g
KCl	0,2g
H <sub>2</sub> O <sub>dd</sub>	1L

Ajustar a pH 7,2 ÷ 7,4

- Solución PBS-Tween 20 0,05%-BSA 1%

Para volumen final 100mL

Reactivo	Cantidad
PBS	100mL
Tween 20	50 $\mu$ L
BSA	1g

Tampón sustrato TMB

Reactivo	Cantidad
Na <sub>2</sub> HPO <sub>4</sub>	7,66g
C <sub>6</sub> H <sub>8</sub> O <sub>7</sub> x H <sub>2</sub> O (ácido cítrico)	10,3g
H <sub>2</sub> O <sub>2</sub> 30%	467 $\mu$ L
H <sub>2</sub> O <sub>dd</sub>	1L

Ajustar a pH 5 antes de adicionar el H<sub>2</sub>O<sub>2</sub>

- Solución de parada

Volumen final: 200mL

Reactivo	Cantidad	Concentración Final
H <sub>2</sub> SO <sub>4</sub> ( $\rho$ =1,84Kg/L; 95-97%)	16,5mL	3N (1,5M)
H <sub>2</sub> O <sub>dd</sub>	183,5mL	

*Western-Blot*

- Tampón sustrato Fosfatasa Alcalina

Reactivo	Cantidad	Concentración Final
Tris	12,1g	100mM
NaCl	8,7g	150mM
MgCl <sub>2</sub>	203mg	1mM
H <sub>2</sub> O <sub>dd</sub>	1L	

Ajustar a pH 9



## Anexo 6. SECUENCIAS DE LOS CEBADORES, GENES CLONADOS Y LAS PROTEÍNAS DEDUCIDAS

Cebadores para la amplificación de los genes que codifican para las 50 primeras proteínas (no descritas previamente) de la lista de priorización y seis conocidas proteínas de superficie y/o factores de virulencia de *M. gallisepticum*, y talla esperada de los productos de PCR.

Gen (GenBank ID)	Cebador	Secuencia	Talla Producto
MGA_0008	MGA_0008F MGA_0008R	ACATGCATGCATGAATGTAGCAGGATATAT TCCCCCGGGTTAAACCTTTTTGTAACTCT	748
MGA_0023	MGA_0023F MGA_0023R	ACATGCATGCATGTCAAATAACGATTCTGA TCCCCCGGGTTATAAAGCTGATACGTTAC	565
MGA_0063	MGA_0063F MGA_0063R	ACATGCATGCTTCTTAGGTAAGAACTACATTAGA TCCCCCGGGCTCTTCTGAGGTTGATCAA	1987
MGA_0084	MGA_0084F MGA_0084R	ACATGCATGCATGAGCATTATGTTTAAAGATCAGCT TCCCCCGGGTTATTTTTTATTACATTCT	1084
MGA_0111	MGA_0111F MGA_0111R	ACATGCATGCATGGAGTATTTTAAAGCTC TCCCCCGGGTTATTTTCTTATCAGAGA	622
MGA_0126	MGA_0126F MGA_0126R	ACATGCATGCATGAATCCTTTCTCACAAATG TCCCCCGGGTTATTTTCATTTTTCTTTGT	1138
MGA_0173	MGA_0173F MGA_0173R	ACATGCATGCATGACATCTGAAGAGAGTTC TCCCCCGGGTTAATAATTTGTGCTTATTT	1357
MGA_0211	MGA_0211F MGA_0211R	ACATGCATGCATGGTTAATGATAATAGTACA TCCCCCGGGTTATTCCTTAGAATCTTTAA	1399
MGA_0213	MGA_0213F MGA_0213R	ACATGCATGCATGATGAAGTGAAAAAATG TCCCCCGGGTTAAAGTTCATTTAATCCAC	676
MGA_0217	MGA_0217F MGA_0217R	ACATGCATGCATGGGTAATCAAAAGCGCTAAAACTGGTTATAA TCCCCCGGGTTAAAAATAACTAAAAAGAAA	568
MGA_0226	MGA_0226F MGA_0226R	ACATGCATGCGCTATGAGTAGAACTAAATACAACCT TCCCCCGGGTTAAAGTGCAGTGTTAACGT	1807
MGA_0241	MGA_0241F MGA_0241R	ACATGCATGCTTAAATCCTGAACCTAACGAT TCCCCCGGGATTTGAAGCAATACCATTAA	3019
MGA_0250	MGA_0250F MGA_0250R	ACATGCATGCTTGGGATATTGTATCTCAGC TCCCCCGGGTTATTTATTTAATTTGTTCT	1147
MGA_0267	MGA_0267F MGA_0267R	ACATGCATGCTCGGAAGCTATGAGTAGAAC TCCCCCGGGTTAAAGTGCAGTGTTAACGT	1813
MGA_0281	MGA_0281F MGA_0281R	ACATGCATGCATGAATTTAGGGTTTAGCAA TCCCCCGGGTTAATTTGTTTAATTGGAA	1219
MGA_0298	MGA_0298F MGA_0298R	ACATGCATGCATGAATAATGAGCAAACTAAACGCT TCCCCCGGGTTATTTTATTGATATAACTT	1006
MGA_0310	MGA_0310F MGA_0310R	ACATGCATGCATGAATAATCAATTCGATTG TCCCCCGGGTTAAATACAAGATTTAAAG	1462
MGA_0319	MGA_0319F MGA_0319R	ACATGCATGCAAGAGAATCTATGGGAATAA TCCCCCGGGGCTTAGATCAAAATCATATC	1243
MGA_0335	MGA_0335F MGA_0335R	ACATGCATGCGTGTTAATTAATATGTCAATAGA TCCCCCGGGCTAGAATACTCTGATTAG	1387
MGA_0398	MGA_0398F MGA_0398R	ACATGCATGCTGTGGTGTCTTCTTCTACTAACATA TCCCCCGGGTTAGTTTAATTTCTTGAAGA	1687
MGA_0403	MGA_0403F MGA_0403R	ACATGCATGCTTGTCTTATTTTTTAATTATTTTA TCCCCCGGGTTATTTAAGTAATTTGATTG	1525
MGA_0416	MGA_0416F MGA_0416R	ACATGCATGCTCGTCTGTCTGTTTTTCAGACGGAATAATATGAC TCCCCCGGGTTAGTTAGTAGCTGGTGAAG	1513
MGA_0472	MGA_0472F MGA_0472R	ACATGCATGCATGGGAATTAACGATACAAA TCCCCCGGGTTATTTGTCTTCTGTTTTAA	580
MGA_0477	MGA_0477F MGA_0477R	ACATGCATGCATGTCTGTAAAAAATTATCAATTA TCCCCCGGGTTATGAATTAACCTCTGACT	1678

Anexo 6. SECUENCIAS DE LOS CEBADORES, GENES  
CLONADOS Y LAS PROTEÍNAS DEDUCIDAS |

Gen (GenBank ID)	Cebador	Secuencia	Talla Producto
MGA_0485	MGA_0485F MGA_0485R	ACATGCATGCAGTGGGGTAATTACTCATCA TCCCCCGGGTTACCTTCAACTGGAACAGT	964
MGA_0487	MGA_0487F MGA_0487R	ACATGCATGCTTGAAAACTTGTACCACC TCCCCCGGGTTAAAATACTCCGATGATT	1270
MGA_0565	MGA_0565F MGA_0565R	ACATGCATGCAACAAATAGATGATAAAAT TCCCCCGGGTTAATTTTGAATTTAAAT	2314
MGA_0648	MGA_0648F MGA_0648R	ACATGCATGCCAGAAAATGCAAGCTCCAGCAACAAAAAGTGCAGAGGA TCCCCCGGGTTAGAAACCGTTAGGGAAAA	2200
MGA_0656	MGA_0656F MGA_0656R	ACATGCATGCATGTTTTAAGTAGTTGTACTGG TCCCCCGGGTTACAGCTCTGTTTGCTCTT	703
MGA_0702	MGA_0702F MGA_0702R	ACATGCATGCATGGTTTTATACGGTGAACCTTTAATAATG TCCCCCGGGTTATTTTTATTGACATTAC	214
MGA_0760	MGA_0760F MGA_0760R	ACATGCATGCATGAACAATATCTTTATTACGCC TCCCCCGGGTTATTTAAATTTGTAAGCA	1078
MGA_0789	MGA_0789F MGA_0789R	ACATGCATGCATGAATCCTTCTTATTATACTAAC TCCCCCGGGCTAACCGATGATTATTGTAT	835
MGA_0797	MGA_0797F MGA_0797R	ACATGCATGCATGAAAAATATTTAAGATTATCA TCCCCCGGGTTAAATTAACCGGAGTTAA	1501
MGA_0801	MGA_0801F MGA_0801R	ACATGCATGCGTGGCTTCAATTTTACGGGGTTAACGGTG TCCCCCGGGTTATTTGTTATCGATCGTTC	1165
MGA_0812	MGA_0812F MGA_0812R	ACATGCATGCGTCTTAAAAAACAAGTAAGG TCCCCCGGGCTATTCTCCAATAACAATTC	325
MGA_0843	MGA_0843F MGA_0843R	ACATGCATGCATGAATAGTAGTTTCCTTTT TCCCCCGGGTTAAATTTCTTTTAGCTT	679
MGA_0849	MGA_0849F MGA_0849R	ACATGCATGCATGTTCAAGAAAAATTTT TCCCCCGGGCTAGGCATGGTTATTCTCTT	2050
MGA_0867	MGA_0867F MGA_0867R	ACATGCATGCTTGAAGGTGGTAGACACCAAGAAGAGTG TCCCCCGGGTTAAACTTTTTTAGTTTGGG	568
MGA_0879	MGA_0879F MGA_0879R	ACATGCATGCGTGAATACAGGCTATAGCCAAATTGATAAACAAG TCCCCCGGGTTAATTTTGGTTCTGATTTA	2152
MGA_0913	MGA_0913F MGA_0913R	ACATGCATGCACAGTAATTAATTCAGGGT TCCCCCGGGTTAGTTTCTTGATAGATAC	754
MGA_0954	MGA_0954F MGA_0954R	ACATGCATGCATAAAAACTGCCGAACCTAAC TCCCCCGGGTTGGCTTAGATTAATAAATG	1411
MGA_0957	MGA_0957F MGA_0957R	ACATGCATGCATGAGTAATCACAAACCCGATTCT TCCCCCGGGTTAAGTATTAGGATTAACCT	547
MGA_0993	MGA_0993F MGA_0993R	ACATGCATGCACTAGAGCGTCACAACCCGT TCCCCCGGGTTATTGATCAGCTGTTCCTA	1813
MGA_1014	MGA_1014F MGA_1014R	ACATGCATGCATGACACCATATCAAAAAGAG TCCCCCGGGTTAAAGTTGATCACTCAGA	1375
MGA_1037	MGA_1037F MGA_1037R	ACATGCATGCATGCCCTTAAATACTAATAATGAA TCCCCCGGGTTATTGTTTGTCTTGACCT	1768
MGA_1079	MGA_1079F MGA_1079R	ACATGCATGCGAATTAATACTAATGAAAA TCCCCCGGGATCATTATCATCACTTAAT	3019
MGA_1161	MGA_1161F MGA_1161R	ACATGCATGCGTTTTTAATTTTTTCAGCAACCTACCAACAAC TCCCCCGGGTTAATTTTCAAGTTAGGTAAT	1687
MGA_1216	MGA_1216F MGA_1216R	ACATGCATGCATGAAGAAATTAACACAAGTGA TCCCCCGGGTTAAATCCTGCATAGATTG	310
MGA_1315	MGA_1315F MGA_1315R	ACATGCATGCATGTATGTTATTTTTCATAACGAGATCACCAACACCAAAG TCCCCCGGGTTATAGGTCAGTAAGATCCA	739
MGA_1318	MGA_1318F MGA_1318R	ACATGCATGCTTGCTGTTTTTAATTAATCTTAT TCCCCCGGGTTAATTTAAATTAGTCTTA	361
<i>gapA</i>	<i>gapAF</i> <i>gapAR</i>	ACATGCATGCTCTGGAGCTAAGCCAAATAA TCCCCCGGGATAACCTGGAAGCGCATCAA	2926
<i>crmA</i>	<i>crmAF</i> <i>crmAR</i>	ACATGCATGCAAGCAATCAGATAAGAGTAAC TCCCCCGGGTGATGGCACTACTCATGAAG	2734
<i>crmC</i>	<i>crmCF</i> <i>crmCR</i>	ACATGCATGCAATCATTCTTCTAGATCAGA TCCCCCGGGCACAAATAAACGGGATAAG	1984
<i>mgc2</i>	<i>mgc2F</i> <i>mgc2R</i>	ACATGCATGCGAATTCACCCACACTTGT TCCCCCGGGTTATCTAGGTCCATTTTGTG	796
<i>oppA</i>	<i>oppAF</i> <i>oppAR</i>	GGGGTACCATTACTAATGATCATGTAAT CCGAGCTCAGAGCTTATGGAACCCACAA	859
<i>vlhA.5.02</i>	<i>vlhA.5.02F</i> <i>vlhA.5.02R</i>	ACATGCATGCATGTCTATTACTAAAAAGACTCAAACCC TCCCCCGGGCTACGTTAGATGGAGAGTAA	803
MGA_0021	MGA_0021F MGA_0021R	ACATGCATGCTTGAACATGATCATAGCTTT TCCCCCGGGTTATAGATTAATCCATTTTA	445
MGA_0046	MGA_0046F MGA_0046R	ACATGCATGCTTGATGAGAAGCTTAAATATTCTTACG TCCCCCGGGTTAGATCGTAACACGGTGTT	889
MGA_0049	MGA_0049F MGA_0049R	ACATGCATGCATGGTTGGGTTTATAGCTTT TCCCCCGGGTTAATCGTTTATGTCAAAC	397
MGA_0050	MGA_0050F MGA_0050R	ACATGCATGCATGAGAAATAAAGACATCTATAA TCCCCCGGGTTAAATTTTCGACTTAAAA	313

Anexo 6. SECUENCIAS DE LOS CEBADORES, GENES  
CLONADOS Y LAS PROTEÍNAS DEDUCIDAS |

Gen (GenBank ID)	Cebador	Secuencia	Talla Producto
MGA_0082	MGA_0082F MGA_0082R	ACATGCATGCATGGTGCCACAAAGATTAGTCCCAAT TCCCCCGGGCTAACTAAAAATAACTCCTAG	715
MGA_0097	MGA_0097F MGA_0097R	ACATGCATGCATGAGATCCAATTGGAACCCCTAGCCA TCCCCCGGGTTATTTAACATCTTGATTGA	481
MGA_0103	MGA_0103F MGA_0103R	ACATGCATGCATGTTTAAATTCACACACGA TCCCCCGGGTTAATGCTTAGGTTGTTTCT	760
MGA_0115	MGA_0115F MGA_0115R	ACATGCATGCATGTCAAAAAAAGTATTAGTAAAAAGTTCGAGCAAGTGC TCCCCCGGGTTAAACCATTCCGCTTGTAT	1135
MGA_0119	MGA_0119F MGA_0119R	ACATGCATGCATGCAATAACAACCTTACTGA TCCCCCGGGTTAAAAGAAGTCACGTGGTT	652
MGA_0274	MGA_0274F MGA_0274R	ACATGCATGCTTTAACCCCTGAAAGTTTCTT TCCCCCGGGCTAAGGATGACTAAATAGAT	2089
MGA_0280	MGA_0280F MGA_0280R	ACATGCATGCTTGAATAATCAAAATATTTT TCCCCCGGGCTAAGTAATAATTACAGTAC	1183
MGA_0323	MGA_0323F MGA_0323R	ACATGCATGCTTGCTGCGAGTGTGGTGAAT TCCCCCGGGTTATCAGCGTAAGCGTGATA	1036
MGA_0337	MGA_0337F MGA_0337R	ACATGCATGCTTGGATATGGAACAGACTTC TCCCCCGGGTTAATTAAACCTTTTGATTCT	1147
MGA_0436	MGA_0436F MGA_0436R	ACATGCATGCATGTCCTTAAAAAAGTCAAAAGAT TCCCCCGGGTTATTCTTCAGTAATATAGA	1306
MGA_0512	MGA_0512F MGA_0512R	ACATGCATGCATGAAAAAGAAGATATTTAA TCCCCCGGGTTAGGCTTGTGCATCTTCTT	466
MGA_0549	MGA_0549F MGA_0549R	ACATGCATGCGTGAAGATTGATCAGGTAAC TCCCCCGGGCTAATTTTTTCTTTAATAA	586
MGA_0551	MGA_0551F MGA_0551R	ACATGCATGCGTGAAATCAAAAAAGTTAATAAG TCCCCCGGGTTATCTCTTTTGGTGATG	865
MGA_0552	MGA_0552F MGA_0552R	ACATGCATGCATGGCGGATCGACCTATGAA TCCCCCGGGTTAACCTCTATAAGCCTTTA	3307
MGA_0554	MGA_0554F MGA_0554R	ACATGCATGCAATAATGTAATAAGACCTCC TCCCCCGGGTTACATTCATTGGTCGATGT	3538
MGA_0556	MGA_0556F MGA_0556R	ACATGCATGCACTGCTGTAGCAATTCGGTTTCATT TCCCCCGGGTTAAATTTTTGTAAAGAAAC	2206
MGA_0557	MGA_0557F MGA_0557R	ACATGCATGCGGTTATAACTGAAAGAATTCA TCCCCCGGGTCTACGCAAGAATGCGAATT	1915
MGA_0558	MGA_0558F MGA_0558R	ACATGCATGCATGTCAGATACAAATAAAA TCCCCCGGGTTAATTTTTATAAAGAGTCT	862
MGA_0566	MGA_0566F MGA_0566R	ACATGCATGCATGCTCGAAATCAAGTTAAT TCCCCCGGGTTATGGAGGAAGTACGATAA	397
MGA_0588	MGA_0588F MGA_0588R	ACATGCATGCACCTTAGTTCAAAGACAAAA TCCCCCGGGTTATGAATTTGATTAGTCT	2002
MGA_0600	MGA_0600F MGA_0600R	ACATGCATGCTTGTACAATATTACAATATGTAATTATGTTTGCT TCCCCCGGGTTATTCGTTTTCTTTAAATG	919
MGA_0674	MGA_0674F MGA_0674R	ACATGCATGCACAAAAGCAAGCACTAGATT TCCCCCGGGTTAGTTTTGTTTAAACGATCG	2182
MGA_0676	MGA_0676F MGA_0676R	ACATGCATGCACGAATAGTGCAAGCAATATCCAGCTGATAGCAAACCG TCCCCCGGGTTATTTCATGAAATGCCAGA	781
MGA_0704	MGA_0704F MGA_0704R	ACATGCATGCATGCAATTAGGAACGATTTTAAATGATCATAGTCTCGC TCCCCCGGGCTATCTACTCTTATTTCTTC	901
MGA_0808	MGA_0808F MGA_0808R	ACATGCATGCGTTAATGAAAGAAATCCAAAAGATCAAAGCGAAG TCCCCCGGGTTAGTTTTTAATTTTCATTGA	952
MGA_0809	MGA_0809F MGA_0809R	ACATGCATGCGCTGATTGATGATCAAAATC TCCCCCGGGTTATCCTTTTTTTAGAAAAAT	949
MGA_0810	MGA_0810F MGA_0810R	ACATGCATGCAAAAGTCAAGTTGTGAGTCA TCCCCCGGGTTAATAAATATCTACTGCAG	814
MGA_0830	MGA_0830F MGA_0830R	ACATGCATGCATGCAAAAAAGAAATATCTT TCCCCCGGGTTATTGGAAGTTTCTAATTA	451
MGA_0878	MGA_0878F MGA_0878R	ACATGCATGCACAAACAAGAATAAGAAAGAGATG TCCCCCGGGTTAGTTTCATTGGTAAGGTCTG	1159
MGA_0956	MGA_0956F MGA_0956R	ACATGCATGCATGTTAAATCGACAGTTTCATT TCCCCCGGGCTATCTGCGATAGCTCATAA	802
MGA_1010	MGA_1010F MGA_1010R	ACATGCATGCATAAAAACTGCCGAACCTAAC ACATGCATGCATAAAAACTGCCGAACCTAAC	1411
MGA_1085	MGA_1085F MGA_1085R	ACATGCATGCATGAAGAAAATTACAAAATA TCCCCCGGGTTAATTATTTATTATTAAAG	1657
MGA_1087	MGA_1087F MGA_1087R	ACATGCATGCTTTGGTAAGAAAACAGAAATG TCCCCCGGGTTAGGAAACATCTTTTAT	526
MGA_1122	MGA_1122F MGA_1122R	ACATGCATGCGTGCAAAATTCAAAAGAACAT TCCCCCGGGTTAAAACATTTTTTTCGTAA	706
MGA_1191	MGA_1191F MGA_1191R	ACATGCATGCGATAAGCTATCATCATCTGAACATGCAAGCAATGAGG TCCCCCGGGCTAGTTTGTACGTTGGTTAT	2578
MGA_1215	MGA_1215F MGA_1215R	ACATGCATGCAAGTGGATATGAAGATAAC TCCCCCGGGTTAATGATCAAAAAGCTTTA	529

Anexo 6. SECUENCIAS DE LOS CEBADORES, GENES  
CLONADOS Y LAS PROTEÍNAS DEDUCIDAS |

Gen (GenBank ID)	Cebador	Secuencia	Talla Producto
MGA_1357	MGA_1357F MGA_1357R	ACATGCATGCGTGAAAACATCTCGTTTGAA TCCCCCGGGCTAAAAAAATAACTTTTCA	301
<i>pvpAa</i>	PvpAaF pvpAaR	ACATGCATGCTGGGGCAAGAGTTAAATAAA TCCCCCGGGTGGCTTCAACATTTTCTCAGCTTGTT	436

**Secuencias de los productos de PCR y las proteínas deducidas a partir de los ORF de los plásmidos recombinantes generados *in silico* (solo para los productos recombinantes obtenidos experimentalmente en el trabajo).**

MGA\_0084

Producto de PCR

acatgcatgcatgagcattatgtttaaaagatcagctaataaaaagcttagaaaaacaagattaagaaaccaggttaaggtgacatta  
ggattgatgggttttttagtggtgctacgactgtgccgttagctaataatcccgtatttgggtgtataataaaaaagaataacgacgag  
acgcgcgaagaattattaagagaagtggttttagcagcttctaagagaagttattaagataaccctcacgcatcatcccaaat  
atcgttcaagacttcattcacctaaccagactgatggtgataagaataatggaattgttaaaacggtttatgacaattaccggttaaat  
gttgaacggttaaaaagtgctaattgatccaagtgattgggttaagttaataatgaatatattataaaaaatggatcgcatctattttaga  
tctactaaactcggtcatgtcgtttatccgatctatcaccaatcatggttaagtgaactaaccctaaagaatcttattataggtacccaaag  
attgggatcactagtgggttttttaaacagtaccatgctgagatcgatggggcttattgttgataaaaaataatttcgcttttagtcagtaac  
taaagctgatctaattgcgatacataaactgtttgtaagtgcctactaaagtaggaatcacgattcgaacagcagctaactctcaactg  
attattgatcctaatacgaaggattagtactaactaatgaagatcttttagaatcatcagctggagcacttagtcctatacaagtatcta  
tgtgaaaactaactcaccagatacagaactattatgattatcaggtgaagaataacctgatcacaagacccaagagattcctagttt  
ccaattacatcaatcttatatagtatcattaatgataacaactagaaaagaatgatattaagcattactttatcaatgggtgatacctgat  
taattaactctttaatgcttcttattatgaacaacagaatgtaataaaaaaataaccggggga

Polipéptido recombinante (42,4 kDa)

MRGSHHHHHHGSACMSIMFKRSANKKLKRNKIKKPKVVTGLMLGFLVVATTVPLANIPFLGLYN  
NKKNNDETRERIIKEKFLAASKEKLLKDNPHAIIPNIVQDFHSPNQTDGDKNNGIVKTVYDNYPF  
KLVERLKSANIDPSDWWVKFNEYIYKNMDRDLFLDLLNSFMSFYPIYHQSWLSELTQRILLYRYQKI  
GITSGFFKQYHAEIDGAYYVDKNNFAFSSVTKADLIAHKLFNGVPTKVGITIRTAANSQLIIDPNTK  
GLVL TNASIFLESSAGALSPYTSIYVKTNSPDTKNYYDYLGEEIPDHKTQEIPSPFITSILYSIINDK  
QLEKNDIKHYFYQWVIVPLINFFNASYYEQQNVIKK

MGA\_0267

Producto de PCR

acatgcatgactagagcgtcacaaacccgttcaaaaaggagaatttgatgacaaagttgttattcagactgcacaaaatcaattttatcc  
attaatgtcagcatttagtcaattggtgatctctacaataaacaatttgctaatacgccagggtttttacctgttgattacaacaaagtgaag  
aaactaatgctactagcgaattacaattaactaataatgttgtaggttagcattcgttcaactcaccttagttccaaatattatattagctgat  
ctaaatgcagcttatcagattaatggattaatcgattattagattatctaacaatccaattatcaacgaaaactatttgacagtgcattta  
taacaactcaataaaaataagtggtatcccaaaagtagtgataaagttatgcaattccgtttaatttaacaaccactgacagtttagtatt  
caacaaacctttaatgaattgttatttagtttagtgatcaaggcgggtggaactgtagataagaattctgcaacctacaagaactcata  
tggaagactttatggagaagattccaaacaaaaaatgaaaagattgcagataaaaaacaacgaaatttacaaggacttactgttga  
cgatcaaaccttttcaaacctagaatcaatgtttgaattttctaagaatttactgaagggttagaactaaagcaagctgaacagttacag  
gccaacaaagagatttaaaagtattttagtatttggtccaaatatttatcaaaaatatctgtgatcaaaattaggttaattctagagattc  
tgattgtgaaatctaaaattacaagacaaccaatttgatttagatttttctaatacgaagaagacagctgatcaagcacaattggagaa  
acttacgatttcttaagaataattacacaacattaaatttaaatgataagcagggttttaaatcaatttttggacaggcggaacatctg  
attgagcgtcatgagatattagaatttcgatactgcttttggtatcgctcctcacgttggtgaaatcagtcagttgatctctctttacgatta  
gaacgtttagaagcactaaagggtgatgttaccacaacagacattaataacgcaaaaaataacttggcttcagcagatgatgttctgtgaa  
aaacacagctgactaaattagataaaaaataatttaataagaacactttctgaattggtggttcttctaattgggatttcaactctagaca  
aagagatgaacaaacaaaaagttcttagagtgatttataaacgatactagaatagctgcacctgatttccgtttaacaatcaatca  
gtaagtcagatttttaaatagtggtcagcgtatgatttaacttcaacaagctctatctacaacagctcaacaagcacttacagatactat  
agctaaagaacaagccgaagtggaaggttaaaaaataacccaactccagaaacttataataatagcgactgatcaagaatttattt  
aaataaagggtttagcaagtttaacggattgattgacattcaaatcagaattaaataaagatctgtaacaattcaattgctttcattc  
atacagatgaaaaagcaaatcgaattctatttctaattggagcagtttcaggaattactagatcagccgatccaactaatttaactaa  
agaaaatctacttagacagattaatgatgcttaggaacagctgatcaataaccggggga

Polipéptido recombinante (63,8 kDa)

MRGSHHHHHHGSACSEAMSRKYNAADADFRKAAGNIGTINQVKNVVENSLQLPKTAPT  
MAPDAGAAAPSAGSEAAEAEASTGTGNTGDTGTAAAGSGSGSSASSSAGSGSQGGGETTS  
TGSGGSSSNTSGQSGSAQANTAGGSGASSNSGSQSSSTQGSAAAGTAASGSTSETAMMTM  
LSGAWKHFDIAIVKAKMNSEKTTAEMGLMKFYTDFKAKFDEVLFEGSKNLRDGTASLEDLRTG  
KTDLETKAKALFTVYNALGTALGQSKTFDTTKKSFEELAEANQGLSTLYKLMNVVDVYNFAKPIEG  
NSYNDLARFFELNGFLGGQFAMNTTTSNVFQLFGNITALLISRELRIIDGAINSVLLTTPSTNREG  
LETLLLTNVNKNSTIKQRYDETVKALNDVKAQTDQWANSPIVTKTNNQSKLEMAAAKANETSK  
PVLTRLVNALYGNSANMPGFAVKGLFRSVNDLKTQLDAVKTLVSANYTSLSDFINVYKMTNTTW  
MEGVGIFVNSQALNGLNDTEIDKLAATNVSDAADTALFTEVKTELKRLVAATTSPLKNGAGQDK  
RSAFYLNQMVRSITSDNQARLASMGEGAQEFVNTAL

MGA\_0281

Producto de PCR

acatgcatgcatgaatttagggtttagcaaatttaaaattggccgaaataaataatgaggattctatcgattatcggttgatctaaatcagat  
atcgatgaacacctagtccaatatgaaaaagaataaaacttttagagggtattatgccttattattggtttgcttcttaacagtttgtaa  
cattagtaattacgcaaaaactgataataacaatatcttctgcaatttgcaaaagtactttacttttaacattttgtttcataacagactg  
attagctcatgcatctactacaagcatgcaacaggccaaccacagataaaattgttaaagcattcttaaaatacttttcacatttaatagt  
attgtcattattatctgtgttcttcatctggtcaagcaataaattcttggtaaccaactaagtgctacgaacgaataatcttgcctcatttcaat  
cactagcttagttaaacaattagattgtcatggttttaagttatttaaaccatttggtgcaattactagtggtttgttactcaaaagaacta  
ttagcttcagtttttgattatcattgttttaattatctgtttgcaataattatctgatcacagaacagtggaattactcgttaagcgagctgcct  
tcttaactcaagtaagtgataactacttaaaattttctcagtttgagctcttattcacctaacaagcttatgcaaggatctggacagtatgaa  
gaagcaatcaacaagatcacagatgctcagattcaagcagcagcaagtgctttaaattcagttgatgcatctaattttgctagcttaagtt  
cggtgatatgttcaaaactttactgaagcctttatttcacaacaataacattaacaacagtaggttatggtgactttgttccccatgcaccaat  
ctcacgggtaattgttagctttatctctttattagctatttcaattattgcaatcccttcaggcattattgctggttcttactgaatgcaaa  
aatcatctgaaagaaataaagtaaaaaactaagaaaaaccctgataataaaaaagataatttcttatgattcctaaaaagaaacaagc  
tgaaactaacattaagaaaaaacattggatcagtgtagctaatccaattaacaaaattaaccggggga

Polipéptido recombinante (46,3 kDa)

MRGSHHHHHHGSACMNLGFSKFKIGRNKLMRILSIIVWSKSDIDEHLVQYEKRIKLFIRGIYALIIVF  
ASLISFVTLVITPKTDNNNIFAQFAKVTLLLTFFVFITDWLHAHAFTYKHATGQPQIKFVKAFLKYFFT  
FNSVIIICVLSSGQAIKFFVTNLSATNELIFASFQSLALVKTIIRLFMVLSLFKPFGAITSVFVTQKKLL  
ASVFLIIIVLILFAIIIWSQETVELLRKRAAFLTQVSDNYLNFSSQFAVFIHLNKLMOGSGQYEEAINKI  
TDAQIQAAASALNSVDASNFASLSSGYVQNFTAEFYFTTITLTTVGYGDFVPHAPISRIVVSFISLL  
AISIIAPSGIIAGSFLTEMQKSSERNKVKTKKNPDNKKDNFFMIPKKKQAEINIKKKPLDQCELIPI  
KQN

MGA\_0319

Producto de PCR

acatgcatgcaagagaatctatgggaataactacatggatctaataccaatcccagttatttgatcctaattgggggaacaaaagaaagtt  
atcttcaaaaccaattaacacaaaagattagcaacgactttttgactcttatttaaggattcttattaaactatgcgccaacaatgtcaac  
tgacgctgatggtttgtcgtttattcaaacctaactaagatccctcaagaacgattttaagtaaccagctaactgagccaacattagatt  
actgatggtgggttagtgcaactaataacgaacaaactaataacgaattcttagctcagatccaagctttgttcaatcagtgagtaac  
taagaaaaaccctaatttagttcaagaattgtttaccagcgatactcctaagatggactaaattcgatcttaattgatcgtactgttagct  
caacagcatttagtagcaagttataattccaagcatttaaaaatccaaagaccaaccctttgatgaaaacaaggggatgagagcttat  
aaccaattcattactagagggttgacagttatgtaaatagtacaactaaagggattaataattccaacaactatcatcagattctggtg  
taagttgttaattgactgcttctgatatttcaatagttttgacgtatcatttagtcagcttatgttcaacaatatttaaaacaaactaataact  
aaccgcgataacgttggtgtagttcaatcagatattgatgagatcaatctgatgaataattttattagagctaagggaatggaacacag  
ccaacacactactcaaaaagattactaataattcattataaaaacaggtgaaacgaatcgaactactgatccatattacaatgcttatgc  
agatttagcagccggaaataaagatatccatgaatctttgaatgaaatgggatgaagacagttgattctcgaaatcagataattcatc  
aacaacgtaattagtaaaaaatgaacgtagtaagttcttatttctcgtggtcagatgggatccatgtgatggcgatcagtggtggg  
gattactactaaacaaagtggtatgactagaaatataacaaaacaaaagagtttctagcttaccgttcattataagatcttcagttaa  
cttatctgaaaatgaaagatatgatttgatctaagccccggggga

Polipéptido recombinante (48 kDa)

MRGSHHHHHHGSACKRIYGNMYMDLIQSQFLDPNGGTKESYLQNQLNQKISNDFDLSLFKDSY  
LNYAPTMSTDADGFVVYSNLTKIPSRITLSNPANWANIRFTDGGFSATNNEQTNNEFLAQIQAFA  
FNQWVTKENPNLVSRIVFTSDTPKDGLNSIFNDRTVSSTALVASYNFQAFKNPKDQPFDENKGM  
RAYNQFITRGLDSYVNSTTKGINIPNNLSSDSGGKLLMTASDMFNSFDVSFSAAYVQYQYKQTN  
NTNRDNAVGVVQSDIDEINLMNNFIRAKGNNGTANTYSQKITNNSLLKTGETNRTTDPYYNAYADL

AAGNKDIHEIFEWNGMKTVDSSKSDNSSTNVMSKNMNASKFLFSRGRDGIHVMIDGGDYLLN  
QSGMTRNITKQKEFLAYRSLRSSVNLSENERDYDFDLSPGSTCSQA

MGA\_0398

Producto de PCR

acatgcatgctgtggtgctgcttctcactaacatgaatctctgttcagttagtagtttcagataatagtccaacttagctgaccaatcattc  
tcagaaactcatataacgggattagagatttctacaaatctgtatacaagaactcttcttagctaacgaccttagttatcagttaac  
aatggtgtgtgaaaacgacctggaacaattgatgattcaagaattaacacttatcgacagattaaaaatgatggatcattgtgtgtag  
ctactggtttcaaccaagaatcagcactaaaccgaattcttgataacacagctttatataacgaattcaaagatttggttcatcttgtgga  
tgggtttatcactaagaaaaatggaactaatatttcagcaataactttccaaactgaaagtgtgcatttttaacggggatcgctgctggag  
tttggtaaataaaaaatagcaatttcttaaaacaacaaaagttggtgatgtagacactatggtattggtgcttatgtaggtttacattccat  
caacaattagttcttaaacggtttagaatgggagcgatcttcttaatacatccaaccaagagtagaaggatttaagaaacttctt  
gagtaagcttaatgcagcagataacaatggtaatcgtgatactttaccagctgataaatctgattcatttaacgcgcagagaacaaaga  
gcaacaacaattactaaccaatgttacaaaatggtgcatctttaatctatccgattgctggtccacaaacagctcttagtaaaaacgtgat  
tctttctaatagaaacacacaccacgcacaattaattggtgttgatacagctcaagaaaatttagcaacaacacaaccattacaaggag  
ctcctggtggaagacaatcgctttctcaacagttaaagcttttagatgtgtgtgtgattcaactttaaaagcaattcaaaaaggtaacca  
gtaaacggcttctatggatggaataactagctacttttagctaacgggtggtgtttcattatcaagcgcaggattagcttactacca  
aacttaactgatctgtttactaaaactagcatacctgtaatgcagatcaaaactactccaagacaagaaatgtctgatcaatcatctggaa  
ctccagcgatgactggaactatgaacggtaagtaattggtggaactgaatctatgatggctatgagtatgatgatggcagaagctgg  
ggcttctacagcactgaatatgccagcaacatcttcaacaccagctgtacaacaactaacaacaaatttctgtaacgacttagttac  
aacaacagaagctaataaaaactatggaattatgcagtacgttaacattttagctggaacttctactttactaccagctgctagtagagaa  
aactgaaaaaattaaaggatgaattaagaatgttcaaaacttctgtggaagctgatgctaaattattaccaattttaacaacagacgttg  
atcctaatacagcttttaagtataagctgcaatcatttctggttcattccaacaagctgaaggtagctttaagccaagcaggcaacaa  
attgtcttcaagaaattaaactaaccggggga

Polipéptido recombinante (60,4 kDa)

MRGSHHHHHHGSACCGAASSLTYESSVQLVVS DNSSTLADQSFSETSYNGIRDYFKSVYKETL  
PLANDPSLSVNNGVWKRPGTIDDSRINTYRQIKNDGSFVAVATGFNQESALNRILDNTALYNEFK  
DFGFIFVDGVITKKNGTNISAITFQTESAFLTGIAAGVLVNKNSNFFKTTKVGDDVDTYGIGAYVGL  
PIPSTISFLNGFRMGAIFFNQYIQPRVEGFKKLSWVSSNAADNNGNRDLPADKSDSFNATEQR  
ATTITNQLLQNGASLIYPIAGPQTALSQNVILSNRNTHHAQLIGVDTAQENLATTQPLQGAPGGKT  
IAFSTVKALDVAVDSTLKAIQKGQPVNGFYGYGWNNLATLANGGVSLSSAGLAYLPNLDTLFTKT  
SIPANADQTTPRQEMSDQSSGTPAMTGTMNGQANGGTESMMAMSMMAEAGASTALNMPA  
TSSTPAATTTNKQISANDLVTTTEANKTMVIMQYVNILAGTSTLLPAASRENWKIKGDELRMFKT  
SVEADAKLLPILTTDVPNTALSHKSAISGSFQQAEGTLLSQAGKQFVFKLN

MGA\_0416

Producto de PCR

acatgcatgctgctgctgctgttttcagacggaataatgacttattcgaagtttttagatatcttaaaagatccacagaatccaaaaacttc  
agctaaagaagttgttatctaggttgaacaaaacgaaatcagacaaccaaccaagctaataaccctaatacaaaagagctt  
gttcaataatttctcatatgctgttgatcttatgatagaataacaccccaacctttgtaataacaatgatagctcaatcgatctaacttta  
agttgtttataatcaagcttctatcgattttaaactcgattgtttaccagctatccattattctagtaacttatataactatcagaagttaata  
aattagataactgagcaactaatttattaccaactgaagcaaatgaagctggtgtgcttctgcttatttgaatcgatggcaacggggtat  
taactggttaatgaacacagataattacattttaccctcgatgattgaattgaatttaaggaatttaaccctgaattatgcaaaaagc  
gcgatttatgatccaagagcttacaatgttgattaccaactaaatgcaaatgagattatgactgataacttattaaatccctcaaaactcta  
acgaagttattgtcttctcaaatattaatatctattgttattatcggttggaattaaatgggtgatgttgcccaaccagggtgcttatattac  
atcaaccagtgatgtaagattccaagcgttagctaataaatttcaatcaatctacaacaccaaatttaactgctccagtatttaatttaaac  
ttaatgatcttggtgtggtgtgagatacaatgttgaacaaagagctagcaacacacaaaccctagccaatcaactaatccagctgatt  
atgtaattagaccaactgcgattgatgtatttccattaggaatcttataatccgaataatttagttatcaatcggtgatctaaaagat  
caaagagttggttttaacaagaatgaactaaccaattagttgatctatctaagatcttgacagctgatcagatgctaacacagaaaaac  
gttaatacagataccctaactaatcagattagtaatatccgtcaaaatactggtttattctttaatttataagtagtactcatttagaataatt  
aatcaaaacgacgtgttgataacaatgcgttgggtgctaatttaaccactattataattggatttttaataagttgactcaagtgaaaacta  
ttcaaagatcgatttagtggtacctaataatgatcctgagattaaggcggaaccactccaccagctagtgatgaataccggaatgaa  
cggtggtatggcaataatgatgctggtgatatgaatggtcaaatgaacgggtgagatggctagtagcactaataacgaaatgccaact  
caccagctactaactaaccggggga

Polipéptido recombinante (57,3 kDa)

MRGSHHHHHHGSACSSAVFRRNNMTYSKFLDILKDPQNPKTSKEVVYLGWNQNEIRQPTQQ  
ANNPNQKSLFNNFSYAVGSYDSNNTPTFVNNNDSSIDLTFKLFYNQAFISILNSISLPAIHYSNLY  
NYLEVNKLDNWATNLLPTEANQAGVLRFSFYESMANGLLTGNEDNYIFTIDLKFEFKELTPELS

AKSAIYDPRAYNVQYQLNANEIMTDNLLNTSNYLTKLFLVLSNINISFAYYIVGLNGDVAPTRDAYIT  
STSDVRFQALANKFQSIYNTKLTPVFNKLKNDLGVVVRYNVEQRASNTTNPSQSTNPADYVIR  
PTAIDSIIPLGILSNPNNLVYQSGDLKDQRVGFKQEWTNQLVDLSKILTADQYANTENVNTDTLTN  
QISNIRQNTGLFFNLLSSTHFRIINQNDVVDNNAFGANLTTYNNWYFNKLTQVENYSKIDLVVPNN  
DPEIKAVTTPASDMNTEMNVGMANNMTGDMNGQMNGEMASSTNNEMPTSPATN

MGA\_0485

Producto de PCR

acatgcatgcagtgagggaattactcatcataaaaattcaaataatgaatatcaattggtaatacaagctgggaaggattagatcctca  
acaccaactcaaccaaagatcccaacgcaccagtagctaataataaacggctgctgctgatcaagctcaagaaaaaccaagac  
aaccagcaaacttagctactttaaaaactgatattgatgacaagatgtcagatacaattggggagtttattcaagcgatcttttaggtaaa  
gataatctaatacgatcaaagaattgcagcgattcaaaatcaaagtgaatgaattgaagagaagtttaataaaacccttattattctca  
gatcaaagcattcttctgaagaatcaaataagattgagataaaaactaaccttcaaaattggtttagatatcggtatccttatgtgcttcagcta  
atgctgaatttaataaaggtaagatcgatttaataacaaaacttatgaaaataagattggggtaatacgggatactaccaactataaaa  
aagaagttactggtgaaggaaactcaattacaccaaagtcagatccacaaaaagctaaagtacaaaaatactacttcagatgaagaa  
ggtaagaacgttttaaaaacttactttaatgctttaaaacaaagtgcacatcaatcttctaaacgactctgatctaccagggttggaag  
gattatgatattaatggtcgtgttactgactaataatgatgctgggtgttctcaaaccacctaagaactttaatagctgggatgactatatta  
tcagaagatcagaccaagatttatcgactttgacttagaacaacaaagatccggctgaacaagatcaacagaaccaacaaaat  
aaccaggctgttgaaaactgtttaccaccaactgttccagttgaaggaatacccgggga

Polipéptido recombinante (36,7 kDa)

MRGSHHHHHHGSACSGVITHHKNSNNEYQLVNQAGEGLDPSTPTQPNPNAPVANNNNGRA  
DQAQEKPRQPANLATLKTIDDDKMSDTIGEFQIAIFLGKDNLDQRIAAIQNSDLSFEEKFNKTY  
YSQIKAFFAKNQNEIKTNPSKFGLDIVYPYVLSANAEFNKGKIVFNNKTYENKIWGNTDITNYKKE  
VTGEGNSITPNADPQKAKVQNTTSDEEGKNVLKTYFNALKQSATSIFLNDSDLPVGVGKDYDING  
RVTDTNDAGVFSNPPKNFNSWDDYIISKIRPRFIDFDLEQNKDPAEQDQQNQNNQAVENLFTT  
NCSSWR

MGA\_0648

Producto de PCR

acatgcatgccagaaaatgaagctccagcaacaaaaagtcagaggagatattcgactttaactaaaggcgatattacgatcgc  
atttaataaaaatcaggataaaaccaagattctcccagtgcaattatgagaagaactatttaattatttctactttcaaattagtcagagt  
gttgacaaaaaagaattagatcacattaatttgatgcaagaactttgttcttaataaccaatcggtactttatcatttgatctagatgatc  
agttgttgacagacctaaagatattaaaaattcgattttagaatcgatggtctgatgacaggtagtgagctagtagtggtatgggtca  
agttctgctagtagcggtacagggtcaggtatagtggttcaactacaactcaacctgagcagatccaaatgactgtccaccagatag  
tagtaaggctaataaaactcaagtagtgaacaaacagtaatacaactagaccaactgaaagccaaagacaacctcgtagtag  
taaaccacctagtttctaatgggttccaggattccctggatcaaatattactaagaaattgttcaaaaacaaaatagttatccagacta  
ttagcaagtagatacaaaactgtagacaaacaaactaatatataagaaatagagaccgtactttctataaggccagggtactttac  
cagtgataacggaacagctcatcaatagttgaccaaggtagcaggatgagtttagattacgctaagaactcgtaacaaacaatcaa  
ctaaaattattcattgcaactaacctacagttattggaattatggaataactaactctaaagagttagacgatttattatcttaccagacc  
caacagggtgtaaaccaagcggatttgcttggtaaatctaataaaactccatcggttactctcaacctaacaatcaaagtggtgatcaa  
ctaaacaaaatgcaggcaccatgcttattattcatataaagaaggagggggcaaaattggtatagaagaggttaaaactactgaag  
cttcagcactcctaaaattgtgttctgctgacttattagatgatcagcgtttaaaccaatataaatctaaataaatgaaaaatgaa  
aagaatgaaaagctaaaaccttaagtctattaacggatctaataacgtgacggtactgtattgtgatgaagaaaaagctagattaa  
agaaattaaccgaatacaatggtaagatttcattctatactgactttgggattttagaattagacgttgatctatcaaaagcagatgaaaa  
cttaaggaatgaataataattcaaccaaagctgtagattctacgtaaatagaataaaaactacacaacacttacctaactataatcca  
gaaacaggtaattatttccctacttttagattacatctctaaaaatctaggattagcatctgttaaacccttctatgcatatggtattgataatgca  
aaagataatctatatagccggtatctaaagacctacaagtcacaaactttctgaatgcagaataatccagcagaaagatatcaagatat  
aacacagggtagtgtaataatcaactacgtattctcaacaagcttgactacctaataaagatctatttaacgtaccaagctatactcgtt  
ctacagactcggcaacatcaaaaacttgatctgatataaataattcaaaattatacaaaaactttgaacagaccgttcattgacagata  
cggatttaattacttactaaattctcatcactatactatggagcttctggttctgtagcttacaatgactttggtcaaaattgttgatctatgatg  
gggtgctagtagcacaaaatttggtgacgtttcaaatcttgaacatttgacaccttagtacaagctgtaacattaaagttcaatctgatga  
aaacattactaattatgcttacaacttaattgacaataaaggtttccgcaccaactagatcataccgaagtaacttaagtctgttctacc  
ctaacggatttacctgacctcaacatctggtacacactactcaagtaactctaaagtaacagcaatttccctaacgggttctaaccggg  
gga

Polipéptido recombinante (81,7 kDa)

MRGSHHHHHHGSACQKMQAPATKSAAEEIFRTLTKGDITIAFNKNQDKTKILPSAINEKNYFNFY  
FQISQSVQKELDHINFDAKNFVPPNNQSGTLSFDLDVSVVDRPKDIKNDFRIDGLMTGSGASS  
GSGSSSASSGSGSGNSGSTTTQPEHDPNDPPDTSKANKTSSEKTVNTRPTESQRQPRDS

SKPPSFPNGFPGFPGSNITKKFVQKQNSYPDYVASRYKTVDKQLIYKEI**W**DRFTFSIRPGTFTSDN  
GTAHPVDQGTG**W**VLDYAKNSSNNQLKFIATNLHVIGNYGNNTNSKELDDLLSYSDPTGVKPS  
GFALGKSNKTPSFTQPNNQSVQDLKQNAGTMLYYSYKEGGGKIGIEEVKTTEAFSTPKIVFAA  
VDYLDHALNQYKSKINEK**W**KE**W**KAKTLSSINGSNSSDGTVFADDEEKARLKKLTEYNGKISFYT  
DFGILELDVDLSKADEKLKE**W**INNSTKAVDSYVNRITKTQHLPNYPETGNFYPTLDYISKNLGLA  
SCKPSYAYGIDNAKDIYIAGYPKDLQSQT**W**MQNNPAERYQDITQGSVNNQLRIPQQA**W**LPNK  
DLFNVPSYTRSTDATS**W**SDINNIQIYTKL**W**NRPFIDRYGFNYFTKFSSLYYGASGSVAYNDF  
GQIVGIYDGVASSTNFGDVSNLGTFAPLVQAANIKVQSDENITNYAYNLIDNKGFPHQTRSYRSN  
LSLFYPNGFT**W**PSTSGTPTQSNSKVTAIFPNGF

MGA\_0760

Producto de PCR

acatgcatgcatgaacaatatctttattacgccacaagcagtgggcgataacacgaaccttgcgatatattattaacttattaattatgttgc  
ttttattaacagttccaacactgctaactctttgttaattgtttataaacccaaggttaaaacaaaaccaaagttaactctttatgcatttcatcc  
gcattattattaagatttgaacattaggtctattgtttgaagcaacagatggggcaacttggttgaacaacaccttaagtagttacacc  
gaggtaatacaaaccttaataagattgccatcattgttgggtgggcaatcttaggtctgacactagtaatctcattcagattttgtacatcca  
ctttaccaacaagatcattgtaaccacacccatgatcattctttacacatcatgaacgagggtgaagatgtacaacacaagatgaaag  
aacacaaacctaatgttaaagcagcttgattagtaatacattgtttattatcacatcgagcgatcgatggtttatcctagggggaacagttt  
ctttattaacccctgaatccaactaaaattaaccttgggtttattattaccttaatactccatattcttggtagggtgtaattattcattaccgtcagat  
ccagtatggtgaaaaaagatttaaagctgcacttcacaactctacacgacattattaattatcccaataatgatagtgaggagcttatttaa  
accgtttcttaagaaacgtagggttgatcattcattagttatgtctagtggtggggaatcatcaccttcttagcagttattgagttagtagcct  
gaatttattcataaccgttcgatgaaaaccgcagactgatataagattgaatctcgtttgctttggtttggtttgcaatcttaactctgtcagt  
acacaacctgtctcactgaacatgatcacgggggatgatgatgatgatgaccacggggccaatagcccatttgattagtagcaaa  
caaagaacttactgctttacaaatttaataacccggggga

Polipéptido recombinante (41,3 kDa)

MRGSHHHHHHGSACMNNIFITPQAVGDNTNLAIFINLLIYVAFLLTVPTLLIFCLMFIKPKLKQNQM  
FNLYAFSSALLLSIGTLGLLFEATDGAT**W**FVNNTLSSYTEVNQTLIKIAIIVGGAILGLTLVISFRFLYI  
HFTKQDHCNHTHDHSLHIMNEGEDVQHKMKEHKPNVKA**W**LVIIILLSHRAIDGFIILGGTVSLLT  
LNPTKINLGFITFNLHILVEVVIHYRQIQYGEKRFKAALHNFYTTLLIPIIMIVGAYLNRFLRNVG**W**II  
PLVNASGGVITFLAVIELVPEFIHNRSMKTAD**W**YKMLISFAFGLVFAILILSVHNLSTEHHDHGM  
MDHDHDHGPIAHLISQTKNLLLYKFK

MGA\_0797

Producto de PCR

acatgcatgcatgaaaaatattttaagattatcaatgtaaccagaaggcaattagcggtatttttagttatttagcgacgattttctagcaa  
gtgctggaatcatcagtttaagacaaacgcaattgtatcctggtggaataacttcgtaattatagcgaaaattatacacgatttctct  
aatttaccgttagtttcttgtattttgtatcaccggttagcaattttataactggtaaagaaagtataagatgataataaactaatcca  
acaaactataagttcaagacactattagctttattgaattgtgcagtttaattgcagattgttttctgataatttaactcgcttattctatccg  
attaataaaacaaattgcacaaattatcaagaacaaatagctacagctgattttactaatcgcagcattcaaacgtggtttgtatcattatta  
gtattaataacaacttctgtgatctaaaattatctgtttgaaatagcagctacacttccatatttcttaacaaagatgtatttcatatcaa  
atatcatttactttgctttatacatcgctttatttattcctttcttaaaacttagttattttaaaccaaaagcaagaacattagtttaattccatttagt  
ggattggtattaaatttaagagaccaacaaagatttaatactgaaggattaatcaccgataataatgaaaacttaaaagttagatcaaaaa  
atcttattatcaataagacagatttccaaattaaattaactcaacagttgtttactgtcatttcttcttttagttggcgtgtgacattcacagc  
attagcttcatattttatttactgtatggatatggatacacactaatcccttctgatacttatgttgccgcttattcaattgacttttaattattca  
atgtagaaaaggaaatattattatcgttattttgatgagcagcagcaattataatgttgatcatctcatcgatttcagtaggggaatcattcta  
gatgataatccattcgttttcctaaagcaatcagaagaataataactttatcattaacttttagttgttattgaaattgcacaatgagcatt  
cttattattagtgtaattttctttatctaataagctgaaggattatatcgtagttgcagatgtccagctgtatatcctccgctgaaccc  
gttaagccagatttaactagtttaatttattctagctcagatttattacaactacaaatctatttctcgtagcattatatctaagttgattctgatct  
atgcaaaagctggtatgttttctaagagaacaagaatctttatgtatatcccttttgaacattattctttagtacctaaacacggtatagaaaaa  
agaaaacaagttataataatttaactgccgttttaatttaacccggggga

Polipéptido recombinante (57,9 kDa)

MRGSHHHHHHGSACMKKYFKIINVTRRQLALFLVYLATIFLASAGIISLMTNAIVSPGGNNFVIYSE  
NLYTISLILPLVFFVIFAITGLAILLTGKESYKYDNKLIQQTYKFKTLLAFIEFVQLIA**W**FFLIILIIAYFY  
PINKTIAQIYQEQIATADFTNRSIQTGLLSFISINNFFADLKLSVLKYAATLPYFFNKDVFSLSNIIYF  
ALYIALFIPFLKLSLFKPKQRTLVLIPFSGLVLSYLRPTRFNHEGLITDNNENLKWYQKILLSNKTVF  
QIKLTSTVVYLSFFFLVGVLTFTALASYFILTGDGYGYTLIPSDTYVAAYSIDFLNYSMLEKEILLSLF  
**W**WAAAIIMLIIFIVFSMGIIIDNPFVFPKAIRKNKYTLSTLVVIEIAQ**W**AFLLLV**W**IFFLSNAS**W**RII



SSVADVPAVYPPPEPVKPDLTSLIYSSSDLFTTTNLFFVALYLS**W**FLIYAKAGMFSKRTRIFMYIPF  
ATLFFVPKHVIEKRKQVNKFNCRLI

MGA\_0843

Producto de PCR

acatgcatgcatgaatagtagtttccttttatcgctcttattagtaattattccaatcgggctaataagttacgttatctataagagaaaaaat  
ctaaagagcctggcagttttcaggaaaaactaaagaagaacgctgtaatgaggtgtgaaaaaccattaagcgttatctccaagataa  
cgaaatgtatggcggtgaaattatgtattctttgtgctaaacgaccatctccaaacgatgatcgtaagttgcacaagcaattaaagaag  
aaacaaagcaatacctattagagcacaactatctaaaaaagataaaaaacaatatttagataatcgtaaaaaagagatggctcgtg  
aacgatattgcatttatttccaaccaaagatgctaggactcaagctacatttgatcctgcaattattgaagcagaagtattaacattacca  
gctaaaagtaagcgtgatgttccagaaagaaaaatccaaattaatgggtgcaagatttcaaaaagaatttctgaattgaaccattaa  
agaataaagaagatgcgcatgtaaaaaagcagaagatgaacgcttaagaaaattagagattaaagagcgcaagaaagctgcta  
aattagctaaaaaagaagcaaaagctaaaaagaaaatttaaccgggggga

Polipéptido recombinante (27,7 kDa)

MRGSHHHHHHGSACMNSSFLFIVLLVIPIGLISYVIYKRKKSKEPGQFSGKTKEERRNEV**W**KTIK  
RYLQDNEMYGREIMYSFVAKRPSNDDRKLHKQFKEETKQYLLEHKLKSKDKKQYLDNRKKEM  
ARERYCIYFQTKDARTQATFDPAIEAEVLTLPKSKRDVPERKIQINGLQDFQKEFS**W**IEPLKNK  
EDARLKAEDERLRKLEIKERKKAALKAKKEAKAKKI

MGA\_0993

Producto de PCR

acatgcatgcactagagcgtcacaaacccgttcaaaaaggagaattgatgacaaagttgtattcagactgcacaaatcaattttatcc  
attaatgcatgcatgtagcaattggtgatctctacaataaacaattgctaatacgccaggtttttacctgttgtaattacaacaaagtga  
aaactaatgctactagcgaattacaattaactaataatgttgtaggttagcattcgttcaactcaccttagttccaaatattatattagctgat  
ctaaatgcagcttatcagattaatggatttaatcgattattagattatctaacaatccaattatcaacgaaaactatttgacagtgcattta  
taacaacttcaataaaataagtggatctacccaaagttagtataaagttatgcaattccgtttaatttaacaaccactgacagtttagtatt  
caacaaacctttaatgaatttgtatttagtttagtgatcaaggcgggtggaactgtagataagaattctgcaacctacaagaacttcata  
tggaagactttatggagaagattccaaacaaaaaatgaaaagattgacagataaaaaacaacgaaatttacaaggacttactgttga  
cgatcaaacttttcaaacctagaatcaatgtttgaattttctaagaaatttactgaaggttagaactaaagcaagctgaacagttacag  
gccaacaaagagatttaaaagtatttatgttagattatggtccaaatatttatcaaaaatatctgtgatcaaaattaggttaattctagagattc  
ttgattgtgaaatctaaaattacaagacaaccaatttgatttagatttttctaataagaaagacagctgatcaagcacaattggagaa  
acttacgatttcttaagaataattacacaacattaaatttaagataagcaggttttaaatcaattattttggaacaggcggaacatctg  
attgagcgtcatgagatattagaatttcgatactgctttggtatcgctcctcacgttggtgataaatcagtcagttgatctctctttacgatta  
gaacgtttagaagcactaaagggtgatgttaccacaagacattaataacgcaaaaaataacttggcttcagcagatgatgttctgtgaa  
aaacacagctgactaaattagataaaaaataatttaataagaacactttctgaattggtggttcttctaattgggatttcaacttctagaca  
aagagatgaacaaacaaaaagttcttagagtgatttataaacgatactagaatagctgcacctgatttccgtttaacaatcaatca  
gtaagtcagattttaaatagtgtgtagcgtgatgatttaacttcaacaagctctatctacaacagctcaacaagcacttacagatactat  
agctaaagaacaagccgaagtggataaggtaaaaaataacccaactccagaaactataataatagcgactgatcaagaatttattt  
aaataaagggtcttagcaagtttaacggattgattgacattcaaatcagaattaaataaagatcctgtaacaattcaattgctttcattc  
atacagatgaaaaagcaaatcgaattctatctatttctaattggagcagtttcaggaattactagatcagccgatccaactaatttaactaa  
agaaaatctacttagacagattaatgatgcttaggaacagctgatcaataaccgggggga

Polipéptido recombinante (68,1 kDa)

MRGSHHHHHHGSACTRASQPVQKGEFDDKVVIQTAQNQFYPLMSAFSQLVDLYNKQFANTPG  
FLPVELQQSEKTNATSELQLTNNVVGSIKNSPLVPNIILADLNAAYQINGFNRLDLNPNPIINEN  
YFDSDIYNNFNKISGSTQSSDKVYAIPFNLTTTDSLNFNKLPLMNLFLSLVDQGGGTVDKNSATYK  
ELHMEDFMEKIPNKK**W**DLQIKNNEIYKGLTVDDQTFNLESMFEFSKKFTEGLELKQARTVTG  
QQRDLKVFMLDYGPNIYQKYL**W**SKLGNRSRDS**W**LWNLKLQDNQFDLDFSNLRKTADQSTIGETY  
DFFKNNTYTLNLNDKQVLKSIYFGTGGTSD**W**AS**W**DIRNFDTAFGIAPHVGNQSVVSPFTIRTF  
RSTKGDVTQQDINNKNFASADDVL**W**KTQLTKLDKNNLNKNTF**W**IGGSSLIGISTSRQRDEQT  
KKFLE**W**LFNNDTRIAAPDSPFNQSVSQILNSVSAYDLTSKQALSTTAQQALTDIAKEQAEVDK  
VKNNPTPETYNNSD**W**SRIYLNKGALASLTD**W**LTFKSELNKDPANNSIAFIHTDEKANRILSILNGA  
VSGITRSADPTNLTKENLLRQINDALGTADQ

MGA\_1161

Producto de PCR

acatgcatgcgtttttaaatttttcagcaacctaccaacaacaacggacaataataatttgaccaacagtcacaaatcttagtgcgtttt  
accaactagtgtattataagaagatttatgatcttagtttagtttagaatttaataattctggtggttataacccttcacgtttgatgtcaataat

gaaggagtaactgctgaaaacggttaactgctcctaccggtgttttgaacgggatgattattgattgacaagctcaaccggtgatgagaa  
cgatcaaaaatgcgaaatgaaccggatatttgcgactaatttacacgtagctgaagcggtattaaatcctttagataataagaactatcgc  
ccagcttgatataagaatgaactaccctaccaggggtgatcaaaccttttattcaactaggtaaatgggatgaaaatctagctgtac  
aaaacaaacacaacccaaaatctttaacttactaccggttaagtaacttacctaaaacagttatacagcaactcttttataaagaatca  
ccaaaatgaattactccgatcgaagccgataaccaggaattcgcgagtacattgatttctgtgtttatcaatcacttttagatctttcatgaa  
cgatgagaaacggagagaaagttcatcgttataatgatgaataccaactgtacaaccggtgaattatgccagcaatgaatgttgctaaa  
gcattatgagataacaaggagttttaaataaccagcaactccacctaataagaagaagaagatgatcaattaactgatcaattac  
cttaccacggttttttgatcgttcaaattacctaaccacaaacgtaagtttaataacttttagtgtttatctgtgtgtatccttattatgtaact  
gaccaacaactccacaatacactaaatttagtgtccatcattatcaagaagacagatcccaatcactagtgattcaagaggttctcctg  
gttgaacgattaacgccttaaatcaactgaactagatggacaaaagatcgctagttcaacatctctaagctcaactggtggaattaga  
gcagggtattataatgctgatattgctagaaattccaattagttatcgtaattgtaatacaacaataatggtatggctatatttcaaaa  
ctcaacttatcagcaggtcatctggatcattagcattaactcaataaccaagcttaggaatctatttggtagcgtgtcagttgatgcc  
aaaaagaagctaccttgggttagtgcctcattttaatccaaataagatccaagtaattgtgttactggttaataacggttttgaaactgat  
acgatccaacatgatctaatatttgtaacgatctaactgctgatgattatggttcatacattagatcattacaaaccttaaaactatcat  
cacgtttattaacaaaaattaaagataatttaccacaaaaagaaggttcaactaatcaagctaaccaacaaacaaacaaactaata  
gatcaactgatgcaacaaaaaagacagttcatcagatgagactaacaagaaccattagctgaattattatcagatatcttcaaaaat  
ttacctaactgaaattaacccggggga

Polipéptido recombinante (63,4 kDa)

MRGSHHHHHHGSACVFNFSSNLPTTDDNNNYGPTVQNLSAFLPTSDYKKIYDLSFSLEFNNSG  
GYNPSRLMSNNEGVTAEVNRPYRVFGTGWLFDWQAQPVDENDQNAKWTFYFATNLHVAEA  
LLNPLDNKNYRPAWYKNELPLPGVDQTLYFNLGKWDENLAVQNKHNPKSLTYLPLSNLPKTVY  
TATSFYKESPKWITPIEADNPGIREYIDFAVLSITLDLSWTMRNGEKVHRYNDEYQLYNRWIMPA  
MNVAKALWDNKGVLNQPATPPNKKEEDDQLTDQLPYHGFFDRSNYTNPNVSLNNFSVYLGGY  
PYYANWPTTPQYTKFSVPSLSRRQIPITSDSRGSPGWITINALNPTELDGQKIASSTLSSTGGIR  
AGIYNADIARNFQLVYRNVKYKQYGYGYIIQNSNLSAGSSGSLALTSNNQALGIYFGTVSVDKAK  
EATFGLVASLFNPNKIQVNVVTGNNVFETDIQPYDLIYGNDLMTDDYGSYIRSLQTLNLSSRLLT  
KIKDNLPKKEGSTNQANQQTNQTNRSTDATKKDSSDETNNKPLAELLSDIFKNLPNWN

crmA

Producto de PCR

acatgcatgcaagcaatcagataagagtaacgataacacgcaattagttaatcaagcaagaacgctagatgctaattctgttagacttg  
caggctctggacaaaatggttcgttgcataacagttcttagagatggttgatgataactttataacagcagctaattggaacaattatcaaat  
tagatagtttactaaaccattatattggttagatctaagtgatgatttctggtatgatacaaaagtaaaacaaatagtttcagattacacaactag  
cagaaatagatttgatcaaaagcaacaagagcatattatgctctgttggttaattgatgaagtaacggtcatttaaaaagaattaact  
aactcaaatagaaattggtatagaacaacaattctaagttgtaattggtggttgataatccagctcacgtaattagattactgatgat  
gggactaaatttaatttcaaaagcaaaactcaagtgaaattgtaattgacttcatttagatgcgccaatcttacctaaagatttacacca  
gattgatataactttacattcaaaagaaagatcttaccaaatgacgttaaactgcagttgttccttgaccagtaggttagagtttaggaa  
caaatgctgatgatgggatgtttgatttgggaatggtcaaaataactaatacagatcctattgctcaaaactaaaaccactactgataatcaa  
aatccttcaacttttaattcaggagcaatgcctggtgcaacaataagatacagattctcaattgaattgcaagcatagaattaaaacatctt  
ccaattagatgaaaaattgtttatccagaatgaactggttcgaagagaataaaaaatattacaagattagctactggaagtttgcaagc  
aacgaaagatattgaattcttgacataccagggactccacaagttactttaaagaagattcagtaacgtaatttcaagactatacttaaa  
ctcagttaatctttatcattcattggtgatagtatttatttgggtactctgaattaccatcattatgatactattcattccaactagattatctg  
atctaaccgcttgaatcaagttaaaacagatgatattgaagctcaagcactgataacggtacaacaacaacggaacaacgacaa  
caactgatacatctagtggttcaacaggtgctggaacaggaaataactactaactcctcaaacagtttcaatcctacttttaataacttct  
gtagtttggaaattgatagtaaaccactctgcaacaaaatagatgaaactaattgagcagatcctaattgtattgaagcaagaatata  
tgctgaatacagattaggtattcaaatgaaattccaataactaatgcaggaaactttatccgaaacacaattggtggtgtgtgtttacttc  
aacagggtcaagagtagtttaagagcttctataacggtgatcaacgtccaactggaaaactccaaccttcttatacgtatttgggtatttag  
gataccaacaaactagaacaggaactttctgatacggaaacataaaactattaaacaacagcccttacgacgtattagatgctgcaag  
agtaggtactgaaaccaatcaatttagaagaactcattaacataccctgttatgggtggatataactgaagaagggtgctagaagtttc  
tctaatactccatataagagcacaaggtgacacaccagaaagccgaagcatctccaatctggtattctgataatacttatgagtac  
attcaatcagtttaggttaggaattagaataacttaaatgttgggttaaagcatcaagcttctaaactcaaatagaccaaatcc  
aaacggtctagaaatgattgctgcaacaacatacttaagatcacaattggattagctagaacatctggattaccaaaccaacaacca  
ttcgaacaactaccaagttattcagatcacctggtgatcagttctcatcaattaagaatattagaacaatctccctggttaaccagttat  
gatacttctattcacaaatgaaaataataaatctagtggttatacattaagattagctgactcaagtaacctgtatgctgcaagctcattca  
gtccaacaagtttaattgacgtaattgaaattggtgtaattcttacccttattagacaattcattctatacagtaaatgctgctggttaattgcat  
tgttctcatcaaacctggttctcctggatcatatactgctgtaaatcatttaatacagaacttatctgatattgctttgaaggttctggtgctaa  
atatacatctgatttctgaggaacaatccaattcaaaccgcatgagtaacttaaaatgggttactagtaagtggtgtagaaacttc  
gttacaacaaacagcttctaaacagtttagtgacttcactcctgctaatgctggttactaactaccgtgtagtggtgatcctgatggtaatt  
aacaacaaaacactacctctaaaagttcagatccaatacttagatggaagtattatgatgtaaatgaagaacaataatttagtaac  
attctctataacaactttgctgtttaccttcatgagtagtgctacacccggggga

Polipéptido recombinante (101,6 kDa)

MRGSHHHHHHGSACKQSDKSNNDNTQLVNQARTLDANSVRLAGLGQNGSLFNTVLRDVDNFI  
TAANGTIIKLSFTKPLYGLDLSDDFAGYKVKQIVSDYTTSRNRFQDQRTTRAYYALLVNDEANVH  
LKRINTNSNRIGNRNNNSKFVIGGVDNPAHVIRFTDDGTFKNFTKQTQGEIVNDFILDAPILPKDL  
HPD**W**YNLYIQRKILPNDVNTAVVP**W**PVGRVSGTNADDGMFDFGNGQITNDPIAQTKTITDNDQ  
NPSTFNSGAMPGANNRYDSQLNVKHRIKTSFQLDEKFVYPE**W**TGSEENKNITRLATGSLPSNE  
RY**W**ILDIPGTPQVTLKEDSVNVFSRLYLNSVNSLSFIGDSIYIFGTSELP**W**YYSFPTRLSDLTAL  
NQVKTDDEIASSTDNGTTTNGTTTTTDTSSGSTGAGTGNTTNTSQTVSNPTLNTYRSFGIDSKP  
TSANKIDET**N**WADPNVIEARIYAEYRLGIQNEIPITNAGNFIRNTIGGVGFTSTGSRVVLRSYNG  
DQRPTGNFQPFYLVFGYLGYYQTRTGTF**W**YGTYKLLNNSPYDVLDAARVGTETNQFRRTSLTY  
PVMGGYLTEEGARSFSNTPYIRAQGDTPESRSIFQSGYSNTYIEYIQSVLGFDIRNNLNVGVK  
ASSFLNSNRPNPNGLEMIAATTYLRSQLGLARTSGLPNQPPFGTTHQVISVSPGDQFSSIKNIRTI  
FPGNQL**W**YFLFTNENNKSSVYTLRLADSSNPDASSSFPTS LIDVNEIGVILPLLDNSFYTVNAA  
GNVALFSSNPGSPGSYAVNTFNQNLSDIAFEGSGAKYTSDF**W**GTIQFKPDEYLIQNGFTSQVA  
RNFVTNQSFNLVDFTPANAGTNRYRVVDPDGNLTNQNLPLKVQIQYLDGKYDDAKLKNNNLV  
TFSYNNFAALPS**W**VVPTPGSTCSQA

crmC

Producto de PCR

acatgcatgcaatcattctctagatcagataaattagtaaccaagcacagtagtgctgataataatcaggattatcagcaactaaaa  
accgaacaaaaatcattatgaactaatagtaataacgattatttctcaccacaaactgggtgtgtaacttgatgttttgcaatgtttatt  
ataactacagtttgcgaccaactcattaattatcaaaccgtgatttctgctgatagtttaagagctgattattacttattgatcaaaaa  
cccgattgttagtattaatgggaacgacttaaaccttaagtgtcagttcacccgcttatgtgtgcaactaaaagatgagaataaaacttt  
aaatgtgcttaagacgttccaattaagtcttatcggttctcaagaaatattgcgacctatacaatcggttgcgttgaactcaactcaagt  
agttcaaacctatcagggtcattagataaaccactaatagcacgttataaataatattcagcatattatgatcatgcatgtagtaatta  
gtgtttattaataatcgcttaggggtgtcgttagtaatgatcaattatccttatggtttatctatttaataatcaaaccaacgacttagtac  
gaataacgacttaaatgtgcacaatcttcccaattatattgggacaaccaacttaattgagcgtaataataatagtaattgtaataca  
agcaactacagcacagttgtgattaacaacattcgtttatacaccagcttactcgttaattggggcaaatgtgtgaagcaaaatttag  
ttaatcctaacggattattaacgccaacaaactgcgatctattcattacatttagggtagtgccagcacctattagtcgactactaata  
acaatcaattcggttataagattagtgtaactactacaatagtgataatcttaatttccaacctattttagtggttgatcagtcacaaact  
aaagtgccagatcactaataatgttaattcttttaagagttcttagtgctctcacacaacaagactaatctatcaagattaattctaag  
acccgaaccgcttatcactattcagttcaagaagatcggttttagtcaaacacaataagttatccaacctagttgggatcttaatacattaa  
agatgagatgttgggttttatactcagctgatagttataatctgctctcacagattaactattatctgattagtgatgatcagtagtgctg  
gtttgatctattaacaactatcgactaaaagatgggaaatgatacctaagtttcttgatccaagacaataatagttggactaatcagat  
tagccaacaattagtaacttaataatgggacattagcaattaactataactaatgtatcaagttaaccaacaaggatcgatcttcttaa  
ccctaatttgggtgagcaagcttttatacttttagtaactcaaaaataatagagtagatctcaacctagcttttatgatcagaaaacccaaa  
gttttctgatgccaccttcttaataacgggtgtgcctagttttgataatatcccaaaactcagatgggtcaaaattatgggggatgattagttta  
aacatcttcatattagtttaataagatctaaacacaactagtgaatgagattaagcaataataattattgcttaataacttatatgatta  
cactcctggtgagatcgattgaaaacctagagtagaagctattgcactaaacgattcaaccttaagattatcaatccaatttcttatttga  
tggaactattacagtcaccaacaaaagttggtaaacttagcattccaacaattatctatgataatcttaagtcttatccggttatgtgtgc  
ccggggga

Polipéptido recombinante (76,3 kDa)

MRGSHHHHHHGSACNHSSRSKLVNQAQLVADNNQDYQQLKTEQNHVYVNTNSNNDYFVLTKT  
GVVKLDVFGNVYYNYSFATNFINYQTVDFVADSLRADYLLIKNPIVSINGNDLNSLSVSSPAYV  
LQLKDENKTLNVLKTFQLSSYRFSRNIAATLYNRIAFESTSSSSNLSGLLDKPTNSTLLNNISAYYD  
HAISKLVFINNRLGVFGSNDQLSL**W**FYLFNINQTNDFSTNNDLNLVTIFPQLYGTTNLN**W**ALNNIS  
NVNQATTAQFVINNIRLYTPAYFVIGANVVKQNLVNPNGLLNAQQTAIYSLHLGLVPAPISAITNNN  
QFGYKISANYNSDNLNYSNLFVSGVISQTKVAQITNNVNLNFEFFSASHTTRLNLSKINSKTRTA  
YHYSVQEDRFSQTNNSPNLVGILIKDEMLGFLYSADSYNLLSQINYLLISDDISMAGFDLLTTIA  
LKDGK**W**YLSFFDPRQSIS**W**TNQISPTISQLNNGTLAINYTKLYQVNQQGSIFLTLILGEQAFYTLV  
TQNNRVDLNLVFDQKTQSFSDATFFNNGVPSFDNIPNSDGSKL**W**GMISFKPSSYLVNQDLNTT  
SAMRLSNNNLLLNNLYDYTPGEID**W**KPRVEAIALNDSTLRLSIQFPYFDGNYSPNQKLVNLAFP  
TIIYDNLNAYPVYVVPSTCSQA

gapA

Producto de PCR

acatgcatgctctggagctaagccaaataacctaaccctgtaaccaagttggggaaatgaattcacaaggccaatctaattctttaga  
gaaagctcgtagatgaagaaacttaacttcacatcatttcaattgacgggtaccaacccagggtcattagtttaactggatcaaaatca

MRGSHHHHHHGSACSGAKPNNLKPVNQVGEMNSQGQSNLLEKARRWRNSNFTSLSIDGTNP  
GALVLTGSKSISRIDL YGNV**I**WTFDPGNTNDLTGKVGFDYDANNRLTAFSGDVPFNVSDLSSKTVV  
EATQDQEDPNVFYLLIPDAAVQQEQKTKDQVFENYFMSDAPATGDTSAEGSATPAGGGSSSS  
AAGGGAVAPAAASSTARLVEEGNSAGMGTMPTASTSETVIDYNSDQNKIPKPKTLDDSSSESE  
SINGGRTYANINTQNNLQGVIVKVNENLFSNENPFAVENMAFIKPKMDVDNYPST**W**TQGSANGK  
MTNVLQFYKHDNPNAVNNRFYRAKYYPKRLETQTTTPLIDSSFSPYEHPE**W**YEDNQFVMP**W**M  
QYITNLGGLYAKDGMVYLFGGNGT**W**VNNESALSIGVFRTKFENRTAEAPGNTKTVGYPYGILLS  
AISFDATRNLGALAPASLGQDVG YHFVPRLAVGGVSSPRGANGNIFLGSAIT**W**GTNGGNFLDTK  
**W**HSPAVIEDAPTTFVTVNSSGALQNSGNPQPTSTPMPNSNGNESIPYR**W**TNSYDYNVSRFAALI  
SKPAGGNTKQVESLFTTALKLDTLNSLPNKFTQENNIFFSYAMLDGRQ**W**SLGTRKDSAW**L**TTNT  
INNFTYNTQQQLASTVAGENANPRNILNALTAKGFDRRDIGNVVYTYSNNTNKFTYYYQVGGAI  
TT**W**PQVQVNYKTSANITYYNLTRDFGSTTPATQDANTVSSKLN GAYLSSTGDQQG**W**YNGSIY  
VKKASFTPSSQGYT**W**QDFKGLTTTASNAVISN**W**TKAGYSIRPDDDTVFNVSKIPFEKEITA AVNV  
RSLDSYYVQLNGETSVNTVARVSPDSSALALNPNRITNPLMNRDNVIGQGAFISRNDIPSSFFEN  
KINDIVTTEADGKEVLD SKYINSIYRTPPQNNPDIRLLVIDRSRATNDFIKLLPQVLVDGEYVA  
VPQANSVFVSDQEFTGFDALPGYPGSTCSQA

acatgcgatgcgaacattcacccacacttggtttacatgaaggtagaaactaactccgttggtccaagaagattacctccgaaccctgatt  
tatccagtagtgggtgcaggtgctgggtgattgtgtttcttactcttgggttagggattgggattccgatcgctaagaaaaagaaaga  
atgatgatccaagaacgtgaagaacacaaaaagatggtgaatcccttgaataattgaagaacaaaataaacagaagcgattga  
gccaactgaagaagttaatactcaagaaccaactcaaccagctggtgtaattagtagctaataaccctcagatggggatcaatcaacct  
cagattaatccgcaatttggctctaatccccacaaagaattaaccacagtgcttgggtggccaatgcaacctaaccaaatgggaat  
gcgaccagggtttaaccaaagccccacaaatgggaggaatgccacctaaccaaatgggaatgcgaccagggtttaaccaaag  
ccccacaaatgggaggaatgccaccaagaccaaacttccctaaccaaatgctaataatgaaccaaccttagaccagggttcagacc  
acaacctggtggtgggtgccgatgggaaataaagctggagggtgggttaatcaccaggtacaccaatgggtccaaaccgcatga  
acttccctaataaggaatgaatcagccccacacatggcaggaccaagagctggtttccaccacaaaatggacctagataacccg  
gggga

Polipéptido recombinante (29,5 kDa)

MRGSHHHHHHGSACEHSPTLVLHEGETNSVGPRKITSEPW<sup>W</sup>FYPVVGAGAGLIVVSLLLGLGIGI  
PIAKKKERMMIQEREEHQKMVESLGIIEEQNKTEAIEPTTEEVNTQEPTQPAGVNVANNPQMGIN  
QPQINPQFGPNPQQRINPQCFGGPMQPNQMGMRPGFNQMPPQMGGMPPNQMGMRPGFNQ  
MPPQMGGMPPRPNFPNQMPNMNQPRPGFRPQPGGGVPMGNKAGGGFNHPGTPMGPNRN  
NFPNQGMNQPPHMAGPRAGFPPQNGPR

oppA

Producto de PCR

gggggtaccattactaatgatcatgtaataatgcattaacatattgattagagtagatcggtgttttcaaaccctgggttaactttattaat  
ctagcgatgattgcagctaagtgtactgtactaaacgtatttaagttaaatgcgttaatgagttacagcattattataagccgctagcata  
gctggatcgctattgacgatcatcataatccgatcaatatcagggttttaagatcggttctgttaagattggatcagttttggagtgcttcta  
gtactaaattagataagtaatactagatatgtgtatgatccgactgggttttcttgaaccgcatatttttgtctgaaactatcaatcccatcggtt  
ttaaagaaggttagtaccatctgtggttgccaccactaaaacggctgtagttttgataaatgatatcgtagctaccttttcaatgtagc  
tggcaaaagatatcttctgtaatagttttctcaatattgacatacccaccaaacgcttctagcatcgctgttttagataagtcgccccacga  
actgttcgctactactattatctaagaagtttagggttactgttttagttctgggtgttttagctttaaacggcttagataaaatcttgctgtttta  
aatcatgagctagatcatgacgaacggcttttcaaagttaaatcttttagataagtgacaacaaagtcatagttttagtagatcaaattctt  
tattgttctagtttagatttttggccttcaaaaaatagtcttagatttcgaccatcataagctttgtgggtccataagctgtgagctcgg

Polipéptido recombinante (35,2 kDa)

MRGSHHHHHHGSACELTAYGTHKAYDGRNLELFFEGQKSKTKNNKEFDLLNYDFVHLSKRFN  
FEKTVRHDLAHDLKTARFYLDRFKAKHPELKQVTLNFDNSSDEQVRGATYKQAMLEAFGGY  
VNIEEKSLPENIFASYIEKGQYDIYQNYDRFSGGQPQDGVLTFFKNDGIDSFRQKNIGFKENPVG  
SYTYLDYLSNLVLEALQKTDPNLTRQTILKPDIDRIMMIVNSDPAMLAAYNNVNSLNAFNLNTFS  
TDHLAAIARLNKVNPGFENTRYSNQYVNALLDYMIISNGTPGRPAAKLN

vlhA.5.02

Producto de PCR

cccccggggaataagaaaagaatcatcttaaagacttttagttgttaggtgcaacatccttcttagtattgggatttctagctgtatgtctatt  
actaaaaagactcaaaccctaaataatggccaaacccaattacaagcagcggaatggagttactgatctaataatgctaaagc  
aaggacattagcttactacaagactatgctaagattgaagctagttatcatctgcttatattgaagctgaaacagttacaataaccta  
atgcaacactagaacaactaaatatggctaaaactaatttagaatcagccatcaaccaagtaatacggataaaacgacttttgataat  
gaacacccaaattagttgaagcatataaagcactaaaaaccactttagaacaacgtgctactaaccttgaaggttagctcaactgct  
tataatcagattcgtaataatttagtggatctatacaataaagctagtagtttaataactaaaacactagatccactaaatgggggaacgc  
tttagattctaatagagattactacagctaataagaatattaataatcagttatcaactattaatgaacaaaagactaatgctgatgcattag  
ctaatagttttataaagaagtattcaaaataataaacaagtttttaggaatgtttacaacactaatgttcaaccttcaaactatagttt  
gttgcttttagtctgatgtaacacctgtcaattataaataatgcaagaagaacgggttgatgaaagcttg

Polipéptido recombinante (30,7 kDa)

MRGSHHHHHHGSACELGTPGNKKRIILKTFSLLGATSFSLSIGISSCMSITKKDSNPNNGQTQLQA  
ARMELTDLINAKARTLASLQDYAKIEASLSSAYIEAETVNNNLNATLEQLNMAKTNLESAINQANT  
DKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLASTAYNQIRNNLVLDLYNKASSLITKTLDPNG  
GTLLDSNEITTANKNINNTLSTINEQKTNADALANSFKEVIQNNKQSFVGMFTNTNVQPSNYSFV  
AFSADVTPVNYKYARRTV

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

## Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

Query = pQE30-gapA-17

Length = 2907

score = 5369 bits (2907), expect = 0.0

identities = 2907/2907 (100%), gaps = 0/2907 (0%)

strand=plus/plus

pQE30-gapA-17	1	tctggagctaagccaaataaccttaaacctgttaaccaagttggggaaatgaattcacaa	60
gapA	1	tctggagctaagccaaataaccttaaacctgttaaccaagttggggaaatgaattcacaa	60
pQE30-gapA-17	61	ggccaatctaatacttttagagaaagctcgtagatgaagaaactctaacttcacatcactt	120
gapA	61	ggccaatctaatacttttagagaaagctcgtagatgaagaaactctaacttcacatcactt	120
pQE30-gapA-17	121	tcaattgacggtaccaacccaggtgcattagttttaactggatcaaaatcaattagccgg	180
gapA	121	tcaattgacggtaccaacccaggtgcattagttttaactggatcaaaatcaattagccgg	180
pQE30-gapA-17	181	attgatttgatggttaacgtgatttgaacgtttgatccaggtaacacaaacgatctaact	240
gapA	181	attgatttgatggttaacgtgatttgaacgtttgatccaggtaacacaaacgatctaact	240
pQE30-gapA-17	241	ggtaagggttgattttatgatgctaacaatagattgactgcattttctggagacgttcct	300
gapA	241	ggtaagggttgattttatgatgctaacaatagattgactgcattttctggagacgttcct	300
pQE30-gapA-17	301	tttaatgtaagtgatctaagctctaaaacagttgtagaagctactcaagatcaagaagat	360
gapA	301	tttaatgtaagtgatctaagctctaaaacagttgtagaagctactcaagatcaagaagat	360
pQE30-gapA-17	361	cctaattgttttctacttattattaattccagatgcagcgggttcaacaagaacaaaagact	420
gapA	361	cctaattgttttctacttattattaattccagatgcagcgggttcaacaagaacaaaagact	420
pQE30-gapA-17	421	aaagatcaagtgtttgaaaactactttatgtctgatgcacctgctactggtgatactagc	480
gapA	421	aaagatcaagtgtttgaaaactactttatgtctgatgcacctgctactggtgatactagc	480
pQE30-gapA-17	481	gctgaagggttctgcaactcctgctggtggtggttcaagtagtagtgctgctggaggaggt	540
gapA	481	gctgaagggttctgcaactcctgctggtggtggttcaagtagtagtgctgctggaggaggt	540
pQE30-gapA-17	541	gctggtgctcctgctgctgctgcgagttcgactgctagacttggtgaagaagggaatagtgt	600
gapA	541	gctggtgctcctgctgctgctgcgagttcgactgctagacttggtgaagaagggaatagtgt	600
pQE30-gapA-17	601	ggtatgggaacgatgactcctactgcttctacttctgaaacagttatagattataatagc	660
gapA	601	ggtatgggaacgatgactcctactgcttctacttctgaaacagttatagattataatagc	660
pQE30-gapA-17	661	gatcaaaataaaaattcctaaacctaaaacactattagacagtagcgaaagttctgaaagt	720
gapA	661	gatcaaaataaaaattcctaaacctaaaacactattagacagtagcgaaagttctgaaagt	720
pQE30-gapA-17	721	atcaatggtggaagaacatatgcgaacattaacactcagaataatttacaagggtgttatt	780
gapA	721	atcaatggtggaagaacatatgcgaacattaacactcagaataatttacaagggtgttatt	780
pQE30-gapA-17	781	gtaaaagttaacgaaaatttattttaattcggaaaatccctttgcagtagaaaaatatggcg	840
gapA	781	gtaaaagttaacgaaaatttattttaattcggaaaatccctttgcagtagaaaaatatggcg	840
pQE30-gapA-17	841	ttcattaagccgaaggatatggttgataattatccttctacttgaacacaaggttctgct	900

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

gapA	841		900
pQE30-gapA-17	901	ttcattaagccgaaggatatggttgataattatccttctacttgaacacaaggttctgct	960
gapA	901	aacggtaaaatgactaacgttcttcaattctacaaacatgataatcctaagctgtgtaac	960
pQE30-gapA-17	961	aatagattctatagagcaaaaatactatcctaaacgttttagaaactcaaacaactactcct	1020
gapA	961	aatagattctatagagcaaaaatactatcctaaacgttttagaaactcaaacaactactcct	1020
pQE30-gapA-17	1021	ctaattgatagttcttttctctccatatgagcatccagaatgatatgaagataatcaattt	1080
gapA	1021	ctaattgatagttcttttctctccatatgagcatccagaatgatatgaagataatcaattt	1080
pQE30-gapA-17	1081	gtaatgccgtgaatgcagtacataacaaatttaggtgggtttatatgctaaagatggaatg	1140
gapA	1081	gtaatgccgtgaatgcagtacataacaaatttaggtgggtttatatgctaaagatggaatg	1140
pQE30-gapA-17	1141	gtgtacctatttgggtggaacgggtacatgagttaacaacgaaagtgcattaagtattggt	1200
gapA	1141	gtgtacctatttgggtggaacgggtacatgagttaacaacgaaagtgcattaagtattggt	1200
pQE30-gapA-17	1201	gttttcagaactaaatttgaaaacagaactgctgaagctccaggaaacactaaaactggt	1260
gapA	1201	gttttcagaactaaatttgaaaacagaactgctgaagctccaggaaacactaaaactggt	1260
pQE30-gapA-17	1261	ggttatccatacgggtattttattatcagcgatttcttttgatgctactagaaatggatta	1320
gapA	1261	ggttatccatacgggtattttattatcagcgatttcttttgatgctactagaaatggatta	1320
pQE30-gapA-17	1321	gcgcttgctcctgcaagtcttgggtcaagatggtggttatcactttgttcctcgtcttgca	1380
gapA	1321	gcgcttgctcctgcaagtcttgggtcaagatggtggttatcactttgttcctcgtcttgca	1380
pQE30-gapA-17	1381	gtgggtggtgtaagttcacctagaggagctaacggtaatatcttcttaggttcagctatt	1440
gapA	1381	gtgggtggtgtaagttcacctagaggagctaacggtaatatcttcttaggttcagctatt	1440
pQE30-gapA-17	1441	acttgaggaacaaacgggtggaatttcttagataactaaatgacacagtcctgctgcatt	1500
gapA	1441	acttgaggaacaaacgggtggaatttcttagataactaaatgacacagtcctgctgcatt	1500
pQE30-gapA-17	1501	gaagatgcacctaactacttttgtgactgttaatagtagtggtgcgcttcagaatagtgga	1560
gapA	1501	gaagatgcacctaactacttttgtgactgttaatagtagtggtgcgcttcagaatagtgga	1560
pQE30-gapA-17	1561	aatccacaaccaacttctactccgatgcctaataagtaacggtaatgaaagcatcccttat	1620
gapA	1561	aatccacaaccaacttctactccgatgcctaataagtaacggtaatgaaagcatcccttat	1620
pQE30-gapA-17	1621	agatgaacgaattcttatgattacaactctgtaagatttgcagctctaattagtaagcca	1680
gapA	1621	agatgaacgaattcttatgattacaactctgtaagatttgcagctctaattagtaagcca	1680
pQE30-gapA-17	1681	gctggtggaacacaaaaacaagttgaatcattattttacaaccgctttaaaattagataca	1740
gapA	1681	gctggtggaacacaaaaacaagttgaatcattattttacaaccgctttaaaattagataca	1740
pQE30-gapA-17	1741	ttaaattctttaccaataaatttactcaagaaaataatatcttcttttagttatgctatg	1800
gapA	1741	ttaaattctttaccaataaatttactcaagaaaataatatcttcttttagttatgctatg	1800
pQE30-gapA-17	1801	ttagatggcgtcaatgaagtttaggtacacgaaaagacagcgcatgattaacaactaat	1860
gapA	1801	ttagatggcgtcaatgaagtttaggtacacgaaaagacagcgcatgattaacaactaat	1860
pQE30-gapA-17	1861	actattaataacttcacttataataactcaacaacaattagcgtctacagtagcaggagaa	1920

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

gapA	1861	actattaataacttcacttataataactcaacaacaattagcgtctacagtagcaggagaa	1920
pQE30-gapA-17	1921	aacgctaataccaagaaatatcttaaacgctttaacaactgcaaaaggggttgatcgaaga	1980
gapA	1921	aacgctaataccaagaaatatcttaaacgctttaacaactgcaaaaggggttgatcgaaga	1980
pQE30-gapA-17	1981	gatattggtaatgtagtatataacttatttctaataataactaataagtttacttattactat	2040
gapA	1981	gatattggtaatgtagtatataacttatttctaataataactaataagtttacttattactat	2040
pQE30-gapA-17	2041	caagttggtggcgcgattacaacttgaccagaagttcaagtaaattacaaaacttcggct	2100
gapA	2041	caagttggtggcgcgattacaacttgaccagaagttcaagtaaattacaaaacttcggct	2100
pQE30-gapA-17	2101	aatattacttactacaatttaactagaactgattttggaagtactactcctgcaactcaa	2160
gapA	2101	aatattacttactacaatttaactagaactgattttggaagtactactcctgcaactcaa	2160
pQE30-gapA-17	2161	gatgcaaataccgatatcatctaaattaaacggcgcttacttatcatcaactggcgatcaa	2220
gapA	2161	gatgcaaataccgatatcatctaaattaaacggcgcttacttatcatcaactggcgatcaa	2220
pQE30-gapA-17	2221	caaggatgatacaacggttcaatctatgttaaaaaagcgagctttacaccaagtagccaa	2280
gapA	2221	caaggatgatacaacggttcaatctatgttaaaaaagcgagctttacaccaagtagccaa	2280
pQE30-gapA-17	2281	ggttatacttgacaagatttcaaagggtttaacaactacagcaagtaacgcagttatttct	2340
gapA	2281	ggttatacttgacaagatttcaaagggtttaacaactacagcaagtaacgcagttatttct	2340
pQE30-gapA-17	2341	aactgaacaaaagctggatacagtagtagaccagatgatgatacagtagtattcaacgtttct	2400
gapA	2341	aactgaacaaaagctggatacagtagtagaccagatgatgatacagtagtattcaacgtttct	2400
pQE30-gapA-17	2401	aagattccttttgaaaaagaaatcacgcgctgctgttaatgtaagatcattagatagttac	2460
gapA	2401	aagattccttttgaaaaagaaatcacgcgctgctgttaatgtaagatcattagatagttac	2460
pQE30-gapA-17	2461	tatgtacaattaaatggtgaaacttcagtttaatactgtagctagagtaagtcctgattct	2520
gapA	2461	tatgtacaattaaatggtgaaacttcagtttaatactgtagctagagtaagtcctgattct	2520
pQE30-gapA-17	2521	agcgcttttagccctaaaccctaacagaattactaaccattgatgaatagagataacgta	2580
gapA	2521	agcgcttttagccctaaaccctaacagaattactaaccattgatgaatagagataacgta	2580
pQE30-gapA-17	2581	atcgggtcaaggtgcgttcattagtagaaatgatattccatcatcattctttgaaaacaaa	2640
gapA	2581	atcgggtcaaggtgcgttcattagtagaaatgatattccatcatcattctttgaaaacaaa	2640
pQE30-gapA-17	2641	attaatgatattgtaactacagaagctgatggtaaagaagtattagatagtaaatacatt	2700
gapA	2641	attaatgatattgtaactacagaagctgatggtaaagaagtattagatagtaaatacatt	2700
pQE30-gapA-17	2701	aattcgatctatagatatatactccacctcaaaacaatcctgatattagattaagattatta	2760
gapA	2701	aattcgatctatagatatatactccacctcaaaacaatcctgatattagattaagattatta	2760
pQE30-gapA-17	2761	gtaattgatcgttctagagcaaccaatgacttcattaagttattacctcaagtattagtt	2820
gapA	2761	gtaattgatcgttctagagcaaccaatgacttcattaagttattacctcaagtattagtt	2820
pQE30-gapA-17	2821	gatggcgaataacggttgctgttccacaagctaataagtggtgttgtgtctgaccaagaattt	2880
gapA	2821	gatggcgaataacggttgctgttccacaagctaataagtggtgttgtgtctgaccaagaattt	2880
pQE30-gapA-17	2881	actgggttttgatgcgcttccaggttat	2907
gapA	2881	actgggttttgatgcgcttccaggttat	2907



# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

Query = MGA_0648				
Length = 2181				
score = 4028 bits (2181), expect = 0.0				
identities = 2181/2181 (100%), gaps = 0/2181 (0%)				
strand=plus/plus				
pQE30-MGA_0648-26	1	cagaaaaatgcaagctccagcaacaaaaagtgcagaggagatatattcgcaactttaactaaa	60	
MGA_0648	1		60	
pQE30-MGA_0648-26	61	ggcgatattacgatcgcatTTtaataaaaaatcaggataaaaaccaagattcttcccagtgca	120	
MGA_0648	61		120	
pQE30-MGA_0648-26	121	attaatgagaagaactattTTtaattatttctactttcaaattagtcagagtgttgaccaa	180	
MGA_0648	121		180	
pQE30-MGA_0648-26	181	aaagaattagatcacattaatTTtgatgcgaagaactttgttcctaataaccaatcgggt	240	
MGA_0648	181		240	
pQE30-MGA_0648-26	241	actttatcatttgatctagatgtatcagttgttgacagacctaaagatattaaaaatttc	300	
MGA_0648	241		300	
pQE30-MGA_0648-26	301	gattttagaatcgatggctctgatgacaggtagtggagctagtagtggtagtggtcaagt	360	
MGA_0648	301		360	
pQE30-MGA_0648-26	361	tctgctagtagcggatcaggttcaggtaaatagtggtcaactacaactcaacctgagcac	420	
MGA_0648	361		420	
pQE30-MGA_0648-26	421	gatccaaatgactgtccaccagatactagtaaggctaataaaaacttcaagtagtgaaaaa	480	
MGA_0648	421		480	
pQE30-MGA_0648-26	481	acagttaatacaactagaccaactgaaagccaaagacaacctcgtgatagtagtaaacca	540	
MGA_0648	481		540	
pQE30-MGA_0648-26	541	cctagttttcctaattgggtttccaggattccctggatcaaataattactaagaaatttggt	600	
MGA_0648	541		600	
pQE30-MGA_0648-26	601	caaaaacaaaatagttatccagactatgtagcaagtagatacaaaactgtagacaaacaa	660	
MGA_0648	601		660	
pQE30-MGA_0648-26	661	ctaataataaaagaaatatgagaccgtactttctctataaggccagggaactttaccagt	720	
MGA_0648	661		720	
pQE30-MGA_0648-26	721	gataacggaacagctcatccaatagttgaccaaggtacaggatgagtttagattacgct	780	
MGA_0648	721		780	
pQE30-MGA_0648-26	781	aagaactcgtcaaacaacaatcaactaaaattattcattgcaactaacctacacggttatt	840	
MGA_0648	781		840	
pQE30-MGA_0648-26	841	ggaaattatggaaataactaactctaaagagtttagacgatttattatcttattcagaccca	900	
MGA_0648	841		900	
pQE30-MGA_0648-26	901	acaggtgttaaaccaagcggatttgctcttggtaaatctaataaaactccatcggtttact	960	

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

MGA_0648	901	 acaggtgttaaaccaagcgatttgctcttggtaaatctaataaaactccatcgtttact	960
pQE30-MGA_0648-26	961	cctcaacctaacaatcaaagtgttgatcaactaaaacaaaatgcaggcaccatgctttat	1020
MGA_0648	961	 cctcaacctaacaatcaaagtgttgatcaactaaaacaaaatgcaggcaccatgctttat	1020
pQE30-MGA_0648-26	1021	tattcatataaagaaggagggggcaaaattggtatagaagagggttaaaactactgaagct	1080
MGA_0648	1021	 tattcatataaagaaggagggggcaaaattggtatagaagagggttaaaactactgaagct	1080
pQE30-MGA_0648-26	1081	ttcagcactcctaaaattgtgtttgctgcagttgactatttagatgatcacgctttaaac	1140
MGA_0648	1081	 ttcagcactcctaaaattgtgtttgctgcagttgactatttagatgatcacgctttaaac	1140
pQE30-MGA_0648-26	1141	caatataaatctaaaaataatgaaaaatgaaaagaatgaaaagctaaaaccttaagtct	1200
MGA_0648	1141	 caatataaatctaaaaataatgaaaaatgaaaagaatgaaaagctaaaaccttaagtct	1200
pQE30-MGA_0648-26	1201	attaacggatctaattcaagtgcaggtactgtatttgctgatgaagaaaaagctagatta	1260
MGA_0648	1201	 attaacggatctaattcaagtgcaggtactgtatttgctgatgaagaaaaagctagatta	1260
pQE30-MGA_0648-26	1261	aagaaattaaccgaataacaatggtaagatttcattctatactgactttgggattttagaa	1320
MGA_0648	1261	 aagaaattaaccgaataacaatggtaagatttcattctatactgactttgggattttagaa	1320
pQE30-MGA_0648-26	1321	ttagacgttgatctatcaaaagcagatgaaaaacttaaggaatgaataaataattcaacc	1380
MGA_0648	1321	 ttagacgttgatctatcaaaagcagatgaaaaacttaaggaatgaataaataattcaacc	1380
pQE30-MGA_0648-26	1381	aaagctgtagattccttacgtaaatagaataaaaaactacacaacacttacctaactataat	1440
MGA_0648	1381	 aaagctgtagattccttacgtaaatagaataaaaaactacacaacacttacctaactataat	1440
pQE30-MGA_0648-26	1441	ccagaaacaggtaattatttccctacttttagattacatctctaaaaatctaggattagca	1500
MGA_0648	1441	 ccagaaacaggtaattatttccctacttttagattacatctctaaaaatctaggattagca	1500
pQE30-MGA_0648-26	1501	tcttgtaaacccttcttatgcatatggtattgataatgcaaaagatatctatatagccggt	1560
MGA_0648	1501	 tcttgtaaacccttcttatgcatatggtattgataatgcaaaagatatctatatagccggt	1560
pQE30-MGA_0648-26	1561	tatcctaagacctacaaagtcaaactttctgaatgcagaataatccagcagaaagatat	1620
MGA_0648	1561	 tatcctaagacctacaaagtcaaactttctgaatgcagaataatccagcagaaagatat	1620
pQE30-MGA_0648-26	1621	caagatataacacagggtagtggttaataatcaactacgtattcctcaacaagcttgacta	1680
MGA_0648	1621	 caagatataacacagggtagtggttaataatcaactacgtattcctcaacaagcttgacta	1680
pQE30-MGA_0648-26	1681	cctaataaagatctatttaacgtaccaagctatactcgttctacagactcggaacatca	1740
MGA_0648	1681	 cctaataaagatctatttaacgtaccaagctatactcgttctacagactcggaacatca	1740
pQE30-MGA_0648-26	1741	aaaacttgatctgatataaataatattcaaatttatacaaaactttgaaacagaccgttc	1800
MGA_0648	1741	 aaaacttgatctgatataaataatattcaaatttatacaaaactttgaaacagaccgttc	1800
pQE30-MGA_0648-26	1801	attgacagatacggatttaattacttcactaaattctcatcactatactatggagcttct	1860
MGA_0648	1801	 attgacagatacggatttaattacttcactaaattctcatcactatactatggagcttct	1860
pQE30-MGA_0648-26	1861	ggttctgtagcttacaatgactttgggtcaaattggttggtatctatgatgggggttgctagt	1920
MGA_0648	1861	 ggttctgtagcttacaatgactttgggtcaaattggttggtatctatgatgggggttgctagt	1920
pQE30-MGA_0648-26	1921	agcacaaaattttgggtgacgtttcaaactcttggaaacatttgcaccttagtacaagctgct	1980

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

MGA_0648	1921	agcacaaatTTTtGGtGacgTTTcaaTcTTggaacattTgcacTTtagtacaagctgct	1980
pQE30-MGA_0648-26	1981	aacattaaagTTcaatctgatgaaaacattactaattatgcttacaacttaattgacaat	2040
MGA_0648	1981		2040
pQE30-MGA_0648-26	2041	aaaggtTTccCGcaccaaactagatcataccgaagtaacttaagtctgttctaccctaac	2100
MGA_0648	2041		2100
pQE30-MGA_0648-26	2101	ggattTaccTgacTTTcaacatctggtacacTactcaaagtaactctaagtaaacagca	2160
MGA_0648	2101		2160
pQE30-MGA_0648-26	2161	ggattTaccTgacTTTcaacatctggtacacTactcaaagtaactctaagtaaacagca	2181
MGA_0648	2161		2181

Query = MGA\_0267  
Length = 1204  
score = 1397 bits (756), expect = 0.0  
identities = 756/756 (100%), gaps = 0/756 (0%)  
strand=plus/plus

pQE30-MGA_0267-8	449	attattgatggtgcaattaattcagttttattaactacacTTTcaactaatagagaaggt	508
MGA_0267	1039		1098
pQE30-MGA_0267-8	509	attattgatggtgcaattaattcagttttattaactacacTTTcaactaatagagaaggt	568
MGA_0267	1099		1158
pQE30-MGA_0267-8	569	cttgaaactTTTactattaactaacgTTaataaaaattcaacaatcaaacaagatacgat	628
MGA_0267	1159		1218
pQE30-MGA_0267-8	629	cttgaaactTTTactattaactaacgTTaataaaaattcaacaatcaaacaagatacgat	688
MGA_0267	1219		1278
pQE30-MGA_0267-8	689	atcgTTacaaaaacaaacaaccaacaatctaaattagaaatggctgctgctaaagctaatt	748
MGA_0267	1279		1338
pQE30-MGA_0267-8	749	gaaactTcaaaaacCTgtattaacaagattagtgaacgctCTTTatggtaattctgctaac	808
MGA_0267	1339		1398
pQE30-MGA_0267-8	809	atgccaggattTgctgtTaaaggTTtattccgctctgTaaacgatCTTaaagactcaatta	868
MGA_0267	1399		1458
pQE30-MGA_0267-8	869	gatgctgtTaaaactCTggtTTctgcaaattacactagctTaaagcgactTcatcaatgtt	928
MGA_0267	1459		1518
pQE30-MGA_0267-8	929	tacaaaatgactaacacaacatgaatggaaggTgttggtatCTTgtTaaacagtcaagct	988
MGA_0267	1519		1578
pQE30-MGA_0267-8	989	ctaaatggTTTaaacgatactgaaattgataaattagctgctacaaatgtTctgatgca	1048
MGA_0267	1579		1638
pQE30-MGA_0267-8	1049	gcagatacagCTTtattcactgaagTtaaaactgaactTaaagagattagtagctgcaact	1108
MGA_0267	1639		1698

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

```
pQE30-MGA_0267-8 1109 atggtgagatcaattacatctgataaccaagctagattagcttctatgggtgaaggtgct 1168
|||||
MGA_0267 1699 atggtgagatcaattacatctgataaccaagctagattagcttctatgggtgaaggtgct 1758
pQE30-MGA_0267-8 1169 caagaatttagcttctacgttaacactgcactttaa 1204
|||||
MGA_0267 1759 caagaatttagcttctacgttaacactgcactttaa 1794
```

score = 830 bits (449), expect = 0.0  
identities = 449/449 (100%), gaps = 0/449 (0%)  
strand=plus/plus

```
pQE30-MGA_0267-8 1 tcggaagctatgagtagaactaaatacaactttgctgcagacgctgattttagaaaagca 60
|||||
MGA_0267 1 tcggaagctatgagtagaactaaatacaactttgctgcagacgctgattttagaaaagca 60

pQE30-MGA_0267-8 61 gcatttggtaatatcggaacaattaaccaagttaaaaacgtttagaaaattcattacag 120
|||||
MGA_0267 61 gcatttggtaatatcggaacaattaaccaagttaaaaacgtttagaaaattcattacag 120

pQE30-MGA_0267-8 121 ttgccaaaaactgctccaactgttactatggctcctgatgcaggtgctgctgctccttct 180
|||||
MGA_0267 121 ttgccaaaaactgctccaactgttactatggctcctgatgcaggtgctgctgctccttct 180

pQE30-MGA_0267-8 181 gctggtagtgaagctgctgaagctgaagcttctacaggaactggtaatactggtgataca 240
|||||
MGA_0267 181 gctggtagtgaagctgctgaagctgaagcttctacaggaactggtaatactggtgataca 240

pQE30-MGA_0267-8 241 ggaactgctgctgctggtggatctggaagcagtggtatcttctgcaagtagcagtgctgga 300
|||||
MGA_0267 241 ggaactgctgctgctggtggatctggaagcagtggtatcttctgcaagtagcagtgctgga 300

pQE30-MGA_0267-8 301 agtggctctcaaggaggagaaactacatctactggaagcggcgatcatcttcaaatact 360
|||||
MGA_0267 301 agtggctctcaaggaggagaaactacatctactggaagcggcgatcatcttcaaatact 360

pQE30-MGA_0267-8 361 agtgggtcaaagcgggttctgcacaagctaacacagctggtggaagtgggtgcatcttcttca 420
|||||
MGA_0267 361 agtgggtcaaagcgggttctgcacaagctaacacagctggtggaagtgggtgcatcttcttca 420

pQE30-MGA_0267-8 421 aatagtggcagtc aaagtagttctacaca 449
|||||
MGA_0267 421 aatagtggcagtc aaagtagttctacaca 449
```

Query = MGA\_0416  
Length = 1494

score = 2760 bits (1494), expect = 0.0  
identities = 1494/1494 (100%), gaps = 0/1494 (0%)  
strand=plus/plus

```
pQE30-MGA_0416-3 1 tcgtctgctgttttcagacggaataatatgacttattcgaagtttttagatatcttaaaa 60
|||||
MGA_0416 1 tcgtctgctgttttcagacggaataatatgacttattcgaagtttttagatatcttaaaa 60

pQE30-MGA_0416-3 61 gatccacagaatcccaaaacttcagctaagaagttgtttatctaggttgaacccaaaac 120
|||||
MGA_0416 61 gatccacagaatcccaaaacttcagctaagaagttgtttatctaggttgaacccaaaac 120

pQE30-MGA_0416-3 121 gaaatcagacaaccaacccaacaagctaataaccctaatacaaaagagcttgttcaataat 180
|||||
MGA_0416 121 gaaatcagacaaccaacccaacaagctaataaccctaatacaaaagagcttgttcaataat 180

pQE30-MGA_0416-3 181 ttctcatatgctgttggatcttatgatagtaataacaccccaacctttgttaataacaat 240
|||||
MGA_0416 181 ttctcatatgctgttggatcttatgatagtaataacaccccaacctttgttaataacaat 240

pQE30-MGA_0416-3 241 gatagctcaatcgatctaacatttaagttgttttataatcaagctttcatatcgatttta 300
```

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

MGA_0416	241	 gatagctcaatcgatctaacatttaagtgtgtttataatcaagctttcatatcgatttta	300
pQE30-MGA_0416-3	301	aactcgattagtttaccagctatccattattctagtaacttatataactatctagaagtt 	360
MGA_0416	301	aactcgattagtttaccagctatccattattctagtaacttatataactatctagaagtt	360
pQE30-MGA_0416-3	361	aataaattagataaactgagcaactaatttattaccaactgaagcaaatcaagctgggtgtg 	420
MGA_0416	361	aataaattagataaactgagcaactaatttattaccaactgaagcaaatcaagctgggtgtg	420
pQE30-MGA_0416-3	421	cttcgttcattttatgaatcgatggcaaacgggttattaactggtaatgaaacagataat 	480
MGA_0416	421	cttcgttcattttatgaatcgatggcaaacgggttattaactggtaatgaaacagataat	480
pQE30-MGA_0416-3	481	tacatttttaccgccgattgatcttaagtttgaatttaaggaattaaccctgaattaagt 	540
MGA_0416	481	tacatttttaccgccgattgatcttaagtttgaatttaaggaattaaccctgaattaagt	540
pQE30-MGA_0416-3	541	gcaaaaagcgcgattttatgatccaagagcttacaatgttgattaccaactaaatgcaaat 	600
MGA_0416	541	gcaaaaagcgcgattttatgatccaagagcttacaatgttgattaccaactaaatgcaaat	600
pQE30-MGA_0416-3	601	gagattatgactgataacttattaaatacctcaaactacttaacgaagttatttgtcctt 	660
MGA_0416	601	gagattatgactgataacttattaaatacctcaaactacttaacgaagttatttgtcctt	660
pQE30-MGA_0416-3	661	tcaaatattaatatctcatttgcttattatatcgttggattaaatgggtgatgttgcccca 	720
MGA_0416	661	tcaaatattaatatctcatttgcttattatatcgttggattaaatgggtgatgttgcccca	720
pQE30-MGA_0416-3	721	accagggatgcttatattacatcaaccagtgatgtaagattccaagcgtagctaataaaa 	780
MGA_0416	721	accagggatgcttatattacatcaaccagtgatgtaagattccaagcgtagctaataaaa	780
pQE30-MGA_0416-3	781	tttcaatcaatctacaacaccaaattaactgctccagttatttaatttaaaacttaatgat 	840
MGA_0416	781	tttcaatcaatctacaacaccaaattaactgctccagttatttaatttaaaacttaatgat	840
pQE30-MGA_0416-3	841	cttgggtgtggttgtgagataacaatgttgaacaaagagctagcaacacacaaaccctagc 	900
MGA_0416	841	cttgggtgtggttgtgagataacaatgttgaacaaagagctagcaacacacaaaccctagc	900
pQE30-MGA_0416-3	901	caatcaactaatccagctgattatgtaattagaccaactgcgattgatagtattattcca 	960
MGA_0416	901	caatcaactaatccagctgattatgtaattagaccaactgcgattgatagtattattcca	960
pQE30-MGA_0416-3	961	ttaggaatccttatctaataccgaataatttagtttatcaatcggggtgatctaaaagatcaa 	1020
MGA_0416	961	ttaggaatccttatctaataccgaataatttagtttatcaatcggggtgatctaaaagatcaa	1020
pQE30-MGA_0416-3	1021	agagttggttttaacaagaatgaactaaccaatttagttgatctatctaagatcttgaca 	1080
MGA_0416	1021	agagttggttttaacaagaatgaactaaccaatttagttgatctatctaagatcttgaca	1080
pQE30-MGA_0416-3	1081	gctgatcagtatgctaacacagaaaacgttaatacagataaccctaactaatcagattagt 	1140
MGA_0416	1081	gctgatcagtatgctaacacagaaaacgttaatacagataaccctaactaatcagattagt	1140
pQE30-MGA_0416-3	1141	aatatccgtcaaaaactggtttattctttaatttattaagtagtactcactttagaata 	1200
MGA_0416	1141	aatatccgtcaaaaactggtttattctttaatttattaagtagtactcactttagaata	1200
pQE30-MGA_0416-3	1201	attaatcaaaacgacggtgttgataacaatgcggttggtgctaatttaaccacttattat 	1260
MGA_0416	1201	attaatcaaaacgacggtgttgataacaatgcggttggtgctaatttaaccacttattat	1260
pQE30-MGA_0416-3	1261	aattggtatttttaataagttgactcaagttgaaaactattcaaagatcgatttagtggtgta	1320

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

MGA_0416	1261	 aattggtattttaataagttgactcaagttgaaaactattcaaagatcgatttagtggtta	1320
pQE30-MGA_0416-3	1321	cctaataatgatcctgagattaaggcggttaaccactccaccagctagtgtgatatgaatacc 	1380
MGA_0416	1321	cctaataatgatcctgagattaaggcggttaaccactccaccagctagtgtgatatgaatacc	1380
pQE30-MGA_0416-3	1381	gaaatgaacggttggtatggcaaataatatgactggtgatatgaatggtcaaatgaacggt 	1440
MGA_0416	1381	gaaatgaacggttggtatggcaaataatatgactggtgatatgaatggtcaaatgaacggt	1440
pQE30-MGA_0416-3	1441	gagatggctagtagcactaataacgaaatgccaaacttcaccagctactaactaa 	1494
MGA_0416	1441	gagatggctagtagcactaataacgaaatgccaaacttcaccagctactaactaa	1494

Query = MGA\_0319

Length = 1224

score = 2261 bits (1224), expect = 0.0

identities = 1224/1224 (100%), gaps = 0/1224 (0%)

strand=plus/plus

pQE30-MGA_0319-5	1	aagagaatctatgggaataactacatggatctaataccaatcccagttatttgatccta 	60
MGA_0319	1	aagagaatctatgggaataactacatggatctaataccaatcccagttatttgatccta	60
pQE30-MGA_0319-5	61	gggggaacaaaagaaagtattcttcaaaaccaattaaacaaaagattagcaacgacttt 	120
MGA_0319	61	gggggaacaaaagaaagtattcttcaaaaccaattaaacaaaagattagcaacgacttt	120
pQE30-MGA_0319-5	121	tttgactctttattttaaggattcttattttaactatgcgccaacaatgtcaactgacgct 	180
MGA_0319	121	tttgactctttattttaaggattcttattttaactatgcgccaacaatgtcaactgacgct	180
pQE30-MGA_0319-5	181	gatggttttgtcgtttatttcaaacctaactaagatcccttcaagaacgattttaagtaac 	240
MGA_0319	181	gatggttttgtcgtttatttcaaacctaactaagatcccttcaagaacgattttaagtaac	240
pQE30-MGA_0319-5	241	ccagctaactgagccaacattagatttactgatgggtgggttttagtgcaactaataacgaa 	300
MGA_0319	241	ccagctaactgagccaacattagatttactgatgggtgggttttagtgcaactaataacgaa	300
pQE30-MGA_0319-5	301	caaactaataacgaattcttagctcagatccaagcttttgcctttcaatcagttagtaact 	360
MGA_0319	301	caaactaataacgaattcttagctcagatccaagcttttgcctttcaatcagttagtaact	360
pQE30-MGA_0319-5	361	aaagaaaaccctaatttagtttcaagaattgtttttaccagcgatactcctaaagatgga 	420
MGA_0319	361	aaagaaaaccctaatttagtttcaagaattgtttttaccagcgatactcctaaagatgga	420
pQE30-MGA_0319-5	421	ctaaattcgatctttaatgatcgtagctgtagctcaacagcattagtagcaagttataat 	480
MGA_0319	421	ctaaattcgatctttaatgatcgtagctgtagctcaacagcattagtagcaagttataat	480
pQE30-MGA_0319-5	481	ttccaagcattttaaaaaatcccaaagaccaacaccttttgatgaaaacaaggggatgagagct 	540
MGA_0319	481	ttccaagcattttaaaaaatcccaaagaccaacaccttttgatgaaaacaaggggatgagagct	540
pQE30-MGA_0319-5	541	tataaccaattcattactagaggggttgacagttatgtaaatagtacaactaaagggatt 	600
MGA_0319	541	tataaccaattcattactagaggggttgacagttatgtaaatagtacaactaaagggatt	600
pQE30-MGA_0319-5	601	aatattccaacaacttatcatcagattctggtggtaagttgttaatgactgcttctgat 	660
MGA_0319	601	aatattccaacaacttatcatcagattctggtggtaagttgttaatgactgcttctgat	660

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

pQE30-MGA_0319-5	661	atgttcaatagttttgacgtatcatttagtgcagcttatgttcaacaatattttaaaca	720
MGA_0319	661	 atgttcaatagttttgacgtatcatttagtgcagcttatgttcaacaatattttaaaca	720
pQE30-MGA_0319-5	721	actaataataactaaccgcgataacgttggtgtagttcaatcagatattgatgagatcaat	780
MGA_0319	721	 actaataataactaaccgcgataacgttggtgtagttcaatcagatattgatgagatcaat	780
pQE30-MGA_0319-5	781	ctgatgaataattttattagagctaagggaatggtaacacagccaacacctactcaca	840
MGA_0319	781	 ctgatgaataattttattagagctaagggaatggtaacacagccaacacctactcaca	840
pQE30-MGA_0319-5	841	aagattactaataattcattattaaaaacaggtgaaacgaatcgaactactgatccatat	900
MGA_0319	841	 aagattactaataattcattattaaaaacaggtgaaacgaatcgaactactgatccatat	900
pQE30-MGA_0319-5	901	tacaatgcttatgcagatttagcagccgaaataaagatatccatgaaatctttgaatga	960
MGA_0319	901	 tacaatgcttatgcagatttagcagccgaaataaagatatccatgaaatctttgaatga	960
pQE30-MGA_0319-5	961	aatgggatgaagacagttgattcttcgaaatcagataattcatcaacaaacgtaatgagt	1020
MGA_0319	961	 aatgggatgaagacagttgattcttcgaaatcagataattcatcaacaaacgtaatgagt	1020
pQE30-MGA_0319-5	1021	aaaaatatgaacgctagtaagttcttattttctcgtggtcgagatgggatccatgtgatg	1080
MGA_0319	1021	 aaaaatatgaacgctagtaagttcttattttctcgtggtcgagatgggatccatgtgatg	1080
pQE30-MGA_0319-5	1081	gcgatcgatggtggggattactacttaaaccaaagtggtatgactagaaataaaca	1140
MGA_0319	1081	 gcgatcgatggtggggattactacttaaaccaaagtggtatgactagaaataaaca	1140
pQE30-MGA_0319-5	1141	caaaaagagtttctagcttaccgttcattattaagatcttcagttaacttatctgaaaat	1200
MGA_0319	1141	 caaaaagagtttctagcttaccgttcattattaagatcttcagttaacttatctgaaaat	1200
pQE30-MGA_0319-5	1201	gaaagatatgattttgatctaagc	1224
MGA_0319	1201	 gaaagatatgattttgatctaagc	1224

Query = MGA\_0843  
Length = 660  
score = 1219 bits (660), expect = 0.0  
identities = 660/660 (100%), gaps = 0/660 (0%)  
strand=plus/plus

pQE30-MGA_0843-2	1	atgaatagtagtttctttttatcgtcttatttagtaattattccaatcgggctaataagt	60
MGA_0843	1	 atgaatagtagtttctttttatcgtcttatttagtaattattccaatcgggctaataagt	60
pQE30-MGA_0843-2	61	tacgttatctataagagaaaaaatctaaagagcctggtcagttttcaggaaaaactaaa	120
MGA_0843	61	 tacgttatctataagagaaaaaatctaaagagcctggtcagttttcaggaaaaactaaa	120
pQE30-MGA_0843-2	121	gaagaacgctcgtaatgaggtgtgaaaaaccattaagcgttatctccaagataacgaaatg	180
MGA_0843	121	 gaagaacgctcgtaatgaggtgtgaaaaaccattaagcgttatctccaagataacgaaatg	180
pQE30-MGA_0843-2	181	tatgggcgtgaaattatgtattcttttggtgctaaacgaccatctccaacgatgatcgt	240
MGA_0843	181	 tatgggcgtgaaattatgtattcttttggtgctaaacgaccatctccaacgatgatcgt	240
pQE30-MGA_0843-2	241	aagttgcacaagcaatttaaagaagaacaaagcaatacctattagagcacaactatct	300
MGA_0843	241	 aagttgcacaagcaatttaaagaagaacaaagcaatacctattagagcacaactatct	300

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

pQE30-MGA_0843-2	301	aaaaaagataaaaaacaatatatttagataaatcgtaaaaaagagatggctcgtgaacgatat	360
MGA_0843	301	aaaaaagataaaaaacaatatatttagataaatcgtaaaaaagagatggctcgtgaacgatat	360
pQE30-MGA_0843-2	361	tgcatttatatttccaaaccaagatgctaggactcaagctacatttgatcctgcaattatt	420
MGA_0843	361	tgcatttatatttccaaaccaagatgctaggactcaagctacatttgatcctgcaattatt	420
pQE30-MGA_0843-2	421	gaagcagaagtattaacattaccagctaaaagtaagcgtgatgttccagaaagaaaaatc	480
MGA_0843	421	gaagcagaagtattaacattaccagctaaaagtaagcgtgatgttccagaaagaaaaatc	480
pQE30-MGA_0843-2	481	caaattaatgggttgcaagattttcaaaaagaattttcttgaattgaaccattaaagaat	540
MGA_0843	481	caaattaatgggttgcaagattttcaaaaagaattttcttgaattgaaccattaaagaat	540
pQE30-MGA_0843-2	541	aaagaagatgcgcgattgaaaaaagcagaagatgaacgcttaagaaaattagagattaaa	600
MGA_0843	541	aaagaagatgcgcgattgaaaaaagcagaagatgaacgcttaagaaaattagagattaaa	600
pQE30-MGA_0843-2	601	gagcgcagaagaaagctgctaaattagctaaaaagaagcaaaagctaaaaagaaaatttaa	660
MGA_0843	601	gagcgcagaagaaagctgctaaattagctaaaaagaagcaaaagctaaaaagaaaatttaa	660
Query = MGA_0398 Length = 1668 score = 3081 bits (1668), expect = 0.0 identities = 1668/1668 (100%), gaps = 0/1668 (0%) strand=plus/plus			
pQE30-MGA_0398-13	1	tgtggtgctgcttcttcactaacatatgaatccttctgttcagttagtagtttcagataat	60
MGA_0398	1	tgtggtgctgcttcttcactaacatatgaatccttctgttcagttagtagtttcagataat	60
pQE30-MGA_0398-13	61	agttcaacttttagctgaccaatcattctcagaaacttcatataacgggattagagatttc	120
MGA_0398	61	agttcaacttttagctgaccaatcattctcagaaacttcatataacgggattagagatttc	120
pQE30-MGA_0398-13	121	tacaaatctgtatacaaaagaaactcttccttttagctaacgacctagtttatcagttaac	180
MGA_0398	121	tacaaatctgtatacaaaagaaactcttccttttagctaacgacctagtttatcagttaac	180
pQE30-MGA_0398-13	181	aatggtgtgtgaaaacgacctggaacaattgatgattcaagaattaacacttatcgacag	240
MGA_0398	181	aatggtgtgtgaaaacgacctggaacaattgatgattcaagaattaacacttatcgacag	240
pQE30-MGA_0398-13	241	attaaaaatgatggatcatttgttgctgtagctactggtttcaaccaagaatcagcacta	300
MGA_0398	241	attaaaaatgatggatcatttgttgctgtagctactggtttcaaccaagaatcagcacta	300
pQE30-MGA_0398-13	301	aaccgaattccttgataaacacagctttatataacgaattcaaagattttggtttcatcttt	360
MGA_0398	301	aaccgaattccttgataaacacagctttatataacgaattcaaagattttggtttcatcttt	360
pQE30-MGA_0398-13	361	gtggatggtgttatcactaagaaaaatggaactaatatttcagcaataactttccaaact	420
MGA_0398	361	gtggatggtgttatcactaagaaaaatggaactaatatttcagcaataactttccaaact	420
pQE30-MGA_0398-13	421	gaaagtgcctgcatttttaacggggatcgctgctggagttttggttaaataaaaatagcaat	480
MGA_0398	421	gaaagtgcctgcatttttaacggggatcgctgctggagttttggttaaataaaaatagcaat	480
pQE30-MGA_0398-13	481	ttctttaaacaacaaaagttggtgatgtagacacttatggtattggtgcttatgtaggt	540
MGA_0398	481	ttctttaaacaacaaaagttggtgatgtagacacttatggtattggtgcttatgtaggt	540
pQE30-MGA_0398-13	541	ttacctattccatcaacaattagtttcttaaacggttttagaatgggagcgatcttcttt	600



# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

MGA_0398	541	ttacctattccatcaacaattagttttcttaaacggttttagaatgggagcgatcttcttt	600
pQE30-MGA_0398-13	601	aatcaatacatccaaccaagagtagaaggatttaagaaacttttcttgagtaagctcta	660
MGA_0398	601	aatcaatacatccaaccaagagtagaaggatttaagaaacttttcttgagtaagctcta	660
pQE30-MGA_0398-13	661	gcagcagataacaatggtaatcgtgatactttaccagctgataaatctgattcatttaac	720
MGA_0398	661	gcagcagataacaatggtaatcgtgatactttaccagctgataaatctgattcatttaac	720
pQE30-MGA_0398-13	721	gcgacagaacaaagagcaacaacaattactaaccaattgttacaaaatgggtgcatcttta	780
MGA_0398	721	gcgacagaacaaagagcaacaacaattactaaccaattgttacaaaatgggtgcatcttta	780
pQE30-MGA_0398-13	781	atctatccgattgctgggtccacaaacagctcttagtcaaaacgtgattctttctaataga	840
MGA_0398	781	atctatccgattgctgggtccacaaacagctcttagtcaaaacgtgattctttctaataga	840
pQE30-MGA_0398-13	841	aacacacaccacgcacaattaattgggtgttgatacagctcaagaaaatttagcaacaaca	900
MGA_0398	841	aacacacaccacgcacaattaattgggtgttgatacagctcaagaaaatttagcaacaaca	900
pQE30-MGA_0398-13	901	caaccattacaaggagctcctgggtggaagacaatcgctttctcaacagttaaagcttta	960
MGA_0398	901	caaccattacaaggagctcctgggtggaagacaatcgctttctcaacagttaaagcttta	960
pQE30-MGA_0398-13	961	gatgttgctgttgattcaacttttaaaagcaattcaaaaagggtcaaccagtaaacggcttc	1020
MGA_0398	961	gatgttgctgttgattcaacttttaaaagcaattcaaaaagggtcaaccagtaaacggcttc	1020
pQE30-MGA_0398-13	1021	tatggatatggatgaaataacttagctacttttagctaacgggtggtgtttcattatcaagc	1080
MGA_0398	1021	tatggatatggatgaaataacttagctacttttagctaacgggtggtgtttcattatcaagc	1080
pQE30-MGA_0398-13	1081	gcaggattagcttacttaccaaaacttaactgatctgtttactaaaactagcatacctgct	1140
MGA_0398	1081	gcaggattagcttacttaccaaaacttaactgatctgtttactaaaactagcatacctgct	1140
pQE30-MGA_0398-13	1141	aatgcagatcaaactactccaagacaagaaatgtctgatcaatcatctggaactccagcg	1200
MGA_0398	1141	aatgcagatcaaactactccaagacaagaaatgtctgatcaatcatctggaactccagcg	1200
pQE30-MGA_0398-13	1201	atgactggaactatgaacgggtcaagctaattgggtggaactgaatctatgatggctatgagt	1260
MGA_0398	1201	atgactggaactatgaacgggtcaagctaattgggtggaactgaatctatgatggctatgagt	1260
pQE30-MGA_0398-13	1261	atgatgatggcagaagctggggcttctacagcactgaatatgccagcaacatcttcaaca	1320
MGA_0398	1261	atgatgatggcagaagctggggcttctacagcactgaatatgccagcaacatcttcaaca	1320
pQE30-MGA_0398-13	1321	ccagctgctacaacaactaacaacaaatttctgctaacgacttagttacaacaacagaa	1380
MGA_0398	1321	ccagctgctacaacaactaacaacaaatttctgctaacgacttagttacaacaacagaa	1380
pQE30-MGA_0398-13	1381	gctaataaaaactatggtaattatgcagtacgttaacatttttagctggaacttctacttta	1440
MGA_0398	1381	gctaataaaaactatggtaattatgcagtacgttaacatttttagctggaacttctacttta	1440
pQE30-MGA_0398-13	1441	ctaccagctgctagtagagaaaactgaaaaattaaagggtgatgaattaagaatgttcaaa	1500
MGA_0398	1441	ctaccagctgctagtagagaaaactgaaaaattaaagggtgatgaattaagaatgttcaaa	1500
pQE30-MGA_0398-13	1501	acttctgtggaagctgatgctaaattattaccaattttaacaacagacgttgatccta	1560
MGA_0398	1501	acttctgtggaagctgatgctaaattattaccaattttaacaacagacgttgatccta	1560
pQE30-MGA_0398-13	1561	acagctttaagtcataagctctgcaatcatttctgggttcattccaacaagctgaagggtacc	1620

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

MGA_0398	1561	 acagctttaagtcataagtcgtcaatcatttctggttcattccaacaagctgaaggtacc	1620
pQE30-MGA_0398-13	1621	ttattaagccaagcaggcacaacaatttgtcttcaagaaattaaactaa	1668
MGA_0398	1621	 ttattaagccaagcaggcacaacaatttgtcttcaagaaattaaactaa	1668

Query = *crmA*  
Length = 2715  
score = 1808 bits (979), expect = 0.0  
identities = 979/979 (100%), gaps = 0/979 (0%)  
strand=plus/plus

pQE30-crmA-16	1	aagcaatcagataagagtaacgataaacacgcaattagttaatcaagcaagaacgctagat	60
<i>crmA</i>	1	 aagcaatcagataagagtaacgataaacacgcaattagttaatcaagcaagaacgctagat	60
pQE30-crmA-16	61	gctaattctgttagacttgcaggctcttgacaaaatgggttcgttgcataacagttctt	120
<i>crmA</i>	61	 gctaattctgttagacttgcaggctcttgacaaaatgggttcgttgcataacagttctt	120
pQE30-crmA-16	121	agagatggttgatgataactttataacagcagctaataaggaacaattatcaaattagatagt	180
<i>crmA</i>	121	 agagatggttgatgataactttataacagcagctaataaggaacaattatcaaattagatagt	180
pQE30-crmA-16	181	tttactaaaccattatatggttttagatctaagtgatgattttgctggatacaaaagtaaaa	240
<i>crmA</i>	181	 tttactaaaccattatatggttttagatctaagtgatgattttgctggatacaaaagtaaaa	240
pQE30-crmA-16	241	caaatagtttcagattacacaactagcagaaatagatttgatcaaagacaaacaagagca	300
<i>crmA</i>	241	 caaatagtttcagattacacaactagcagaaatagatttgatcaaagacaaacaagagca	300
pQE30-crmA-16	301	tattatgctctgttggttaatgatgaagctaacgcttcattttaaagaattaataactaac	360
<i>crmA</i>	301	 tattatgctctgttggttaatgatgaagctaacgcttcattttaaagaattaataactaac	360
pQE30-crmA-16	361	tcaaatagaattggtaatagaacaacaattctaagtttgtaattgggtggtgttgataat	420
<i>crmA</i>	361	 tcaaatagaattggtaatagaacaacaattctaagtttgtaattgggtggtgttgataat	420
pQE30-crmA-16	421	ccagctcacgtaattagatttactgatgatgggactaaatttaattttacaaagcaaact	480
<i>crmA</i>	421	 ccagctcacgtaattagatttactgatgatgggactaaatttaattttacaaagcaaact	480
pQE30-crmA-16	481	caaggtgaaattgttaatgacttcatttttagatgcgccaatcttacctaaagatttacac	540
<i>crmA</i>	481	 caaggtgaaattgttaatgacttcatttttagatgcgccaatcttacctaaagatttacac	540
pQE30-crmA-16	541	ccagattgatataaacttatacattcaaagaaagatcttaccaaagacgtaaacactgca	600
<i>crmA</i>	541	 ccagattgatataaacttatacattcaaagaaagatcttaccaaagacgtaaacactgca	600
pQE30-crmA-16	601	gttggttccttgaccagtaggttagagttagtggacaaatgctgatgatgggatgtttgat	660
<i>crmA</i>	601	 gttggttccttgaccagtaggttagagttagtggacaaatgctgatgatgggatgtttgat	660
pQE30-crmA-16	661	tttggaatgggtcaaataactaatacagatcctattgctcaaactaaaaccactactgat	720
<i>crmA</i>	661	 tttggaatgggtcaaataactaatacagatcctattgctcaaactaaaaccactactgat	720
pQE30-crmA-16	721	aatcaaaatccttcaacttttaattcaggagcaatgcctgggtgcaaacaatagatacgat	780
<i>crmA</i>	721	 aatcaaaatccttcaacttttaattcaggagcaatgcctgggtgcaaacaatagatacgat	780
pQE30-crmA-16	781	tctcaattgaatgtcaagcatagaattaaaacatctttccaattagatgaaaaatttggt	840
<i>crmA</i>	781	 tctcaattgaatgtcaagcatagaattaaaacatctttccaattagatgaaaaatttggt	840

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

pQE30-crmA-16	841	tatccagaatgaactggttctgaagagaataaaaaatattacaagattagctactggaagt	900
crmA	841		900
pQE30-crmA-16	901	ttgccaagcaacgaaagatattgaattcttgacataccagggactccacaagttacttta	960
crmA	901		960
pQE30-crmA-16	961	aaagaagattcagttaacg	979
crmA	961		979
score = 2134 bits (1155), expect = 0.0 identities = 1155/1155 (100%), gaps = 0/1155 (0%) strand=plus/plus			
pQE30-crmA-16	979	ggataccaacaaactagAACAGGAacttttctgatacggAACATATAaactatttaaacaac	1038
crmA	1561		1620
pQE30-crmA-16	1039	agcccttacgacgtatttagatgctgcaagagtaggtactgaaaccaatcaatttagaaga	1098
crmA	1621		1680
pQE30-crmA-16	1099	acttcattaacataccctgttatgggtggatatctaactgaagaaggtgctagaagtttc	1158
crmA	1681		1740
pQE30-crmA-16	1159	tctaatactccatatataagagcacaaggtgacacaccagaaagccgaagcatcttccaa	1218
crmA	1741		1800
pQE30-crmA-16	1219	tctggctattctgataatacttatgagtacattcaatcagttttaggatttgatggaatt	1278
crmA	1801		1860
pQE30-crmA-16	1279	agaaataacttaaagtgttggtggttaaagcatcaagcttcttaaactcaaatagaccaa	1338
crmA	1861		1920
pQE30-crmA-16	1339	ccaaacggtctagaaatgattgctgcaacaacatacttaagatcacaaattggattagct	1398
crmA	1921		1980
pQE30-crmA-16	1399	agaacatctggattaccaaaccaacaaccattcggacaactcaccaagttatttcagta	1458
crmA	1981		2040
pQE30-crmA-16	1459	tcacctggtgatcagttctcatcaattaagaatattagaacaatcttcctggttaaccag	1518
crmA	2041		2100
pQE30-crmA-16	1519	ttatgatacttcttattcacaaatgaaaataataaatctagtgtttatacattaagatta	1578
crmA	2101		2160
pQE30-crmA-16	1579	gctgactcaagtaaccctgatgcgtcaagctcattcagttccaacaagtttaattgacgtt	1638
crmA	2161		2220
pQE30-crmA-16	1639	aatgaaattggtgtaattcttacctttattagacaattcattctatacagtaaatgctgct	1698
crmA	2221		2280
pQE30-crmA-16	1699	ggtaatggtgattgttctcatcaaaccctggttctcctggatcatatactgctgtaaat	1758
crmA	2281		2340

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

pQE30-crmA-16	1759	acattttaatcagaacttatctgatattgcttttgaaggttctggtgctaaatatacatct	1818
crmA	2341		
		acattttaatcagaacttatctgatattgcttttgaaggttctggtgctaaatatacatct	2400
pQE30-crmA-16	1819	gatttctgaggaacaatccaattcaaaccgatgagtacttaattcaaatgggttcact	1878
crmA	2401		
		gatttctgaggaacaatccaattcaaaccgatgagtacttaattcaaatgggttcact	2460
pQE30-crmA-16	1879	agtcaagtggctagaaacttcgttacaaccaaagcttcttaaacagtttagttgacttc	1938
crmA	2461		
		agtcaagtggctagaaacttcgttacaaccaaagcttcttaaacagtttagttgacttc	2520
pQE30-crmA-16	1939	actcctgctaatactgctggtactaactaccgtgtagtggttgatcctgatggtaatttaaca	1998
crmA	2521		
		actcctgctaatactgctggtactaactaccgtgtagtggttgatcctgatggtaatttaaca	2580
pQE30-crmA-16	1999	aacaaaaacctacctctaaaagttcagatccaatacttagatggtaagtattatgatgct	2058
crmA	2581		
		aacaaaaacctacctctaaaagttcagatccaatacttagatggtaagtattatgatgct	2640
pQE30-crmA-16	2059	aaattaaagaacaataatttagtaacattctcttataacaactttgctgctttaccttca	2118
crmA	2641		
		aaattaaagaacaataatttagtaacattctcttataacaactttgctgctttaccttca	2700
pQE30-crmA-16	2119	tgagtagtgcctaca	2133
crmA	2701		
		tgagtagtgcctaca	2715

Query = crmC

Length = 1965

score = 3629 bits (1965), expect = 0.0

identities = 1965/1965 (100%), gaps = 0/1965 (0%)

strand=plus/plus

pQE30-crmC-17	1	aatcattcttctagatcagataaattagtttaaccaagcacagttagttgctgataataat	60
crmC	1		
		aatcattcttctagatcagataaattagtttaaccaagcacagttagttgctgataataat	60
pQE30-crmC-17	61	caggattatcagcaactaaaaaccgaacaaaatcattatgtaactaatagtaataacgat	120
crmC	61		
		caggattatcagcaactaaaaaccgaacaaaatcattatgtaactaatagtaataacgat	120
pQE30-crmC-17	121	tattttgttctcaccaaaactgggtgttgtaaaacttgatgtttttggcaatgtttattat	180
crmC	121		
		tattttgttctcaccaaaactgggtgttgtaaaacttgatgtttttggcaatgtttattat	180
pQE30-crmC-17	181	aactacagttttgcgaccaacttcattaattatcaaaccggttgattttggtgctgatagt	240
crmC	181		
		aactacagttttgcgaccaacttcattaattatcaaaccggttgattttggtgctgatagt	240
pQE30-crmC-17	241	ttaagagctgattattattacttattgatcaaaaaccgattgttagtattaatgggaac	300
crmC	241		
		ttaagagctgattattattacttattgatcaaaaaccgattgttagtattaatgggaac	300
pQE30-crmC-17	301	gacttaaaactcattaagtgtcagttcaccgcgttatgtgttgcaactaaaagatgagaat	360
crmC	301		
		gacttaaaactcattaagtgtcagttcaccgcgttatgtgttgcaactaaaagatgagaat	360
pQE30-crmC-17	361	aaaactttaaatgtgcttaagacggttccaattaagttccttatcggttctcaagaaatatt	420
crmC	361		
		aaaactttaaatgtgcttaagacggttccaattaagttccttatcggttctcaagaaatatt	420

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

pQE30-crmC-17	421	gcgaccttataacaatcggtattgcgtttgaatcaacttcaagtagttcaaacctatcaggg	480
<i>crmC</i>	421	gcgaccttataacaatcggtattgcgtttgaatcaacttcaagtagttcaaacctatcaggg	480
pQE30-crmC-17	481	ctattagataaaccaactaatagcacgttattaaataatatttcagcatattatgatcat	540
<i>crmC</i>	481	ctattagataaaccaactaatagcacgttattaaataatatttcagcatattatgatcat	540
pQE30-crmC-17	541	gcgattagtaaattagtggtttattaataatcgcttaggggtgttcggtagtaataatgatcaa	600
<i>crmC</i>	541	gcgattagtaaattagtggtttattaataatcgcttaggggtgttcggtagtaataatgatcaa	600
pQE30-crmC-17	601	ttatccttatggttttatctatttaataattaatcaaaccaacgacttttagtacgaataac	660
<i>crmC</i>	601	ttatccttatggttttatctatttaataattaatcaaaccaacgacttttagtacgaataac	660
pQE30-crmC-17	661	gacttaaatttggtcacaatctttccccaattatatgggacaaccaacctaattgagcg	720
<i>crmC</i>	661	gacttaaatttggtcacaatctttccccaattatatgggacaaccaacctaattgagcg	720
pQE30-crmC-17	721	ttaaataatattagtaaatgttaataagcaactacagcacagtttgtagtaataacaacatt	780
<i>crmC</i>	721	ttaaataatattagtaaatgttaataagcaactacagcacagtttgtagtaataacaacatt	780
pQE30-crmC-17	781	cgtttatacaccccagcttacttcgtaattggggcaaatggttgtagcaaaaatttagtt	840
<i>crmC</i>	781	cgtttatacaccccagcttacttcgtaattggggcaaatggttgtagcaaaaatttagtt	840
pQE30-crmC-17	841	aatcctaacggattattaaacgccaacaaactgcgatctattcattacatttaggggta	900
<i>crmC</i>	841	aatcctaacggattattaaacgccaacaaactgcgatctattcattacatttaggggta	900
pQE30-crmC-17	901	gtgccagcacctattagtgcgattactaataacaatcaattcggttataagattagtgct	960
<i>crmC</i>	901	gtgccagcacctattagtgcgattactaataacaatcaattcggttataagattagtgct	960
pQE30-crmC-17	961	aactactacaatagtgataatcttaattattccaacctatttgtagtggggtagatcagt	1020
<i>crmC</i>	961	aactactacaatagtgataatcttaattattccaacctatttgtagtggggtagatcagt	1020
pQE30-crmC-17	1021	caaactaaagttgccagatcactaataatgttaatactttttaatgagttccttagtgct	1080
<i>crmC</i>	1021	caaactaaagttgccagatcactaataatgttaatactttttaatgagttccttagtgct	1080
pQE30-crmC-17	1081	tctcacacaacaagacttaattctatctaagattaattctaagaccgaaccgcttatcac	1140
<i>crmC</i>	1081	tctcacacaacaagacttaattctatctaagattaattctaagaccgaaccgcttatcac	1140
pQE30-crmC-17	1141	tattcagttcaagaagatcggttttagtcaaacaacaatagttatcccaacctagttggg	1200
<i>crmC</i>	1141	tattcagttcaagaagatcggttttagtcaaacaacaatagttatcccaacctagttggg	1200
pQE30-crmC-17	1201	atcttaatcattaaagatgagatggtggggtttttataactcagctgatagttataatctg	1260
<i>crmC</i>	1201	atcttaatcattaaagatgagatggtggggtttttataactcagctgatagttataatctg	1260
pQE30-crmC-17	1261	ctctcacagattaactattatctgattagtgatgatatcagtatggctgggtttgatcta	1320
<i>crmC</i>	1261	ctctcacagattaactattatctgattagtgatgatatcagtatggctgggtttgatcta	1320
pQE30-crmC-17	1321	ttaacaactatcgactaaaagatgggaaatgatacctaagtttctttgatccaagacaa	1380
<i>crmC</i>	1321	ttaacaactatcgactaaaagatgggaaatgatacctaagtttctttgatccaagacaa	1380
pQE30-crmC-17	1381	tcaattagttggactaatcagattagtgccaacaattagtgcaacttaataatgggacatta	1440
<i>crmC</i>	1381	tcaattagttggactaatcagattagtgccaacaattagtgcaacttaataatgggacatta	1440

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

pQE30-crmC-17	1441	gcaattaactataactaaactgtatcaagttaaccaacaaggatcgatcttcttaacccta	1500
<i>crmC</i>	1441		1500
pQE30-crmC-17	1501	attttgggtgagcaagctttttatacttttagtaactcaaaataatagagtagatctcaac	1560
<i>crmC</i>	1501	attttgggtgagcaagctttttatacttttagtaactcaaaataatagagtagatctcaac	1560
pQE30-crmC-17	1561	ctagtcttttatgatcagaaaaacccaaagtttttctgatgccaccttctttaataacggt	1620
<i>crmC</i>	1561	ctagtcttttatgatcagaaaaacccaaagtttttctgatgccaccttctttaataacggt	1620
pQE30-crmC-17	1621	gtgcctagttttgataatatcccaaactcagatgggtcaaaattatgggggatgattagt	1680
<i>crmC</i>	1621	gtgcctagttttgataatatcccaaactcagatgggtcaaaattatgggggatgattagt	1680
pQE30-crmC-17	1681	tttaaaccatcttcatattttagttaatcaagatctaaacacaactagtgcattgagatta	1740
<i>crmC</i>	1681	tttaaaccatcttcatattttagttaatcaagatctaaacacaactagtgcattgagatta	1740
pQE30-crmC-17	1741	agcaataataatatttattgcttaataaacttatatgattacactcctgggtgagatcgattga	1800
<i>crmC</i>	1741	agcaataataatatttattgcttaataaacttatatgattacactcctgggtgagatcgattga	1800
pQE30-crmC-17	1801	aaacctagagtagaagctattgcactaaacgattcaaccttaagattatcaatccaattt	1860
<i>crmC</i>	1801	aaacctagagtagaagctattgcactaaacgattcaaccttaagattatcaatccaattt	1860
pQE30-crmC-17	1861	ccttattttgatgggaactattacagtcccaacccaaaagtggtaaacttagcattccca	1920
<i>crmC</i>	1861	ccttattttgatgggaactattacagtcccaacccaaaagtggtaaacttagcattccca	1920
pQE30-crmC-17	1921	acaattatctatgataatcttaatgcttatcccgtttatggtgtg	1965
<i>crmC</i>	1921	acaattatctatgataatcttaatgcttatcccgtttatggtgtg	1965

Query = *mgc2*

Length = 777

score = 1435 bits (777), expect = 0.0

identities = 777/777 (100%), gaps = 0/777 (0%)

strand=plus/plus

pQE30-mgc2-32	1	gaacattcacccacacttggtttacatgaaggtgaaactaactccgttggtccaagaaag	60
<i>mgc2</i>	1		60
pQE30-mgc2-32	61	attacctccgaaccctgattttatccagtagtgggtgcaggtgctgggtgattggtgtt	120
<i>mgc2</i>	61	attacctccgaaccctgattttatccagtagtgggtgcaggtgctgggtgattggtgtt	120
pQE30-mgc2-32	121	tctttactcttgggttttagggattgggattccgatcgctaagaaaaaagaaagaatgatg	180
<i>mgc2</i>	121	tctttactcttgggttttagggattgggattccgatcgctaagaaaaaagaaagaatgatg	180
pQE30-mgc2-32	181	atccaagaacgtgaagaacacccaaaagatggttgaatcccttgggaataattgaagaacaa	240
<i>mgc2</i>	181	atccaagaacgtgaagaacacccaaaagatggttgaatcccttgggaataattgaagaacaa	240
pQE30-mgc2-32	241	aataaaacagaagcgattgagccaactgaagaagttaataactcaagaaccaactcaacca	300
<i>mgc2</i>	241		300
pQE30-mgc2-32	301	gctgggtgtaattgtagctaataaccctcagatggggatcaatcaacctcagattaatccg	360
<i>mgc2</i>	301	gctgggtgtaattgtagctaataaccctcagatggggatcaatcaacctcagattaatccg	360

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS

pQE30-mgc2-32	361	caatttggtcctaataccccacaagaattaacccacagtgctttggtggcccaatgcaa	420
<i>mgc2</i>	361	caatttggtcctaataccccacaagaattaacccacagtgctttggtggcccaatgcaa	420
pQE30-mgc2-32	421	cctaaccaaatgggaatgcgaccaggggtttaaccaaagccccacaaatgggaggaatg	480
<i>mgc2</i>	421	cctaaccaaatgggaatgcgaccaggggtttaaccaaagccccacaaatgggaggaatg	480
pQE30-mgc2-32	481	ccacctaaccaaatgggaatgcgaccaggggtttaaccaaagccccacaaatgggagga	540
<i>mgc2</i>	481	ccacctaaccaaatgggaatgcgaccaggggtttaaccaaagccccacaaatgggagga	540
pQE30-mgc2-32	541	atgccaccaagaccaaacttccctaaccaaatgcctaatatgaaccaacctagaccaggt	600
<i>mgc2</i>	541	atgccaccaagaccaaacttccctaaccaaatgcctaatatgaaccaacctagaccaggt	600
pQE30-mgc2-32	601	ttcagaccacaacctggtggtgggtgccgatgggaaataaagctggaggtgggtttaat	660
<i>mgc2</i>	601	ttcagaccacaacctggtggtgggtgccgatgggaaataaagctggaggtgggtttaat	660
pQE30-mgc2-32	661	caccaggtacaccaatgggtccaaaccgatgaacttccctaatacaggaatgaatcag	720
<i>mgc2</i>	661	caccaggtacaccaatgggtccaaaccgatgaacttccctaatacaggaatgaatcag	720
pQE30-mgc2-32	721	ccccacacatggcaggaccaagagctggttttccaccacaaaatggacctagataa	777
<i>mgc2</i>	721	ccccacacatggcaggaccaagagctggttttccaccacaaaatggacctagataa	777

Query = *oppA*  
Length = 843  
score = 1557 bits (843), expect = 0.0  
identities = 843/843 (100%), gaps = 0/843 (0%)  
strand=plus/plus

pQE30- <i>oppA</i>	1	acagcttatggaaccacaaagcttatgatggtcgaaatctagaactatTTTTTgaaggc	60
<i>oppA</i>	1	acagcttatggaaccacaaagcttatgatggtcgaaatctagaactatTTTTTgaaggc	60
pQE30- <i>oppA</i>	61	caaaaatctaaaactaagaacaataaagaatttgatctactaaactatgactttgtgtt	120
<i>oppA</i>	61	caaaaatctaaaactaagaacaataaagaatttgatctactaaactatgactttgtgtt	120
pQE30- <i>oppA</i>	121	cacttatctaaaagatttaactttgaaaagaccgttcgtcatgatctagctcatgattta	180
<i>oppA</i>	121	cacttatctaaaagatttaactttgaaaagaccgttcgtcatgatctagctcatgattta	180
pQE30- <i>oppA</i>	181	aaaacagcaagattttatctagaccgttttaagctaaacacccagaactaaaacaagta	240
<i>oppA</i>	181	aaaacagcaagattttatctagaccgttttaagctaaacacccagaactaaaacaagta	240
pQE30- <i>oppA</i>	241	accctaaacttcttagataatagtagtgacgaacaagttcgtggggcgacttatctaaaa	300
<i>oppA</i>	241	accctaaacttcttagataatagtagtgacgaacaagttcgtggggcgacttatctaaaa	300
pQE30- <i>oppA</i>	301	caagcgatgctagaagcggttggtgggtatgtcaatattgaagaaaaatcattaccagaa	360
<i>oppA</i>	301	caagcgatgctagaagcggttggtgggtatgtcaatattgaagaaaaatcattaccagaa	360
pQE30- <i>oppA</i>	361	aatatctttgccagctacattgaaaaagggtcagtagatcatcttatcaaaactacgac	420
<i>oppA</i>	361	aatatctttgccagctacattgaaaaagggtcagtagatcatcttatcaaaactacgac	420
pQE30- <i>oppA</i>	421	cgttttagtggtggccaaccacaagatgggggtactaaccttctttaaaaacgatgggatt	480
<i>oppA</i>	421	cgttttagtggtggccaaccacaagatgggggtactaaccttctttaaaaacgatgggatt	480

# Anexo 7. COMPARACIÓN DE LAS SECUENCIAS DE LOS INSERTOS SECUENCIADOS CON LAS SECUENCIAS AMPLIFICADAS|

pQE30-oppA	481	gatagtttcagacaaaaaaatatacgggttcaaagaaaaccagtcggatcatacacatat	540
oppA	481	gatagtttcagacaaaaaaatatacgggttcaaagaaaaccagtcggatcatacacatat	540
pQE30-oppA	541	ctagattacttatctaatttagtactagaagcactccaaaaaactgatccaaatcttaca	600
oppA	541	ctagattacttatctaatttagtactagaagcactccaaaaaactgatccaaatcttaca	600
pQE30-oppA	601	agacaaacgatcttaaaacctgatattgatcggtattatgatgatcgatcaatagcgatcca	660
oppA	601	agacaaacgatcttaaaacctgatattgatcggtattatgatgatcgatcaatagcgatcca	660
pQE30-oppA	661	gctatgctagcggcttataataatgctgttaactcattaaacgcatttaacttaaatagc	720
oppA	661	gctatgctagcggcttataataatgctgttaactcattaaacgcatttaacttaaatagc	720
pQE30-oppA	721	tttagtacagatcacttagctgcaatcatcgctagattaaataaagttaaccagggttt	780
oppA	721	tttagtacagatcacttagctgcaatcatcgctagattaaataaagttaaccagggttt	780
pQE30-oppA	781	gaaaacacccgatactctaataatggttaatgcattattagattacatgatcattagt	840
oppA	781	gaaaacacccgatactctaataatggttaatgcattattagattacatgatcattagt	840
pQE30-oppA	841	aat 843	
oppA	841	aat 843	



# **Capítulo 1**

## **REVISIÓN BIBLIOGRÁFICA**

## **Capítulo 2**

# **PREDICCIÓN DE ANTÍGENOS DE SUPERFICIE DE *M. gallisepticum* CON EL EMPLEO DE Bact-VAP**

## **Capítulo 3**

# **OBTENCIÓN Y CARACTERIZACIÓN TOPOLÓGICA Y ANTIGÉNICA DE LAS PROTEÍNAS OBTENIDAS**

## **Capítulo 4**

### **DISCUSIÓN GENERAL**

## AGRADECIMIENTOS

- A mi tutora la Dra. Siomara Martínez por la ayuda brindada en la realización de esta tesis
- A Maricela por darme una mano cada vez que lo necesité.
- En general a mis compañeros de grupo...simplemente por soportarme.
- A Abdulahi, Jorge, Made y Alejo que con sus consejos y señalamientos, hicieron de este trabajo, un documento un millón de veces mejor.
- A Ivette y Belkis por sus inapreciables revisiones.
- A Isel por su invaluable y molesto tiempo de impresora impresora y a Yanet por sus cartuchos.
- A Uffito y Dulce por el apoyo incondicional en los momentos difíciles.
- Al IFS, CONICET y ELAP por las becas que permitieron la realización de este trabajo.
- A los doctores Fernán Agüero y José Pérez-Casal por la ayuda que me brindaron tanto técnica como materialmente, así como por sus sabios consejos.
- Sin dudas un agradecimiento especial para el Dr. Andrew Potter (Andy), por su gran apoyo de todo tipo.
- A Santi, Tracy, Teresa y Musa por hacerme la vida más fácil lejos de mi casa y los míos y por toda la ayuda que alguna vez me brindaron.
- A Mario Pablo que vino a salvar un “imposible” en este trabajo.
- A Adrián, que desde que era un estudiante está trabajando para esta tesis y no ha dejado de hacerlo aunque ya no pertenezca al laboratorio.
- A Yandy, quien a pesar de las presiones que ha tenido sobre él, ha sabido sacar fuerzas para hacer lo que hiciese falta para que las cosas salieran.
- A Yaday, sin dudas, por estar ahí incondicionalmente, para hacer lo que hiciera falta, desde una transformación, hasta cargar una impresora por todo el CENSA o buscarme unos plátanos para mis niños si yo no tenía tiempo.
- A todos los que en las múltiples predefensas estuvieron ahí y me hicieron señalamientos que de mucho me sirvieron.

- A la Nena, Sunia y María que muchas veces tuvieron que asumir “tareas extras” para que yo me pudiera dedicar a esta tesis.
- A mi suegra que se pasó todo el tiempo dándome ánimos.
- A mis padres, también a mi papá que ya no está y a quien le hubiera encantado ver este momento,, a ellos mis agradecimientos por sufrir o disfrutar a la par mía (incluso más) cada uno de mis fracasos o éxitos, por darme lo mejor que pudieron darme, por poner sus corazones siempre a mi lado. Gracias a ellos soy quien soy, por lo tanto gracias a ellos hoy aquí me estoy haciendo Doctor y entonces, ¿por qué no?, esta es también su Doctorado.
- A Ivette, sin dudas mucho amor hay que sentir para apoyarme como lo hizo, sobre todo porque juro que normalmente no es sencillo soportarme, imagínense en medio de una Doctorado
- Y a mis niños que han sufrido las carencias de un padre poco dedicado en estos tiempos y que han sido mi mejor inspiración

A mis niños, Ivette y mis padres