BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

INSTITUTO DE CIENCIAS

ESTUDIO *IN SILICO* DE LA TRANSCRIPCIÓN DE GENES QUE CODIFICAN LAS CONDROITIN LIASAS EN *Avibacterium paragallinarum*

Tesis que para obtener el grado de:

Maestría en Microbiología

Con opción en:

Bioquímica y Genética Microbiana

PRESENTA:

HÉCTOR HUGO MORENO CASTILLO

DIRECTOR: Dr. CANDELARIO VAZQUEZ CRUZ

CO-DIRECTOR: Dr. ERASMO NEGRETE ABASCAL

Puebla, Pue. DIC 2022.

"*The future belongs to those who believe in the beauty of their dreams*"

- Eleanor Roosevelt.

Agradecimientos:

A mis padres: Por todo su esfuerzo, dedicación y sabiduría para formar a un hijo, individuo y profesionista dedicado a su vocación y que ejercerá con los mismos principios como lealtad, honestidad, sentido del deber entre otros, que lo llevará a ser un ciudadano y profesionista íntegro; Por todas aquellas conversaciones y palabras que poco a poco fueron forjando a alguien que se superaría en sus días por venir.

A mi hermana: Quien siempre me alentó a llegar lejos en la vida; A quien veo como un modelo a seguir debido a su fortaleza para alcanzar sus metas y objetivos; Por siempre hacerme ver mis errores de una manera directa e impulsarme a superarlos.

A Brenda Elisa Márquez Barrientos: En un modo muy especial, por aquellas palabras de aliento y amor en los días difíciles, porque cuando todo parecía ir cuesta abajo, me dió las fuerzas para no rendirme y superar muchos obstaculos; Por nuestro vínculo que nos inspira a ser mejores y cumplir nuestros sueños.

A mi abuela Lourdes: Gracias por brindar algo muy importante, el deseo de superarse a si mismo, de buscar el conocimiento y el gusto por aprender siempre todo aquello que ignoramos, porque siempre me inspiraste a que el conocimiento por más duro que sea obtenerlo, siempre será un tesoro y por darle a mi padre todas esas enseñanzas que fueron pasadas y que seguirán siendo otorgadas de generación en generación.

A Elena Cobos y Miguel Ávalos: Por otorgarme su gran amistad en mi paso por el laboratorio MIMOCE-ICUAP, por ser mis mentores y enseñarme mucho de la disciplina académica e investigación.

Al Doctor Candelario Vazquez Cruz: Por darme la oportunidad de cursar este posgrado bajo su tutela y guía; además de la oportunidad de desempeñarme como un estudiante digno de su confianza y del mismo modo, por ser paciente y siempre enseñarme todo aquello que desconocía.

A mi comité revisor:

- Dra. Patricia Sánchez Alonso
- Dra. Lucía Soto Urzúa
- Dra. Alejandra Espinoza Texis
- Dra. Norma Elena Rojas Ruiz
- Dr. Erasmo Negrete Abascal
- Dr. Miguel Ángel Avalos Rangel

Agradezco a todos y cada uno de ustedes sus valiosas observaciones en la redacción del trabajo, tomaré cada uno de estos puntos en cuenta en mi formación profesional.

Proyecto apoyado por UNAM-DGAPA-PAPIIT IN204122.

A CONACYT por beca Hugo Moreno Castillo No. 1048779.

Al proyecto CONACYT CB-2015-01-259209

Al proyecto BUAP 100103133-VIEP 2022

Se agradece a la Vicerrectoría de Investigación y Estudios de Posgrado por el apoyo otorgado para la conclusión de esta tesis del Eje IV. Modelo de Investigación abierta y compartida. Objetivo 13. Formar recuersos humanos que impacten positivamente el contexto social y científico como consecuencia de su accionar en una comunidad para lograr una educación desarrolladora de la transformación. Indicador establecido en el Plan de Desarrollo Institucional 2021-2025.

BUAP

Puebla, Pue. a 6 de diciembre 2022.

A LA ACADEMIA DEL POSGRADO
EN MICROBIOLOGÍA
CICM-ICUAP
PRESENTE

Por este conducto los abajo firmantes, integrantes del Comité revisor de Tesis de Maestría del alumno **Hèctor Hugo Moreno Castillo,** les informamos que hemos revisado el escrito titulado:

"Estudio *in silico* de la transcripción de los genes que codifican las condroitin liasas en *Avibacterium pargallinarum*"

A nuestro juicio, el alumno Hèctor Hugo Moreno Castillo puede proceder a la digitalización de la tesis y a la presentación del examen de grado correspondiente.

Sin más que agregar, nos despedimos afectuosamente de ustedes.

Atentamente.
"Pensar Bien Para Vivir Mejor"

M.C. Alejandra Espinosa Texis — **D.C. María Patricia G. Sánchez Alonso**

D.C. Lucia Soto Urzúa — **D.C. Norma Elena Rojas Ruiz**

D.C. Miguel Angel Avalos Rangel

Posgrado en Ciencias Microbiológicas
Instituto de Ciencias (ICUAP)

Edificio IC11,
Ciudad Universitaria
Col. San Manuel, Puebla, Pue. C.P. 72570
01 (222) 229 55 00 Ext. 2522
posgrado.microbiologia@correo.buap.mx

INDICE GENERAL

INDICE GENERAL **6**
ÍNDICE DE IMAGENES **8**
ÍNDICE DE TABLAS **9**
ÍNDICE DE GRÁFICAS **10**
RESUMEN **11**
ABSTRACT **12**
1. INTRODUCCIÓN **13**
2. MARCO TEÓRICO **13**
2.1 Coriza Infecciosa 13
2.2 *Avibacterium paragallinarum* 14
2.2.1 Epidemiología *14*
2.2.2 Clasificación de la bacteria A. paragallinarum *14*
2.2.3 Requerimientos de crecimiento *15*
2.2.4 Propiedades bioquímicas *15*
2.2.5 Genómica *16*
2.2.6 Serotipificación *17*
2.2.7 Factores de Virulencia *17*
2.2.8 Genotipificación *20*
2.3 Condroitín sulfato 20
2.4 Condroitín Liasas 21
3. ANTECEDENTES **21**
4. JUSTIFICACIÓN **22**
5. OBJETIVOS **23**
5.1 Objetivo General 23
5.2 Objetivos Partículares 23
6. HIPÓTESIS **23**
7. ESQUEMA DE TRABAJO **24**
8. MATERIALES Y MÉTODOS **25**
8.1 Identificación de los factores sigma y reguladores transcripcionales codificados en el genoma de AVPG 2015. 25
8.2 Predicción de los sitios potenciales de reconocimiento de proteínas iniciadoras de la transcripción en el genoma de AVPG2015. 25
8.3 Análisis de las secuencias consenso del reconocimiento por los diferentes Factores Sigma putativos de AVPG2015. 26
8.4 Reconocimiento de regiones que contienen genes de profago en el genoma de AVPG2015. 26
8.5 Identificación de potenciales secuencias promotoras y secuencias blanco de DNA que reconocen reguladores transcripcionales en genes chl1 y chl2. 26

8.6 Exploración de sitios de reconocimiento de factores sigma y factores transcripcionales, en un grupo pequeño de genes, que codifican proteínas de conservación celular y de factores de virulencia. 26
8.7 Deducción de la organización transcripcional de genes en el genoma de AVPG2015. 27
8.8 Elaboración de redes de regulación transcripcional de ciertos genes de AVPG2015. 27
8.9 Predicción de la estructura tridimensional de las proteínas Chl1 y Chl2. 28

9. RESULTADOS 28

9.1 EXPLORACIÓN DEL GENOMA DE AVPG2015 (ACCESO CP058307) PARA DISTINGUIR ENTRE GENES CROMOSOMALES Y DE BACTERIOFAGOS 28

9.2 PROTEÍNAS QUE ACTÚAN COMO REGULADORES TRANSCRIPCIONALES EN AVPG2015 31

9.3 SITIOS Y SECUENCIAS POTENCIALMENTE REGULADORAS DE LA TRANSCRIPCIÓN DE LOS GENES ANOTADOS EN EL GENOMA CP058307 32

9.3.2 Proteínas reguladoras de la transcripción presentes en AVPG 2015 34

9.4 ABUNDANCIA DE SECUENCIAS RECONOCIDAS POR REGULADORES TRANSCRIPCIONALES PREDICHOS EN EL GENOMA DE AVPG2015 35

9.5 FACTORES SIGMA Y REGULADORES TRANSCRIPCIONALES QUE PROBABLEMENTE REGULAN A LOS GENES DE AVPG2015 36

9.6 IDENTIFICACIÓN Y REPRESENTACIÓN DE POSIBLES SECUENCIAS CANÓNICAS DE SECUENCIAS BLANCO DE FACTORES SIGMA CODIFICADOS EN EL GENOMA DE AVPG2015. 39

9.7 UNIDADES TRANSCRIPCIONALES QUE CONFORMAN EL GENOMA DE AVPG2015 42

9.8 POTENCIALES SECUENCIAS REGULADORAS EN CIS DE LOS GENES *CHL1 Y CHL2* 48

9.8.1 Análisis de las secuencias cis en la región líder de *chl*1 48
9.8.2 Análisis de las secuencias *cis* en la región líder de *chl*2 50

9.9 REDES DE REGULACIÓN TRANSCRIPCIONALES PREDICHAS PARA LOS GENES DE CONDROITÍN LIASAS . 51

9.10 PREDICCIÓN DE ESTRUCTURA TRIDIMENSIONAL DE LAS PROTEÍNAS CHL1 Y CHL2 53

10. DISCUSIÓN 60

11. CONCLUSIÓN 64

12. PERSPECTIVAS 65

ABREVIATURAS Y GLOSARIO DE TÉRMINOS

AVPG	*Avibacterium paragallinarum*
Chl y *chl*	Condroitin liasa (Proteína y gen)
IC	Coriza Infecciosa
ERIC-PCR	Enterobacterial Repetitive Intergenic Consensus
HA	Antígeno Hemaglutinina
DO	Densidad Óptica
ORF	Marco de Lectura Abierto
Secuencia logo	Representación gráfica de un patrón, proveniente de un alineamiento multiple de secuencias de DNA o proteínas
Unidad Transcripcional	Segmento o segmentos de ADN transcrito a un único o múltiples ARNs
*gyr*A, *atp*D, *rec*N, *rpo*B	Genes de conservación celular (Housekeeping) de AVPG 2015
fur, *pep*N, *cdt*A, *cdt*B, *cdt*C, *omp*A, *fha*B, *avx*A, *htmp*B	Genes de virulencia de AVPG 2015

ÍNDICE DE IMAGENES

- Imagen 1: Mapa físico del genoma de AVPG2015, en el que se resaltan las posiciones de las secuencias codificantes de elementos regulatorios y de profagos…….26
- Imagen 2: Logo de las secuencias canónicas consenso reconocidas por la proteína sigma 70……36
- Imagen 3: Logo de las secuencias canónicas consenso reconocidas por la proteína sigma 32……37
- Imagen 4: Logo de las secuencias canónicas consenso reconocidas por la proteína sigma 24……37
- Imagen 5: Región 5´ rio arriba del marco de lectura abierto del gen *chl*1. Secuencia de 500 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción……..43
- Imagen 6: Región 5´ rio arriba del marco de lectura abierto del gen *chl*1. Secuencia de 3000 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción……..44
- Imagen 7: Región 5´ rio arriba del marco de lectura abierto del gen *chl*2. Secuencia de 500 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción……..45
- Imagen 8: Región 5´ rio arriba del marco de lectura abierto del gen *chl*2. Secuencia de 3000 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción……..45

- Imagen 9: Red de regulación transcripcional predicha de la proteína Sigma 70 por el programa bioinformático String……47
- Imagen 10: Modelo propuesto de la red de regulación transcripcional que controla la expresion de la proteína Chl1……48
- Imagen 11: Modelo propuesto de la red de regulación transcripcional que controla la expresión de la proteína Chl2……48
- Imagen 12: Estructuras tridimensionales de la proteína Chl1. Modelos generados con el programa SwissModel. A) Chl1.M1, B) Chl1.M2……49
- Imagen 13: Estructura tridimensional de la proteína Chl2. Modelo generado con el programa SwissModel……...51
- Imagen 14: Estructura de la enzima Chl1 de AVPG. Modelos tridimensionales generados mediante el programa ITasser……..52
- Imagen 15: Estructura de la enzima Chl2 de AVPG. Modelos tridimensionales generados mediante el programa ITasser……..52
- Imagen 16: Superposición de dos estructuras tridimensionales, obtenidas de dos cristales de Condroitin liasas, una de Proteus vulgaris y otra de Bacteroides. La superposición se hizo por medio del programa PyMol……..53
- Imagen 17: Superposición de dos estructuras tridimensionales, una del cristal de Condroitin liasa 1 de Proteus vulgaris y otra de Chl1 de AVPG predicha por medio de ITasser. La superposición se hizo por medio del programa PyMol……..54

ÍNDICE DE TABLAS

- Tabla 1: Clasificación Taxonómica de AVPG……...11
- Tabla 2: Características fenotípicas de AVPG (Blackall et,. al 2017)……12
- Tabla 3: Características químicas de los diferentes tipos de Condroitin sulfato…...16
- Tabla 4. Genes de AVPG 2015 utilizados para deducción de regulación transcripcional……...23
- Tabla 5: Segmentos de Profago anotados en el genoma de AVPG2015……...25
- Tabla 6: Factores Sigma codificados en el genoma de AVPG2015……..27
- Tabla 7: Proteínas reguladoras de los factores Sigma codificadas en el genoma de AVPG2015……27
- Tabla 8: Predicción de los posibles reguladores transcripcionales en AVPG2015...29
- Tabla 9: Reguladores transcripcionales que se encuentran codificados en el genoma de AVPG2015 (registro CP058307)……..30
- Tabla 10: Información condensada de factores sigma y reguladores transcripciones que potencialmente actúan en las secuencias reguladoras de los genes *chl*1 y *chl*2……33
- Tabla 11: Información condensada de factores sigma y reguladores transcripciones que potencialmente actúan en las secuencias reguladoras de nueve genes relacionados con la virulencia bacteriana……...34

- Tabla 12: Clases de probables secuencias reconocidas por sigma 70……..35
- Tabla 13: Clases de probables secuencias reconocidas por sigma 32……..36
- Tabla 14: Predicción de secuencias de reconocimiento por factores sigma y reguladores transcripcionales……38
- Tabla 15: Organización transcripcional de los genes de AVPG2015……....39
- Tabla 16: Organización de las probables unidades transcripcionales de 13 genes de AVPG2015……42
- Tabla 17: Resultados estadísticos predictivos para los diferentes modelos propuestos de la proteína Chl1…….50
- Tabla 18: Resultado estadístico predictivo para el modelo propuesto de la proteína Chl2……...51

ÍNDICE DE GRÁFICAS

- Gráfico 1: Frecuencia de secuencias blanco en el genoma de AVPG2015, que son reconocidas por reguladores trasncripcionales……..31
- Gráfico 2: Porcentaje de genes regulados por los factores Sigma codificados en el genoma de AVPG2015……..35
- Gráfico 3: Frecuencia de unidades transcripcionales en el genoma de AVPG2015..41

RESUMEN

La Coriza Infecciosa causada por *Avibacterium paragallinarum* es una enfermedad costosa para los sistemas productivos de carne de ave y huevo, debido a la afectación múltiples etapas de la crianza y producción. Para conocer cómo enfrentar al metabolismo de la bacteria patogénica *A. paragallinarum* cepa 2015 se analizó la estructura de su cromosoma con un enfoque transcriptómico *in silico*. Se diseccionó el genoma en el core del DNA bacteriano y del DNA de profago, ya que este último tiene menos probabilidad de ejercer un efecto directo en la patogénesis de la bacteria. La disección también sería interesante para visualizar el tipo de reguladores trans que pudieran reconocer a las secuencias cis-reguladoras o blanco. Se identificaron los genes putativos codificantes de los factores sigma 70, 24 y 32; *rpo*D, *rpo*E y *rpo*H respectivamente, los cuales serían los selectores de los promotores respectivos, que construyen tres tipos de regulones sigma. Se identificaron también los genes de 54 reguladores transcripcionales putativos que potencialmente van a modificar la fuerza de expresión de los tres posibles regulones sigma. De todos los posibles sitios de reconocimiento de factores sigma y reguladores transcripcionales, el 25.7% corresponden a sigma70, ArgR y ArcA, siendo aparente el dominio de regulación por estas tres proteínas, donde el 74.3% de secuencias blanco corresponden para 51 reguladores potenciales. Dos genes que codifican la actividad de condroitín-liasa, potencialmente importantes para la virulencia de *A. paragallinarum* 2015 se han propuesto previamente, denominados *chl1* y *chl2*. En su secuencia líder se encontró una posible regulación por sigma 70 para el gen *chl1* en tanto que en la región líder proximal de *chl2* no se halló un promotor dependiente de sigma, por lo cual se infiere que *chl2* forma parte de una unidad policistrónica. Los reguladores potenciales que controlan a *chl1* son *crp*, *arg*R2, *ihf*, *lrp*, *fnr* y *arc*A, acomodados en probables 4 promotores. En tanto que para *chl2* se observó menor participación de reguladores en la región de hasta 3000 bases río arriba, de *gvc*A, *arg*R2, *ihf* y *lrp*, acomodados en hasta 3 promotores posibles en una región grande de 3000 bases río arriba. Además, se generaron predicciones de la estructura tridimensional de ambas proteínas Chl1 y Chl2 en diferentes programas bioinformáticos, que nos proporcionó múltiples modelos con diferentes valores estadísticos QMEAN, C-Score entre otros, que dieron validez a los modelos tridimensionales presentados y que serán de utilidad para elucidar más a detalle datos importantes como el sitio activo de la enzima o elementos acerca de su estructura.

ABSTRACT

Infectious Coryza caused by *Avibacterium paragallinarum* is a costly disease for poultry and egg production systems, due to the affectation in multiple stages of breeding and production. To find out how to deal with the metabolism of the pathogenic bacterium *A. paragallinarum* strain 2015, the structure of its chromosome was analyzed with an in silico transcriptomic approach. The genome was dissected into the core of bacterial DNA and prophage DNA, since the latter is less likely to exert a direct effect on bacterial pathogenesis. The dissection would also be interesting to visualize the type of trans regulators that could recognize *cis*-regulatory or target sequences. Putative genes encoding sigma factors 70, 24 and 32 were identified; *rpo*D, *rpo*E and *rpo*H respectively, which would be the selectors of the respective promoters, which would build three types of sigma regulons. The genes of 54 putative transcriptional regulators that are potentially going to modify the strength of expression of the three possible sigma regulons were also identified. Of all the possible recognition sites of sigma factors and transcriptional regulators, 25.7% correspond to sigma70, ArgR and ArcA, the regulation domain being apparent by these three proteins, where 74.3% of white sequences correspond to 51 potential regulators. Two genes encoding chondroitin-lyase activity, potentially important for the virulence of *A. paragallinarum* 2015, have been previously proposed, named *chl*1 and *chl*2. In its leader sequence, a possible regulation by sigma 70 was found for the *chl*1 gene, while in the proximal leader region of *chl*2 a sigma-dependent promoter was not found, for which it is inferred that *chl*2 is part of a polycistronic unit. The potential regulators that control *chl*1 are *crp*, *arg*R2, *ihf*, *lrp*, *fnr* and *arc*A, accommodated in 4 probable promoters. While for *chl*2, lesser participation of regulators was demonstrated in the region of up to 3000 bases upstream, of *gvc*A, *arg*R2, *ihf* and *lrp*, accommodated in up to 3 possible promoters in a large region of 3000 bases upstream. In addition, predictions of the three-dimensional structure of both Chl1 and Chl2 proteins were generated in different bioinformatics programs, which did not present models with multiple statistical values QMEAN, C-Score among others, which validated the three-dimensional models presented and which will be useful for elucidate in more detail important data such as the active site of the enzyme or elements about its structure.

1. INTRODUCCIÓN

Existen varias enfermedades infecciosas en las aves producidas por diferentes agentes etiológicos: bacterias, virus y parásitos. Estas patologías representan un riesgo para el sector avícola por las grandes pérdidas económicas por el uso de medicamentos, la muerte de aves y la reducción de la producción de huevo. Entre estas patologías se puede nombrar la colibacilocis aviar cuyo agente etiológico es *Escherichia coli*; viruela aviar producida por el virus *Variola avium;* tricomoniasis causada por los parásitos *Trichomonas gallinae.*

La coriza infecciosa producida por *Avibacterium paragallinarum* se caracteriza por signos de escurrimiento nasal, edema facial, diarrea y anorexia en aves (13). Cuando *A. paragallinarum* no está asociado con otros patógenos la enfermedad se considera de duración corta y no hay necesidad de dar tratamiento con antibióticos. Sin embargo, es común la infección natural simultánea con otras bacterias provocando la "IC complicada", con una duración mayor que requiere tratamiento farmacológico en la parvada.

Esta enfermedad es una afección del tracto respiratorio superior en pollos, gallinas (*Gallus gallus*) y otras aves, la que ocasiona importantes pérdidas económicas para los avicultores, porque retrasa el crecimiento, en consecuencia hay pérdida de peso o el aumento de sacrificios de aves con coriza infecciosa. Esta enfermedad se reporta en todo el mundo, en México se conoce de brotes en Sonora, Jalisco, Puebla, Yucatán, Michoacán y Morelos (2).

Esta investigación se enfoca en un estudio molecular transcripcional *in silico* de *Avibacterium paragallinarum* agente etiológico de la coriza infecciosa.

2. MARCO TEÓRICO

2.1 Coriza Infecciosa

La coriza Infecciosa, también llamada moquillo (2) es una enfermedad que afecta las vías respiratorias superiores de varios tipos de aves. El agente etiológico de esta enfermedad es *Avibacterium paragallinarum*; esta bacteria anteriormente era conocida como *Haemophilus gallinarum*, pero fue reclasificada al ampliar su caracterización fenotípica y genotípica con métodos moleculares (4).

Los signos clínicos de la coriza infecciosa son: secreción nasal, hinchazón facial, lagrimeo, anorexia y diarrea (22), aunque se ha observado que la coriza infecciosa se complica en

presencia de otros patógenos o por estrés del cuidado o manejo, observando mas signos como artritis y septicemia. En la enfermedad complicada se ha observado la coinfección con bacterias como *Mycoplasma gallisepticum, Mycoplasma synoviae, Pasteurella spp., o Salmonella spp.* (13). La coriza infecciosa (CI) es de corta duración y cede sin tratamiento, en cambio la CI complicada puede prolongarse hasta 6 semanas y puede requerir tratamiento con antimicrobianos (2).

El diagnóstico de la enfermedad está basado en antecedentes de propagación de la enfermedad, síntomas clínicos, cultivo del patógeno *Avibacterium paragallinarum* (AVPG), pruebas bioquímicas o pruebas moleculares. Para el cultivo de AVPG se realiza una toma de muestra del seno infraorbitario mediante hisopado (81). Para el diagnóstico molecular se han reportado el uso de técnicas como PCR de secuencias repetitivas intergénicas de enterobacterias (ERIC-PCR) (46) y la secuenciación del gen que codifica para el RNAr 16S; logrando con ello la identificación rápida y precisa de las bacterias que participan en la coriza infecciosa simple o complicada (1).

2.2 *Avibacterium paragallinarum*

2.2.1 Epidemiología

La coriza infecciosa es una enfermedad cosmopolita (61), de la cual se han reportado casos importantes en EE.UU, Argentina, India, Marruecos, Tailandia y México (2). La infección simple y la recuperación del patógeno se ha reportado en Argentina, Australia, Bulgaria, Canadá, Egipto, Gran Bretaña, Guatemala, Holanda, India, Indonesia, Iraq y Suiza, entre otros países (62). En México se han reportado casos en Sonora, Jalisco, Michoacán, Yucatán, Morelos y Puebla. El único país considerado libre de CI es Nueva Zelanda (46).

La incidencia de la CI puede variar de acuerdo al tipo de aves, la edad del huésped, la línea genética del huésped, la cepa bacteriana y los factores ambientales, como el hacinamiento, el clima y las enfermedades concurrentes (63).

2.2.2 Clasificación de la bacteria *A. paragallinarum*

A. paragallinarum pertenece a la familia *Pasteurallaceae* (2), la cepa tipo ATCC 29545 = NCTC 11296 es Gram negativa, oxidasa y catalasa negativa, es pleomórfica predominando la estructura cocobacilar (83). La bacteria tiene un tamaño de 1 a 3 µm de longitud por 0.4 a 0.8 µm y movilidad nula (27), es clasificada taxonómicamente como se indica en la tabla 1.

Tabla 1. Clasificación Taxonómica de *A. paragallinarum*

Clasificación Taxonómica	
Superreino	Procaryotae
Reino	Eubacteria
División	Gracilicutes
Clase	Protobacteria
Familia	*Pasteurellaceae*
Género	*Avibacterium*
Especie	*paragallinarum*

2.2.3 Requerimientos de crecimiento

Los medios más comunes para el cultivo de *A. paragallinarum* son infusión cerebro-corazón, agar triptosa o infusión de Carne de Pollo, suplementados con suero de pollo o de caballo (2, 11, 22). *A. paragallinarum* crece mejor en atmosfera de CO_2 al 5%, por lo que se considera microaerofílico, aunque no es un requisito esencial para su cultivo, y la temperatura de incubación puede ser de 34 a 42 grados centígrados (12). *A. paragallinarum* adicionalmente necesita como suplementos en el medio de cultivo el factor V (NAD) (64) en su forma reducida u oxidada (11), asi como, cloruro de sodio (12).

2.2.4 Propiedades bioquímicas

A. paragallinarum muestra propiedades bioquímicas que lo distinguen de otras bacterias del género *Avibacterium*, entre las que podemos destacar la utilización de maltosa, D-manitol y D-sorbitol y la incapacidad de fermentación de lactosa, arabinosa, galactosa o trehalosa y la prueba catalasa negativa (12). En la tabla 2 se resumen las propiedades bioquímicas de especies filogenéticamente relacionadas incluyendo *A. paragallinarum*.

Tabla 2. Características fenotípicas de *A.paragallinarum* (Blackall y cols 2017).

Taxon	*Avibacterium gallinarum*	*Avibacterium paragallinarum*	*Avibacterium volantium*	*Avibacterium avium*	*Avibacterium endocarditis*
Catalase	+	-	+	+	+
Symbiotic growth	-	V	+	+	+
Acid form					
L-arabinose	-	-	-	-	+
D-galactose	+	-	+	+	+
Maltose	+	+	+	-	V
D-mannitol	-	+	+	-	V
D-sorbitol	-	+	V	-	-
Trehalose	+	-	+	+	+
α-glucosidase	+	-	+	+	+

2.2.5 Genómica

Hasta noviembre de 2022, existen 45 genomas secuenciados de la especie *A. paragallinarum* y reportados en la base de datos de secuencias genéticas de GenBank de los cuales 15 estan totalmente circularizados y 30 en forma fragmentada o de draft (45). Las tecnologías de secuenciación masiva usadas para conocer la secuencia del genoma de esta bacteria son 454, Illumina, PacBio y Oxford-MinIon, resultando más eficientes las dos últimas, pero tambien son mas caras y escasas hasta el momento. Como ejemplo, la secuencia de la cepa 221 de *A. paragallinarum* como genoma circular tiene una longitud de 2,685,568 pb y un contenido medio de G + C del 41%, además, la predicción del número de genes que codifican proteínas es de 2994 genes con una longitud media de 783 pb. Se detectaron 13 copias de genes ARNr 5S, 16S y 23S (29).

En la Benemérita Universidad Autónoma de Puebla se han secuenciado tres genomas de la bacteria *Avibacterium pargallinarum* (2015, ESV135 y Cl, números de acceso CP050316.1, CP058307.1 y LAEN01000001:LAEN01000154 respectivamente) de los cuales ESV135 y 2015 estan completamente circularizados (Cobos Justo, manuscrito en preparación). El

contenido de GC del genoma de las cepas conocidas va de 40.8 a 41.2% y la longitud de genoma esta en el rango de 2.41 a 2.88 Mb (45).

2.2.6 Serotipificación

A. paragallinarum se ha estudiado con dos esquemas de serotipificación diferentes (2), el esquema de serotipificación de Page, el cual se desarrolló mediante el uso de una prueba de aglutinación en placa y se han reconocido tres serogrupos, A, B y C (13). Se sabe que los 3 serogrupos presentan distintas serovariedades y que cada una de estas se localiza de diferente manera por región y pais (12). Otra clasificación serológica reconocida es la desarrollada por Kume, la cual originalmente se baso en pruebas de Inhibición de la hemaglutinación (25) utilizando células tratadas y sonicadas con KCN, suero hiperinmune de conejo y eritrocitos de pollo fijados con glutaraldehído. A esta clasificación le corresponden los serogrupos I, II y III equivalentes a A, B y C de Page. Mas recientemente se reconocen múltiples serovares por serogrupo, así se identifican para A (A-1,A-2,A-3 y A-4), B (B-1) y C (C-1,C-2,C-3 y C-4), esta clasificación se ha propuesto debido a que muchos aislados no eran tipificables por esquemas de Page (12, 13, 27).

2.2.7 Factores de Virulencia

Los factores de virulencia pueden ser definidos como componentes de un microorganismo que hacen más agresivo al microorganismo patógeno en su hospedero, por tanto contribuyen a provocar enfermedades y son la base de la patogenicidad del organismo, además pueden ser importantes para su viabilidad o dispersión (20). Estos tienen múltiples funciones que incluyen la evasión de la respuesta inmune, la colonización, adquisición de nutrientes e inmunosupresión. *A. paragallinarum* posee varios factores de virulencia, tales como: la cápsula bacteriana, las metaloproteasas, la toxina citoletal distendente, la toxina RTX, la hemaglutinina, las fimbrias y las vesículas de membrana (18, 28, 52, 54, 55, 58, 59).

2.2.7.1 Cápsula Bacteriana

La cápsula bacteriana se encuentra presente en la superficie de microorganismos patógenos (20), juega un rol clave en la patogenicidad mediante la evasión de la respuesta inmune y colonización bacteriana. En *A. paragallinarum* se ha reportado la presencia de dos regiones que contienen genes necesarios para la biosíntesis de la cápsula bacteriana y que contienen genes específicos para la síntesis de condroitina y heparosano resultando estos los principales

componentes químicos de la cápsula (16). Las mutantes que carecen de cápsula son menos virulentas que las cepas de tipo silvestre (WT) (58).

2.2.7.2 Metaloproteasas

Las metaloproteasas son enzimas que como mecanismo principal catalizan la hidrólisis de enlaces peptídicos en proteínas cuyo mecanismo catalítico requiere la presencia de un metal, como Zn o Co; juegan un papel clave en la virulencia incluyendo la colonización, evación de respuesta inmune o adqusición de nutrientes (20). Se ha reportado la presencia de estos factores de virulencia en miembros del género *Pasteurellaceae* como *A. paragallinarum*, entre otros, que son capaces de degradar inmunoglobulinas de tipo IgG de animales, lo que genera evasión de la respuesta inmune del huesped (22, 59).

2.2.7.3 Toxina citoletal distindente (CDT)

La toxina citoletal distindente es una citotoxina presente en bacterias Gram negativas específicamente a la clase gammaproteobacterias (60). Se ha reportado la presencia de estas toxinas en miembros de la familia Pasterellaceae como *Haemophilus ducreiy* o *Avibacterium paragallinarum.* En AVPG la CDT, se encuentra codificado en el operón *cdt*ABC que contiene tres genes que codifican tres subunidades de proteínas CdtA, CdtB y CdtC . La subunidad B posee actividad DNAsa I, la subunidad A y C son necesarias para la adhesión a superficies celulares e introducción al interior de la célula (28).

2.2.7.4 Toxinas RTX

Las Toxinas RTX han sido clasificadas como parte de la familia de toxinas formadoras de poros de bacterias Gramnegativas. Inducen mediadores inflamatorios o realizan efectos citotóxicos en las células inmunitarias del huésped dando como resultado necropsis e inflamación (15). Se ha demostrado que *A. paragallinarum* es capaz de producir una toxina RTX que tiene una fuerte actividad citotóxica (12) denominada AvxA el cual se encuentra codificado en un operón RTX, que está conformado por un gen activador *avx*C, el gen *avx*A que codifica la toxina RTX y el gen *avx*BD el cual codifica un sistema de screción tipo I (55).

2.2.7.5 Hemaglutinina

La hemaglutinina es considerada el antígeno más importante de *A. paragallinarum* debido a que es la distintiva en la serología de la especie y juega un rol importante en la virulencia de

AVPG (54). Las cepas de los diferentes serogrupos A, B y C poseen determinantes antigénicos con capacidad hemoaglutinante de células eritrocitarias de varias especies animales. Se han reportado dos proteínas como componentes de HA potenciales, la proteína HagA y la proteína HMTp210, las cepas mutantes que carecen de esta última pierden la capacidad de adherencia y por lo tanto disminuye su virulencia (26). El antígeno de Hemaglutinina (HA) es específico en los diferentes serogrupos y es considerado un antígeno protector debido a que existe una correlación entre el título del antícuerpo de inhibición de hemaglutinación y protección (26).

2.2.7.6 Fimbrias

Las fimbrias son apéndices similares a capilares con una longitud de 1-2 um y diámetro de 2 a 8 nm que están presentes en microorganismos Grampositivos y Gramnegativos, con función de adhesión a las células huesped (6). Las fimbrias han sido clasificadas en base al tipo de proteína Fimbrial: Curly, Chaperonas y Fimbrias Tipo 4.

Las proteínas fimbriales de *A. paragallinarum* son consideradas un factor de virulencia y potenciales antígenos para vacunas contra AVPG. Se ha identificado un gen que codifica para la proteína fimbrial FlfA en *A. paragallinarum*, este antígeno es capaz de conferir protección con anticuerpos anti FlfA en pollo y las cepas mutantes carentes del antígeno son menos virulentas (18).

2.2.7.7 Vesículas de Membrana

Las vesículas de membrana son producidas principalmente por microorganismos Gramnegativos y se desarrollan en la superficie de la membrana externa que son liberadas durante el metabolismo celular, y cuyas funciones incluyen la colonización de tejidos, obtención de nutrientes o interacciones huesped-microorganismo (17). El contenido de las vesículas de membrana puede incluir compuestos variados como lípidos, lipopolisacáridos, fosfolípidos, proteínas así como se ha reportado la presencia de ADN o ARN en su superficie (53). Dentro de la familia de las Pasterellaceas, se ha descrito la presencia de vesículas de membrana externa en Avibacterium paragallinarum, cuya su función se cree que está vinculada a la secreción de factores de virulencia y por lo tanto, al desencadenamiento de una respuesta inmune en el huesped (52).

2.2.8 Genotipificación

Para la genotipificación de *A. paragallinarum* se han desarrollado dos técnicas moleculares, basadas específicamente en la PCR, una convencional y otra en tiempo real (12).

La PCR convencional, llamada HP-2 PCR, se desarrolló mediante una técnica de clonación al azar, es un método validado en el uso en colonias en agar o moco obtenido del seno infraorbital de aves vivas (13), la PCR de tiempo real posee una mayor sensibilidad analítica que la PCR convencional. Ha sido recomendado un método de PCR que utiliza cebadores universales que es Para amplificar el gen que codifica para el RNAr 16S seguido de digestión con enzimas de restricción (11).

2.3 Condroitín sulfato

O también llamado sulfato de condroitina, es un glucosaminoglucano sulfatado (GAG) el cual esta conformado por cadenas de polisacáridos no ramificados de longitud variable que contienen dos monosacáridos alternos: ácido D-glucurónico (GlcA) y N-acetil-D-galactosamina (GalNAc), y por lo general, se encuentran unidos a proteínas como parte de proteoglicanos (49).

Cabe destacar, que existen diferentes tipos de condroitín sulfato, y estos se clasifican en su mayoría conforme a su origen, específicamente hablando de fuentes terrestres y marinas, y todas con una alta variabilidad en términos de longitud de la cadena y patrón de sulfatación. El sulfato de condroitina de animales terrestres (las cuales cuentan con un peso molecular aproximado de 45 kDa) está compuesto casi exclusivamente por unidades no sulfadas (O) y monosulfadas (A y C), mientras que en las especies marinas (las cuales cuentan con un peso de 70 kDa, y tienden a ser cadenas más largas) la proporción de unidades disulfadas (D, E y B) es mayor (50)(51).

Tabla 3. Características químicas de los diferentes tipos de condroitin sulfato.

Condroitín sulfato	Tipo de Sulfatación	Nombre Sistémico
sulfato de condroitina A	Carbono 4 del azúcar N-acetilgalactosamina (GalNAc)	condroitina-4-sulfato
sulfato de condroitina C	Carbono 6 del azúcar GalNAc	condroitina-6-sulfato
sulfato de condroitina D	Carbono 2 del ácido glucurónico y del 6 del azúcar GalNAc	condroitina-2,6-Sulfato

sulfato de condroitina E	Carbonos 4 y 6 del azúcar GalNAc	condroitina-4,6-Sulfato

Otro punto importante es que el sulfato de condroitina B, el cual actualmente se conoce como sulfato de dermatán o dermatán sulfato, ya no se clasifica como sulfato de condroitina.

2.4 Condroitín Liasas

Esta enzima, la cual tiene número de acceso EC 4.2.2.20 (47), se conoce como condroitina ABC eliminasa o condroitina ABC liasa, pertenece al grupo de enzimas conocidas como polisacaridos liasas.

Esta enzima degrada una variedad de glicosaminoglicanos como condroitin sulfato y dermatán sulfato aunque la enzima también puede actuar sobre el hialuronan en menor medida (31).

Hasta el momento se encuentran catalogadas cuatro tipos de enzimas condroitín liasas, que son A, B, C y AC, cada una de las cuales actúan sobre un sustrato diferente, siendo la condroitina-4-sulfato (sulfato de condroitina A) el sustrato de la Condroitin Liasa A; La condroitina-6-sulfato (sulfato de condroitina C) es el sustrato sobre el que actúa la condroitín liasa C, el dermatán sulfato es el sustrato de la condroitín liasa B y por último la condroitín liasa AC actúa sobre dos sustratos diferentes, el condroitín sulfato A y C (48).

El mecanismo de acción de la enzima, consiste en la escisión endolítica de enlaces (1->4)-beta-galactosaminicos entre la N-acetilgalactosamina y el ácido D-glucurónico o el ácido L-idurónico para producir una mezcla de oligosacáridos insaturados delta4 de diferentes tamaños que finalmente se degradan a tetra y disacáridos insaturados delta4 (32).

3. ANTECEDENTES

En el centro de investigaciones en Ciencias Microbiológicas se está desarrollando investigación de la genómica de *A. paragallinarum* y la identificación estructural de los genes de virulencia de acuerdo a lo establecido por Horta G., Cobos E. y cols (no publicado). Parte de esta investigación está enfocada en la parte funcional de los genes de las condroitin liasas como fue establecido por Cobos E., Barcenas G, Ramos F. y cols (no publicado). Se ha realizado la secuenciación de catorce genomas de *A. paragallinarum* a partir de cultivos puros. De estos, tres genomas estan reportados en el GenBank, número de acceso CP050316.1, CP058307.1 y LAEN01000001:LAEN01000154. Se ha hecho el ensamblado

de los fragmentos genómicos mediante programas bioinformáticos para la construcción de un genoma circular, ejemplos CP050316.1, CP058307.1. Además, la búsqueda de los genes que codifican factores de virulencia ha dado como resultado, la confirmación de la mayoría de los genes que están implicados en la virulencia, aunque, existen otros genes que no están descritos en *A. paragallinarum* como factores de virulencia. Uno de estos es el gen que codifica la condroitinasa o condroitin liasa. Se describió en *A. paragallinarum* 2015 la presencia de 2 genes que codifican la condroitinasa, uno quizas esta relacionado con una condroitinasa que actúa sobre polisacáridos cortos y otro gen que codifica una enzima que escinde polisacáridos largos. Dependiendo del polisacárido disponible es probable que la bacteria active el gen correspondiente. Además, se ha realizado una búsqueda y comparación de genes de condroitin liasas en otras bacterias patógenas, y se ha observado que están presentes los genes de condroitinasa, sugiriendo que estos genes pueden estar relacionados con la virulencia. Se han diseñado primers que permiten amplificar por PCR los genes *chl*1 y *chl*2 de la condroitinasa, así se comprobó que ambos genes están presentes en todas las cepas en general de *A. paragallinarum* en el laboratorio MIMOCE-ICUAP. Se prepararon además medios de cultivo que tuvieran condroitina (Agar sangre o BHI) para analizar el crecimiento de esta bacteria, observando un mayor crecimiento y cambios en la morfología de las colonias. Se hicieron curvas de crecimiento de *A. paragallinarum* en medios líquidos enriquecidos con sulfato de condroitina, demostrando que este microorganismo crece mejor cuando está enriquecido con condroitina; la DO con medios enriquecidos puede llegar a 1 ó 1.1 a las 9 horas mientras que en BHI sólo alcanza una DO máxima de 0.8 a las 14 horas de crecimiento como fue reportado por Cobos E. y cols, 2021 (84). Recientemente Barcenas G. y cols, 2020 (79) demostraron la existencia de los dos operones de condroitin liasas y se inició la clonación de los genes con el fin de expresar proteinas etiquetadas para la inducción de anticuerpos anticondroitin liasa en animales de laboratorio.

4. JUSTIFICACIÓN

A. paragallinarum infecta a las aves y las enferma de coriza infecciosa. Problema que es de interés científico y económico. Por tanto, es importante conocer los factores de virulencia bacterianos que se producen *in vivo* durante la CI que padecen aves y que puedan ser blanco de estudio para posibles soluciones que ayuden a controlar esta enfermedad. Actualmente debido a la pandemia de SARS-COV2, nos hemos enfocado al estudio *in silico* de la

transcripción en AVPG para apoyar los resultados que se obtengan a partir de las investigaciones comparativas de los genomas de AVPG y de los experimentos de qRT-PCR con cultivos de AVPG y su comparación con los estudios previos de expresión heteróloga de genes de condroitin liasa en *E. coli.*

5. OBJETIVOS

5.1 Objetivo General

Objetivo: Explorar los eventos regulatorios posibles que afectan la expresión de los genes del genoma de *A. paragallinarum* 2015 con énfasis en los que regulan las proteínas Chl1 y Chl2.

5.2 Objetivos Partículares

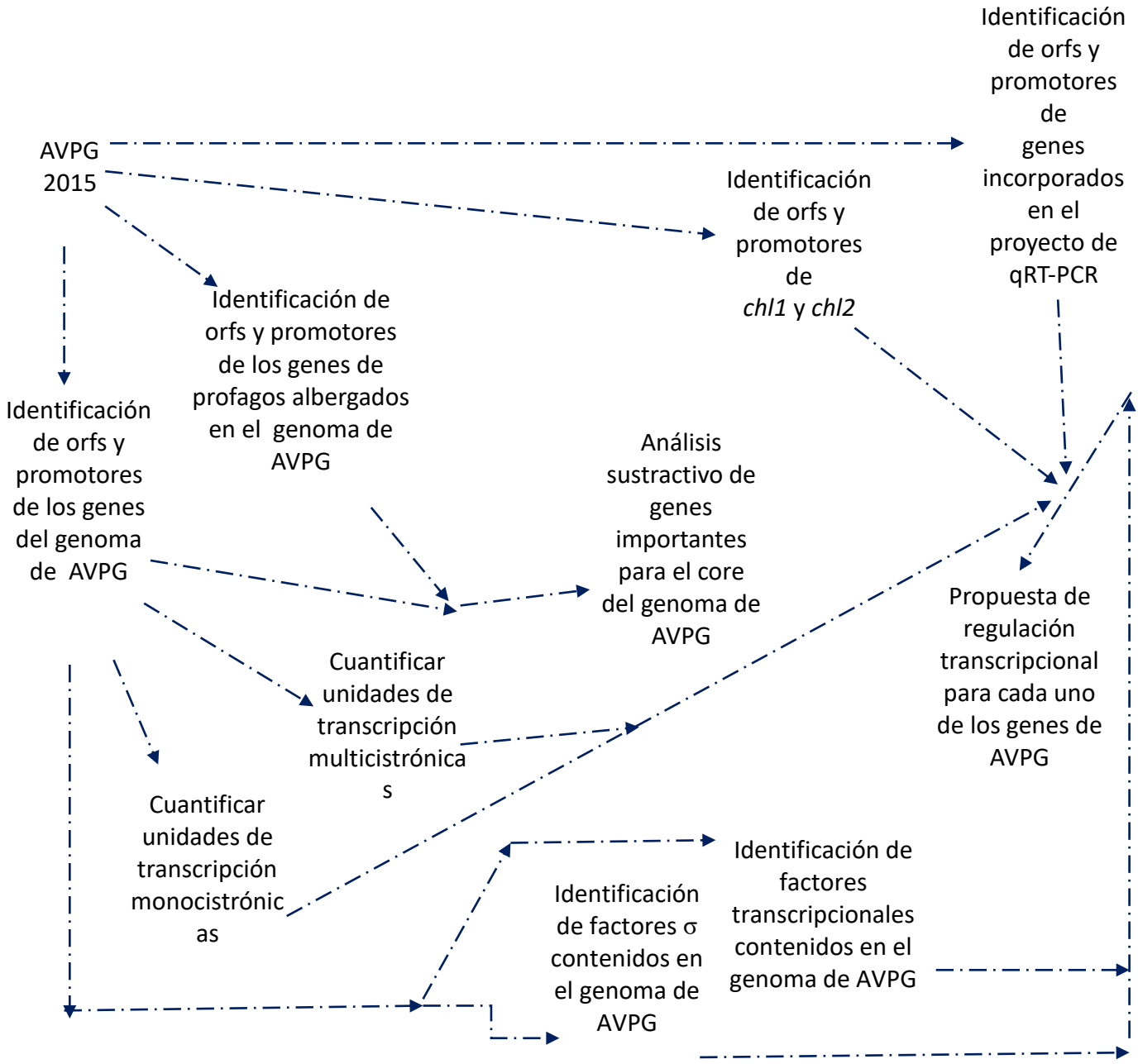
- Explorar el genoma de AVPG 2015 para distinguir entre genes cromosomales y de profagos.
- Identificar todas las proteínas que actúan como reguladores transcripcionales en AVPG2015.
- Identificar todas las secuencias potencialmente *cis*-reguladoras de la transcripción de los genes anotados en CP058307.
- Determinar la frecuencia de los sitios potenciales de inicio de la transcripción.
- Identificar las potenciales secuencias reguladoras en *cis* de los genes en AVPG2015, mediada por factores sigma.
- Determinar las secuencias consenso y el logo respectivo de las cis-reguladoras.
- Deducir la organización de genes en el genoma de AVPG2015.
- Identificar las potenciales secuencias reguladoras en cis de los genes *chl*1 y *chl*2.
- Proponer un modelo de regulación transcripcional asociado a los genes *chl*1 y *chl*2.
- Proponer un modelo de estructura tridimensional de las proteínas Chl1 y Chl2.

6. HIPÓTESIS

Se sugiere que los dos genes de condroitin liasas (*chl*1 y *chl*2) identificados en el genoma de *Avibacterium paragallinarum*, contribuyen a la virulencia de esta bacteria. Como consecuencia de la expresión de las condroitin liasas *in vivo* durante la coriza infecciosa es posible que se establezca una respuesta inmune contra estas enzimas y contra *A. paragallinarum*. En tal caso sería posible usar estas condroitin liasas como inmunógenos para

prevenir la infección de *A. paragallinarum*. Sin embargo estan pendientes los estudios de expresión heteróloga y homóloga de los genes *chl*1 y *chl*2 de AVPG para consolidar la hipótesis y para elegir el mejor sistema de expresión de estos antígenos.

7. ESQUEMA DE TRABAJO



8. MATERIALES Y MÉTODOS

8.1 Identificación de los factores sigma y reguladores transcripcionales codificados en el genoma de AVPG 2015.

Se trabajó con el genoma de *A. paragallinarum* cepa 2015 (GenBank acceso CP058307) se hizo una búsqueda bioinformática de los genes que codifican factores sigma así como genes que codifican reguladores de la transcripción en el genoma de AVPG2015.

Se hicieron análisis de alineamientos de secuencias proteína-proteína tanto en la plataforma de BLAST (56), como en la plataforma de RAST (57), se usaron secuencias de *A. paragallinarum* o de *E. coli* con anotaciones correctas para factores sigma o de reguladores transcripcionales. Cuando no hubo anotaciones precisas en los alineamientos se consideraron como validos los resultados que tenían un porcentaje de identidad mayor al 50-60% y valores E mayores a 1e-18.

8.2 Predicción de los sitios potenciales de reconocimiento de proteínas iniciadoras de la transcripción en el genoma de AVPG2015.

En *E. coli* y en otras pocas bacterias se ha investigado cómo se unen las proteínas iniciadoras de la transcripción, tales como factores sigma y proteínas que regulan a la holoenzima de la RNA polimerasa. La información está registrada para llevar a cabo la búsqueda predictiva de secuencias semejantes en el DNA de otros microorganismos. La predicción de sitios potenciales de reconocimiento o unión para el control del inicio de la transcripción en el genoma de *A. paragallinarum* 2015, se hizo por medio la plataforma de BPROM (65). Para esta búsqueda predictiva se usaron dos rangos de secuencias, 500 pb y 3000 pb río arriba del ORF de los 2490 genes anotados por RAST y GenBank con el fin de analizar el poder predictivo del programa BPROM para cada gen y saber si la búsqueda de 3000 pb contenía los sitios de la secuencia de 500pb. Era importante conocer el poder predictivo del programa BPROM para reajustar el tamaño del segmento de DNA que se pretendía analizar y la confiabilidad del resultado de la predicción.

8.3 Análisis de las secuencias consenso del reconocimiento por los diferentes Factores Sigma putativos de AVPG2015.

Todas las secuencias predichas para el reconocimiento de los factores sigma contenidos en el genoma de *A. paragllinarum* 2015, σ70 o Sigma A, σ32 o Sigma H y σ24 o Sigma E, se agruparon según el tipo de factor sigma, se alinearon con Clustal Omega para identificar los consensos y se procesaron con WebLogo para representar gráficamente los consensos. WebLogo (66).

8.4 Reconocimiento de regiones que contienen genes de profago en el genoma de AVPG2015.

Según los registros de RAST y GenBank existen anotaciones de genes de profago (bacteriofagos residentes) en el genoma de *A. paragallinarum* 2015. Del registro de Genbank CP058307 se separaron los genes de profago, se registraron las posiciones en el genoma, el tamaño de la región y el número de genes por región. La información se procesó por Excel y se representó en un mapa físico.

8.5 Identificación de potenciales secuencias promotoras y secuencias blanco de DNA que reconocen reguladores transcripcionales en genes chl1 y chl2.

En el genoma de A. paragallinarum 2015 se han identificado dos probables genes para la utilización o metabolismo de sulfatos de condroitina. Los genes se han nombrado *chl*1 y *chl*2 y se deduce que codifican las proteínas Chl1 (tamaño de 3006 bases) y Chl2 (Tamaño 3099 bases). Se predijo con BPROM los probables sitios de reconocimiento de factores sigma y reguladores transcripcionales, utilizando dos tamaños de DNA líder de los ORF identificados (3000 y 500 pb) y elaborar los mapas físicos de las secuencias predichas en cada gen que codifica la condroitin liasa.

8.6 Exploración de sitios de reconocimiento de factores sigma y factores transcripcionales, en un grupo pequeño de genes, que codifican proteínas de conservación celular y de factores de virulencia.

Para saber y contrastar el tipo de regulación probable que se deduce de genes de conservación celular "housekepping" y de genes de virulencia, se eligieron los genes de *A. paragallinarum 2015* que se indican en la tabla 4, y de ellos se analizaron las regiones 5´de 500 y 3000 pb río arriba de los genes para conocer el tipo de proteínas transcripcionales que tienen probables sitios de reconocimiento en sus respectivos promotores.

Tabla 4. Genes de AVPG 2015 utilizados para deducción de regulación transcripcional.

Genes de conservación celular	Genes de virulencia
*gyr*A, *atp*D, *rec*N, *rpo*B	*chl*1,*chl*2, *fur*, *pep*N, *cdt*A, *cdt*B, *cdt*C, *omp*A, *fha*B, *avx*A, *htmp*B

8.7 Deducción de la organización transcripcional de genes en el genoma de AVPG2015.

Se analizó el número de genes que se encontraban en el mismo sentido, así como el tamaño de las regiones intergénicas, para agrupar los genes en las categorías de monocistrónicos o policistrónicos, de acuerdo al número de genes contiguos predicho con el mismo sentido transcripcional.

8.8 Elaboración de redes de regulación transcripcional de ciertos genes de AVPG2015.

La base de datos String (https://string-db.org/) y el programa interfase "on-line" permitió predecir cómo se relacionan los genes por sus posibles interacciones en el proceso de regulación de la expresión genética, además se exploraraaron las potenciales redes de regulación que pudieran estar operando en control de los genes *chl*1 y *chl*2. El algoritmo toma en consideración para ensamblar la red de regulación, interacciones conocidas, o predichas por medio de: la vecindad de genes, la coocurrencia, la presencia en bases de datos, minería de datos, homología, la coexpresión, la experimentación (genética, bioquímica, biofísica, etc) (67).

8.9 Predicción de la estructura tridimensional de las proteínas Chl1 y Chl2.

La estructura terciaria de las condroitín liasas 1 y 2 se predijo por medio de las interfases "on-line" de los programas computación predictiva: Swiss Model (https://swissmodel.expasy.org), I-Tasser (https://zhanggroup.org/I-TASSER/) y PyMol. Para correr el programa en SwissModel workspace, se usaron las secuencias en formato fasta (68,69), y se obtuvieron resultados, PDB view (70), QMEANDisCo (72, 73), QMEAN (74), QSQE (75) e IDT (76), mapas de identidad, y secuencias relacionadas por identidad de dominios cristalizados. En I-Tasser se introdujo la secuencia de aminoácidos de ambas proteínas en formato FASTA y se generaron modelos predictivos en base a modelos de proteínas cristalizadas contenidos en su base de datos (77).

Para el programa PyMOL se introdujeron los modelos obtenidos de I-Tasser para compararlos y sobreponerlos con modelos de proteínas cristalizadas de condroitin liasa extraídos de la base de datos de proteínas PDB de dos microorganismos: *Bacteroides thethainomicron* y *Proteus vulgaris* (78).

9. RESULTADOS

9.1 Exploración del genoma de AVPG2015 (acceso CP058307) para distinguir entre genes cromosomales y de bacteriofagos

Teniendo como antecedente el trabajo de Roodt y col. (2012) y de Horta y col. (2017), respecto a la presencia de una secuencia con homología al fago Mu, otra al fago Lambda y una tercera secuencia similar al fago HP2 en *Avibacterium paragallinarum* cepa Modesto (Secuencia de referencia en Genbank: NZ_JPZA00000000.1), se registraron todos los sitios del genoma de AVPG2015 que contienen genes de profago, muchos sin identificar, y se registraron 7 segmentos a lo largo del genoma; las cuales se detallan en la tabla 5.

Tabla 5. Segmentos de Profago anotados en el genoma de AVPG2015.

Segmento	Posición en el genoma	Tamaño	Número de Genes
1	80595-88281	7,686 pb	10
2	92386-192753	10,0367 pb	143
3	398452-474897	76,445 pb	85
4	1555658-1591653	35,995 pb	38
5	1598708-1642101	43,393 pb	68
6	2344828-2368759	23,931 pb	24
7	2511453-2532300	20,847 pb	31

En estos segmentos de DNA de profago se encuentran varios genes que codifican para proteínas con funciones virales específicas, entre las que podemos destacar: proteínas de cápside, proteínas de la cola del fago, integrasas, transposasas y varios genes que codifican proteínas hipotéticas y proteínas con dominio DUF (Proteínas de dominio desconocido).

Las secuencias de genes de profago mapeadas en siete regiones tienen la probabilidad de que contribuyan a la regulación transcripcional local en el cromosoma de *Avibacterium paragallinarum*, ya sea por su conformación estructural o bien por su influencia en *cis* sobre los genes vecinos. Basados en el mapa construido (imagen 1) se descarta una función más general basada en factores sigma, o proteínas reguladoras de la transcripción codificadas por los profagos. El mapa cromosomal con la ubicación de los loci virales se muestra en la imagen siguiente con color rojo. En color morado se indican los genes que codifican los factores sigma. En color amarillo amarillo se representan la posición de los genes que codifican proteínas que actúan como reguladores transcripcionales en el genoma AVPG2015 y en color verde se representan proteínas que regulan a los factores sigma.

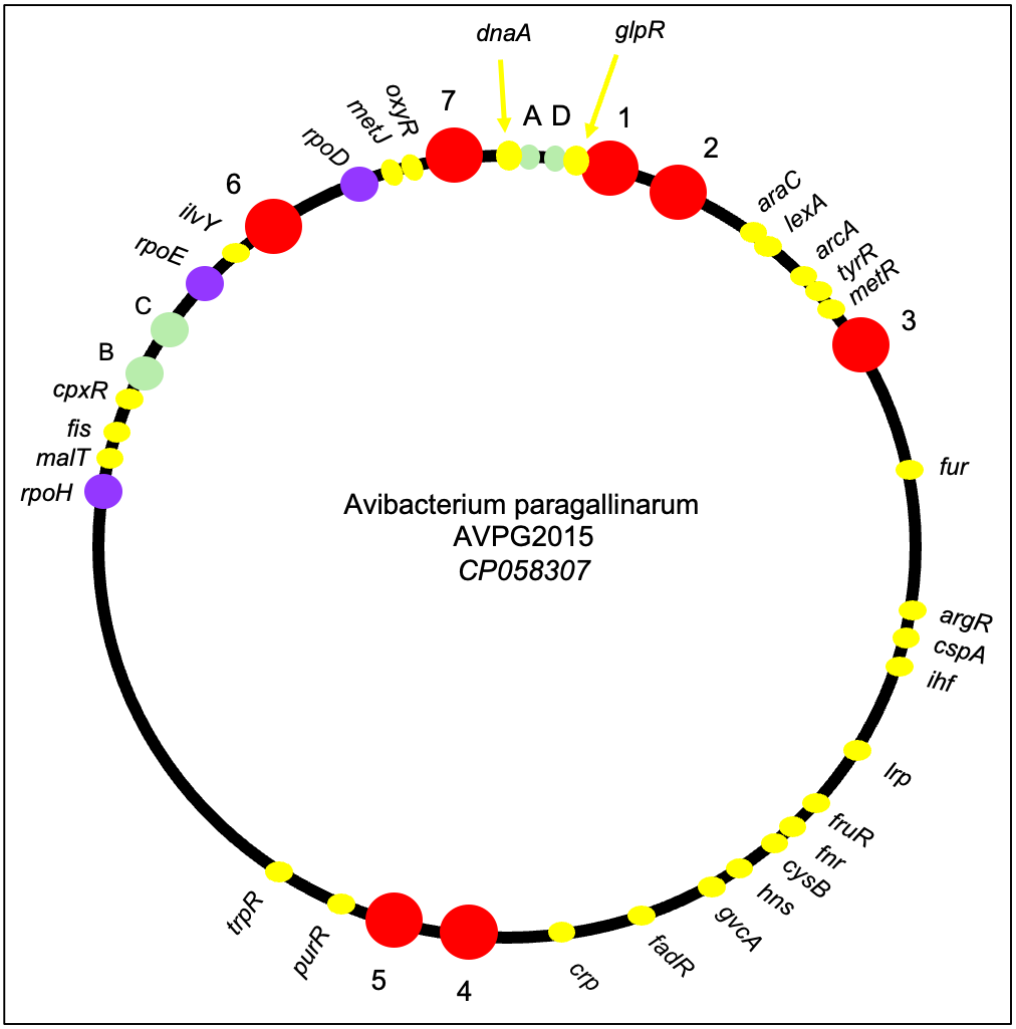


Imagen 1. Mapa físico del genoma de AVPG2015, en el que se resaltan las posiciones de las secuencias codificantes de elementos regulatorios y de profagos. Los círculos rojos representan las posiciones de los segmentos del fago en el genoma AVPG2015, los círculos morados representan la posición de los factores Sigma que están codificados en el genoma, los círculos azules representan la posición de los genes que codifican las proteínas que pueden afectar la función de los factores sigma descritos y, por último, los círculos amarillos representan la posición de los genes que codifican proteínas que actúan como reguladores transcripcionales en el genoma AVPG2015.

Basándonos en la posición de los genes que codifican factores sigma, éstos están en la porción cromosomal y no en la fágica. Por tanto, ningún fago codifica para factores sigma conocidos que contribuya a la regulación transcripcional en AVPG2015.

9.2 Proteínas que actúan como reguladores transcripcionales en AVPG2015

Las principales proteínas que regulan la transcripción son los factores sigma. Mediante la exploración de la anotación del genoma de AVPG2015 se encontraron 3 genes que codifican para factores sigma, así como 4 proteínas que regulan a los factores sigma mencionados y se enlistan en las tablas 6,7.

Tabla 6. Factores Sigma codificados en el genoma de AVPG2015

Gen	Factor σ	Tamaño	Posición en el genoma
gene="*rpo*D" locus_tag="VY92_012095"	σ70 o Sigma A	1862 a.a	2480982..2482844
gene="*rpo*H" locus_tag="VY92_009315"	σ32 o Sigma H	845 a.a	1916884..1917729
gene="*rpo*E" locus_tag="VY92_011725"	σ24 o Sigma E	575 a.a	2394545..2395120

Tabla 7. Proteínas reguladoras de los factores Sigma codificadas en el genoma de AVPG2015

Gen	Proteína codificada	Tamaño	Posición en el genoma
gene="*rse*P" locus_tag="VY92_001350"	Proteasa de Sigma E	1331 a.a	225260..226591
gene="*rse*B" locus_tag="VY92_011715"	Proteína reguladora de Sigma E	959 a.a	2392889..2393848
locus_tag="VY92_011720"	Proteína de regulación negativa de Sigma E	575 a.a	2393924..2394499
locus_tag="VY92_001800"	Regulador transcripcional que interactúa con Sigma 54	938 a.a	330518..331456

9.3 Sitios y secuencias potencialmente reguladoras de la transcripción de los genes anotados en el genoma CP058307

Previamente se mostró la ausencia de expresión funcional de ambos genes de condroitin liasa de AVPG 2015, clonados en *E. coli* de manera independiente. Esta falta de expresión podría ser debida a la necesidad de ambos genes en la misma construcción (79). Como posibilidad de la falta de expresión se puede considerar la necesidad de la presencia de ambas enzimas Chl, para observar mayor crecimiento como se observa para *A. paragallinarum.* En el trabajo anterior referido, también pudo presentarse una regulación alterada por fenómenos de secreción de las enzimas o bien por ausencia de transcripción. El último aspecto resulta muy llamativo dado que en el trabajo anterior se usaron incluso clonaciones de segmentos largos de DNA de AVPG, que contenían los ORF con sus respectivos extremos 5´de hasta 2 kb, por eso se pretende analizar en este trabajo el tipo de secuencias *cis*-reguladoras de las regiones 5´ líder de los ORF de los genes *chl*1 y *chl*2 de *A. paragallinarum* por medio de la exploración y uso de la base de datos de promotores de *E. coli* con la herramienta BPROM (65). El trabajo bioinformático consistió en hacer el estudio sistemático en todo el genoma y de allí extraer la información para los genes *chl*1 y *chl*2.

La herramienta BPROM (65) sabemos que hace la predicción de posibles promotores bacterianos, proporciona las posibles posiciones de interacción de factores sigma y otros reguladores transcripcionales. En otras investigaciones con el uso de esta herramienta se ha reportado que la transcripción de genes bacterianos regulados por promotores sigma 70, que es la clase principal de promotores de *E. coli*, tiene una precisión y especificidad del 80%, valor alto y relativamente confiable para estudios de predicción en nuestro modelo de trabajo. Se obtuvo información predictiva de promotores con diferentes tamaños de secuencias de DNA, considerando los 2490 genes del genoma de AVPG2015. Se trabajó primeramente con segmentos de 120,000 bases y posteriormente con fragmentos de 3000 y 500 bases respectivamente. Por la anotación del genoma de *A. paragallinarum* 2015 sabemos que existen 65 reguladores transcripcionales, los sitios potenciales totales de reconocimiento en el DNA se presentan en la tabla 8.

Tabla 8. Predicción de los posibles reguladores transcripcionales en AVPG2015.

Sitio de reconocimiento	# de Sec en Profagos	# de Sec en Cromosoma	Sitio de reconocimiento	# de Sec en Profagos	# de Sec en Cromosoma
ada X	2	28	*marR* X	3	23
crp	97	860	*melR* X	1	19
fis	58	623	*metJ*	21	149
Fnr	60	500	*metR*	29	369
Fur	29	272	*modE* X	3	22
Hns	22	218	*nagC* X	50	403
ihf	147	1223	*narL* X	9	82
Lrp	127	1088	*narP* X	5	24
tus X	11	81	*ntrC* X	2	16
araC	10	91	*ompR* X	36	308
arcA	91	691	*oxyR*	14	91
argR	144	1097	*pdhR* X	8	57
carP X	5	120	*phoB* X	59	543
cpxR	35	266	*purR*	60	461
cspA	2	5	*rhaS* X	0	2
cynR X	9	67	*rpoE*	3	15
cysB	7	53	*rpoN* X	3	36
cytR X	4	32	*soxS* X	38	321
deoR X	22	208	*torR* X	2	10
dnaA	9	87	*trpR*	4	24
fadR	2	16	*tyrR*	41	382
farR X	2	31	*argR2*	112	768
fhlA X	11	57	*flhCD* X	16	44
fruR	0	2	*phoB3* X	0	6
galR X	1	15	*rpoH2*	19	245
gcvA	16	126	*rpoH 3*	27	154
glpR	19	156	*rpoD15*	71	606
hipB X	7	68	*rpoD16*	130	907
ilvY	2	11	*rpoD17*	201	2033

lacI X	1	15	*rpoD18*	84	579
lexA	102	698	*rpoD19*	30	297
malT	1	5	*rpoS17* X	18	135
			rpoS18 X	0	18

X. El símbolo indica que los reguladores están ausentes en AVPG, por lo que no son funcionales de acuerdo al la red de regulación conocida para *E. coli.* La predicción se hizo por medio de SofBerry y el programa Bprom.

9.3.2 Proteínas reguladoras de la transcripción presentes en AVPG 2015

Dado que la base de datos de regulación transcripcional de BPROM está basada en *E. coli* se pudo predecir el tipo de secuencias que pueden ser reconocidas en el fondo genético de *E. coli*. Así se predijeron 53 posibles blancos de regulación por proteínas reguladoras de la transcripción. Para confirmar la presencia de los reguladores predichos se hizo la búsqueda con BLASTP por secuencia de aminoácidos en las proteínas anotadas en el genoma de AVPG 2015. Se tomó como real la presencia de reguladores transcripcionales si se obtenía un resultado de identidad mayor al 60%, y se concluyó con este corte que sólo encuentran 27 genes que codifican proteínas de regulación transcripcional en AVPG; los datos se muestran en la tabla 9, en la que se indica también la posición del gen en el genoma.

Tabla 9. Reguladores transcripcionales que se encuentran codificados en el genoma de AVPG2015 (registro CP058307)

Gen	Tamaño	Posición
crp	611 pb	1520212..1520823
fis	299 pb	2087183..2087482
fnr	773 pb	993958..994731
fur	446 pb	510208..510654
hns	395 pb	1056835..1057230
ihf	296 pb	770745..771041
lrp	479 pb	842287..842766

araC	896 pb	258553..259449
arcA	713 pb	304743..305456
argR	467 pb	707455..707922
cpxR	683 pb	2038483..2039166
cspA	221 pb	718089..718310
cysB	971 pb	997933..998904
dnaA	1364 pb	988..2352
fadR	725 pb	1322926..1323651
fruR	986 pb	960425..961411
gcvA	890 pb	1009634..1010524
glpR	758 pb	53311..54069
ilvY	878 pb	2312120..2312998
lexA	635 pb	267898..268533
malT	641 pb	2122967..2123608
metJ	317 pb	2505416..2505733
metR	929 pb	360391..361320
purR	1010 pb	1061523..1062533
trpR	299 pb	1100113..1100412
tyrR	938 pb	330518..331456
oxyR	902 pb	2063409..2064311

9.4 Abundancia de secuencias reconocidas por reguladores transcripcionales predichos en el genoma de AVPG2015

En esta parte se exploró el número de secuencias blanco que pueden ser reconocidas en el genoma de *A. paragallinarum* por las proteínas reguladoras del fondo genético de *E. coli*. El resultado del número de secuencias reconocidas por los 65 reguladores de *E. coli* en el genoma de *A. paragallinarum* 2015 se muestra en la gráfica 1, donde destacan además de los blancos para *rpo*D17 y *rpo*D16, los blancos para los reguladores *arg*R, *arg*R2 y *arc*A, el parámetro de búsqueda fueron segmentos de 120 mil bases. Los reguladores mencionados en este apartado están codificados en el genoma de *A. paragallinarum 2015* por lo que la

predicción tiene mucha certeza funcional. Los sitios consenso encontrados coinciden con la predición de los genes de las proteínas que los reconocen albergados en AVPG 2015.

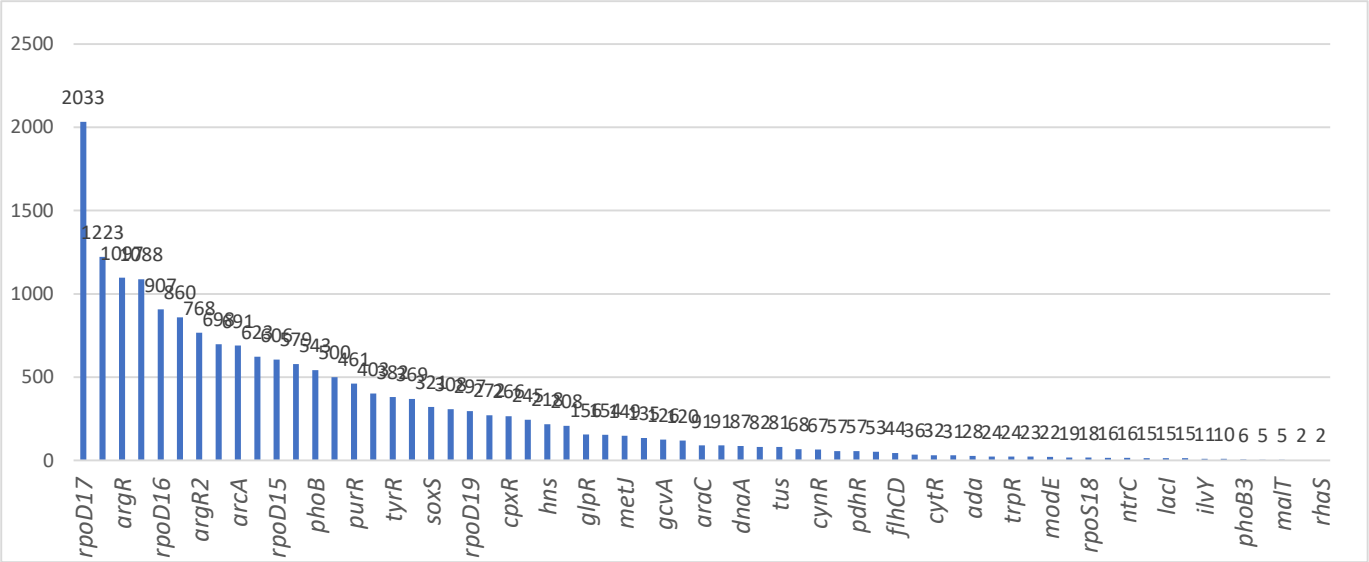


Gráfico 1. Frecuencia de secuencias blanco en el genoma de AVPG2015, que son reconocidas por reguladores trasncripcionales.

9.5 Factores Sigma y Reguladores Transcripcionales que probablemente regulan a los genes de AVPG2015

Para predecir si los genes *chl*1 y *chl*2 pueden seguir un comportamiento de expresión transcripcional similar al de otros factores de virulencia de *A. paragallinarum*, se exploraron las regiones líderes de 500 y 3000 bases para los genes que se muestran en las tablas 10 y 11. En la tabla 10 se muestra el resultado comparativo de los genes *chl* y de cuatro genes "housekeeping". Con el tamaño de 3000 bases de la región líder 5´ se observa la presencia de secuencias blanco para RpoD, excepto para *chl*2 y con 500 bases además de faltar en *chl*2 también falta en *atp*D. El faltante en *atp*D significa que no se predice un promotor cercano y que existe con otro consenso o se usa lo predicho con 3000 bases. Quizás algo similar ocurra con *chl*2, aunque, el promotor puede ser más distal a las 3000 bases o de momento no es posible predecirlo con la información de la base de datos para *E. coli*. Respecto a los factores transcripcionales se observa que son del tipo de los que reestructuran la estructura

tridimensional del genóforo, que responden a la represión catabólica. También destaca la posibilidad de que se esté respondiendo a nivel de nutrición de tipo aminoacídico, ya que se detectan sitios para TyrR y ArgR.

Tabla 10. Información condensada de factores sigma y reguladores transcripciones que potencialmente actúan en las secuencias reguladoras de los genes *chl*1 y *chl*2.

	Rango de búsqueda: 3000 pb			Rango de búsqueda 500 pb		
Gen	Posibles secuencias promotoras	Factor Sigma	Reguladores Transcripcionales	Posibles secuencias promotoras	Factor Sigma	Reguladores transcripcionales
*chl*1	4	*rpo*D (3)	*crp* (4), *arg*R2 (3), *ihf* (1), *lrp (1), fnr* (1)	2	*rpo*D (1)	*tyrR* (1), *ihf* (1), *arcA* (1), *fnr* (2)
*chl*2	3	N/A	*gcv*A (1), *argR2* (2), *ihf* (1)	2	N/A	*lrp* (1)
Genes housekeeping incorporados en el proyecto qRT-PCR						
*gyr*A	8	*rpo*D (6)	*arg*R (1), *hns* (1), *ihf* (4), *gcv*A (1), *lrp* (1), *cys*B (1), *pur*R (1), *arg*R2 (1), *lex*A (1)	2	*rpo*D (1)	*arg*R (1), *hns* (1), *ihf* (1)
*atp*D	8	*rpo*D (2)	*hns* (1), *crp* (1), *ihf* (1)	1	N/A	N/A
*rec*N	7	*rpo*D (8)	*fis* (2), *arc*A (2), *crp* (2), *pur*R (1), *glp*R (1), *ara*C (1), *ihf* (3), *arg*R (1), *fur* (1), *fis* (2), *tyr*R (1), *lex*A (1)	2	*rpo*D (2)	*tyr*R (1), *ihf* (1), *arcA* (1)
*rpo*B	7	*rpo*D (7)	*tyr*R (2), *lrp* (2), *arc*A (1), *metR* (1), *lex*A (2)	1	*rpo*D (2)	*tyr*R (2), *ihf* (1)

*Se muestra la diferencia de resultado predictivo al analizar 500 y 3000 pb de la región UTR. Se muestra el resultado comparativo de predicción en 4 genes HouseKepping

En la tabla 11 se presenta la predicción de secuencias reguladoras en *cis* presentes en genes relacionados con la virulencia de *A. paragallinarum*. Se propone que los genes de virulencia pudieran estar regulados por RpoD o por RpoD y RpoH; la activación por RpoD es más probable si se toman en cuenta secuencias cortas de 500 bases, en tanto que si se toman en cuenta secuencias de 3000 bases se propone la participación de los dos factores sigma. Es

interesante ver que RpoD pudiera controlar varios de estos genes de virulencia como el caso de *fur*, *cdt*C, *hmtp*210. Tienen una dualidad de expresión *pep*N, *cdt*A, *cdt*B, *avx*A, *fha*B, destacando que en *cdt*A se puede predecir como blanco para RpoD y RpoH, tomando en cuenta 500 y 3000 bases respectivamente, lo que significa que pudiera existir una regulación compleja de *cdt*A, en tanto que es menos complicada para los genes vecinos del operón *cdt*B y *cdt*C. En los genes de virulencia también se observa que cada gen pudiera ser influenciado por otros reguladores transcripcionales, como ocurre para los genes housekeeping y *chl*1 y *chl*2, aunque cada gen de virulencia tiene su propia complejidad de regulación.

Tabla 11. Información condensada de factores sigma y reguladores transcripciones que potencialmente actúan en las secuencias reguladoras de nueve genes relacionados con la virulencia bacteriana.

	Rango de búsqueda: 3000 pb			Rango de búsqueda: 500 pb		
Genes de virulencia incorporados en el proyecto qRT-PCR						
Gen	Posibles secuencias promotoras	Factores Sigma	Reguladores transcripcionales	Posibles secuencias promotoras	Factores Sigma	Reguladores Transcripcionales
fur	8	*rpo*D (4)	*arg*R2 (1), *ihf* (3), *met*R (1)	1	N/A	N/A
*pep*N	8	*rpoD* (3) *rpo*H (1)	*fis* (1), *crp* (2), *arc*A (1), *tyr*R (1), *fnr* (2), *arg*R (3), *lrp* (1)	1	*rpo*D (2)	*tyrR (1), ihf (1), arcA (1), purR (2), hns (1)*
*cdt*A	7	*rpo*D (8) *rpo*H (1)	*ihf (6), tyrR (2), argR (2), cpxR (2), fis (1), fur (1), lexA (3)*	2	*rpo*D (1) *rpo*H (2)	*arcA (1), argR (2), cpxR (1), lrp (2)*
*cdt*B	8	*rpo*D (5) *rpo*H (2)	*arcA (1), argR (2), cpxR (1), lrp (2), glpR (1), fnr (1), crp (1), lexA (3)*	2	*rpo*D (1)	*fur (1)*
*cdt*C	8	*rpo*D (7)	*arcA (3), gcvA (1), argR (1), ihf (6), hns (1), crp (2), tyrR (1), lexA (1)*	1	*rpo*D (1)	*fis (1), cpxR (1)*
*omp*A	7	*rpo*D (5)	*oxyR (1), metR (1), lrp (2), argR2 (2), fis (1), ihf (1), fnr (1)*	1	*rpo*D (1)	*ihf (1), argR2 (1), fnr (1)*
*fha*B	7	*rpo*D (6) *rpo*H (1)	*purR (2), hns (1), cpxR (1), arcA (3), fis (1), ihf (1), tyrR (1), argR (1), argR2 (1), crp (5)*	2	*rpo*D (2)	*arcA (1), fnr (1), argR (1), hns (1)*
*avx*A	7	*rpo*D (11) *rpo*H (1)	*lrp (6), argR (3), fis (2), glpR (1), crp (1), argR2*	1	*rpo*D (2)	*fis (1)*

			(3), ihf (1), metR (3), lexA (1)			
*hmtp*210	7	*rpo*D (6)	*argR2 (3), lrp (4), cpxR (1), fnr (3), argR (3), arcA (2), purR (1), ihf (2), tyrR (1), glpR (1)*	1	*rpo*D (2)	*fnr (1), argR2 (1), argR (1), ihf (1), arcA (1)*

9.6 Identificación y representación de posibles secuencias canónicas de secuencias blanco de factores sigma codificados en el genoma de AVPG2015.

Basándose únicamente en los tres factores sigma anotados para AVPG2015 se hizo el conteo de los genes que son regulados por estos. Se predijeron 1644 genes regulados por el factor sigma 70, 245 genes regulados por sigma H y 13 genes regulados por sigma E, para un total de 1901 promotores de un total global de 2490 genes; esto significa que los 591 restantes son regulados por promotores distantes al ORF o por otro medio no predecible por la herramienta BPROM, de acuerdo a la estrategia de búsqueda con 500 bases del 5´. Los porcentajes predichos de genes controlados por factores sigma los podemos ver en la gráfica 2: sigma 70 reconoce el 66.02% de secuencias, sigma 32 reconoce el 9.83%, sigma E reconoce el 0.522% y el 23.73% de genes de AVPG2015 tienen otro tipo de regulación, o su promotor no se ubica en las distancias analizadas.

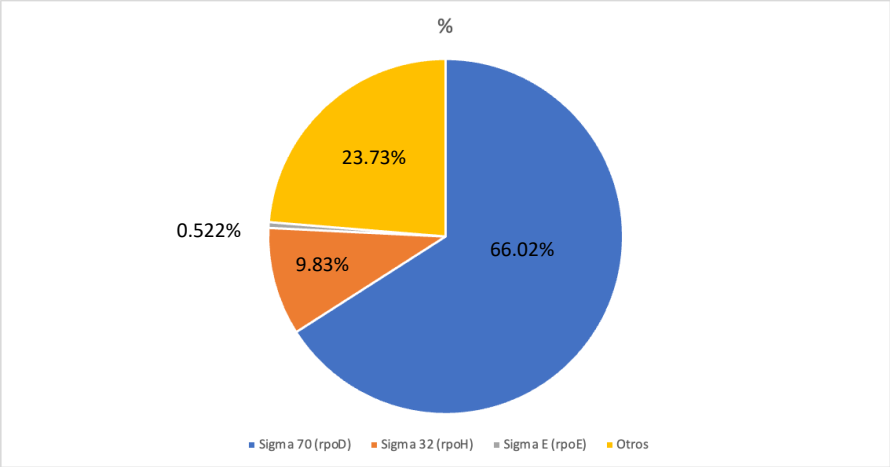


Gráfico 2: Porcentaje de genes regulados por los factores Sigma codificados en el genoma de AVPG2015.

La herramienta BPROM además nos indicó que puede haber isoformas de reconocimiento del factor sigma 70 o sitios para RpoD. El programa ofrece 5 versiones de secuencias blanco llamadas *rpo*D15, *rpo*D16, *rpo*D17, *rpo*D18 y *rpo*D19 las cuales se buscaron en las secuencias del DNA de AVPG2015. Los resultados se muestran en la tabla 11, observándose que la secuencia en *cis* reconocida con más frecuencia es para *rpo*D17 (5' MOTIVO 3') seguida de *rpo*D16 (5' MOTIVO 3') y con menor frecuencia la de *rpo*D19 (5' MOTIVO 3').

Tabla 12. Clases de probables secuencias reconocidas por Sigma 70.

Tipo de Secuencias reconocidas por Sigma 70 o *rpo*D	Número de Secuencias existentes
*rpo*D15	376
*rpo*D16	551
*rpo*D17	971
*rpo*D18	357
*rpo*D19	199

La base de datos de BPROM tiene registro para dos motivos de reconocimiento para *rpoH*, en las versiones *rpoH2* y *rpoH3*. La predicción de estas secuencias en el genoma de *A. paragallinarum* 2015 se muestra en la tabla 13, donde de observa que la más frecuente es *rpo*H2 con 156 secuencias blanco y *rpo*H3 sólo con 100.

Tabla 13. Clases de probables secuencias reconocidas por Sigma 32.

Tipo de Secuencias reconocidas por Sigma 32 o *rpo*H	Número de Secuencias existentes
*rpo*H2	156
*rpo*H3	100

Las secuencias de reconocimiento predichas para cada uno de los tres factores sigma se separaron de acuerdo con el factor sigma participante en la transcripción. Se alinearon las secuencias con Clustal Omega y se representó la secuencia consenso por medio de la herramienta WebLogo. Los consensos para sigma 70 y las versiones que tiene BPROM se muestran en la imagen 2; para el consenso de los promotores del tipo *rpo*D16 y *rpo*D17, que

fueron los más abundantes en la predicción con BPROM, se observa que los nucleótidos de adenina y timina (A y T) son muy frecuentes en una diada con la primera mitad del consenso rica en T y la segunda rica en A.

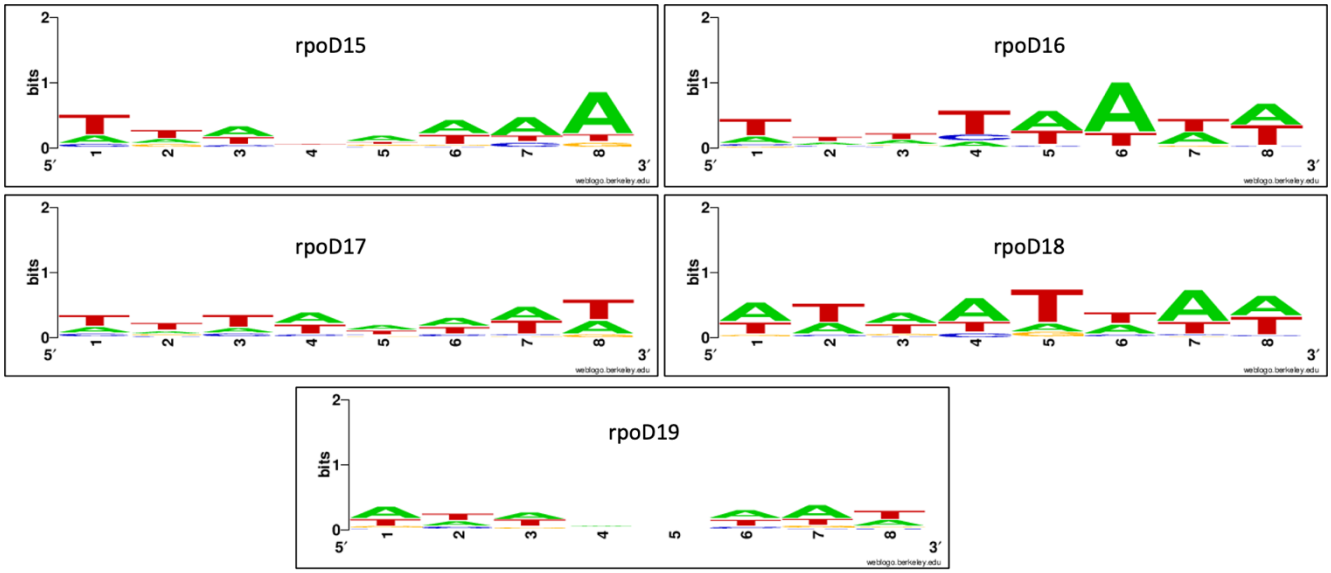


Imagen 2. Logo de las secuencias canónicas consenso reconocidas por la proteína sigma 70.

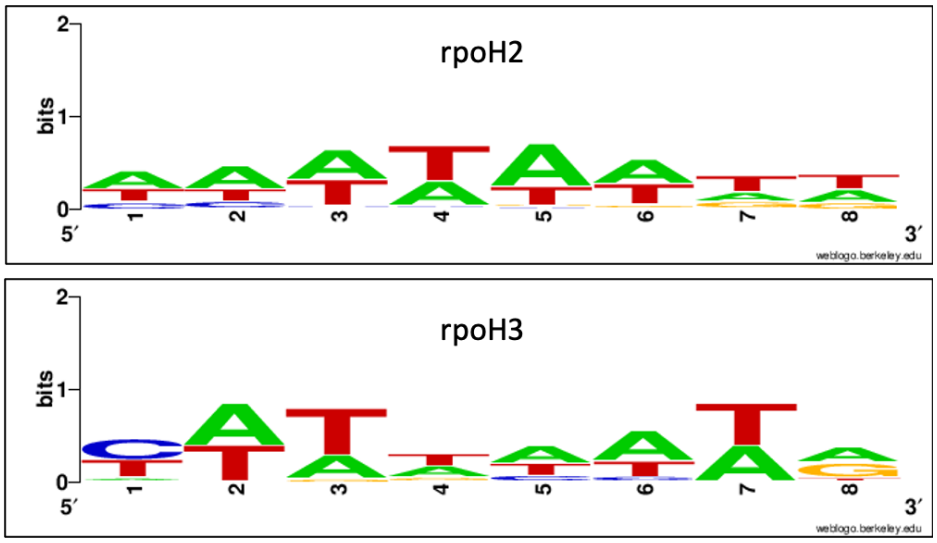


Imagen 3. Logo de las secuencias canónicas consenso reconocidas por la proteína sigma 32.

Las dos versiones de la secuencia de reconocimiento para RpoH, llamadas *rpo*H2 y *rpo*H3, tienen alto contenido de A y T, siendo más abundante la presencia de A; las cuatro bases centrales parecen mostrar una diada como la de *rpo*D, pero más corta con la presencia de G

o C en la parte más externa de la secuencia de reconocimiento. El resultado se muestra en la imagen 3.

La versión del consenso de reconocimiento para RpoE se muestra en la imagen 4. Este consenso se obtuvo sólo con 13 secuencias y se puede observar en la región central mayor frecuencia de T y una frecuencia más variada de otros nucleótidos en otras posiciones (A, G y C).

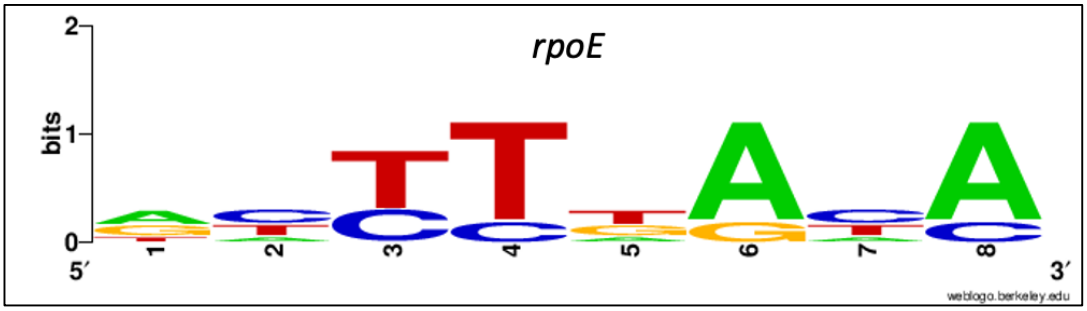


Imagen 4. Logo de las secuencias canónicas consenso reconocidas por la proteína sigma 24.

9.7 Unidades Transcripcionales que conforman el genoma de AVPG2015

Para predecir si hay presencia de unidades transcripcionales monocistrónicas y policistrónicas se cotejaron las predicciones de BPROM en segmentos de DNA de 500 bases en comparación con las de segmentos de 120,000 bases. En la tabla 14 se puede observar la frecuencia con la que se presentan las secuencias concenso predichas para cada regulador transcripcional en tramos de 12 kpb y 500 pb de DNA genómico. Se propone que en la región de 500 bases se predicen las unidades monocistrónicas por la cercanía, en tanto que en la región larga se predicen las policistrónicas u operones de acuerdo con cada regulador transcripcional. Se puede observar que el numero siempre es menor para tramos de 500 bases, lo que significa que se pierden sitios predichos al considerar la rigidez de 500 bases río arriba de cada ORF y puede haber reguladores en segmentos más largos por ejemplo considerando ORFS de longitudes mayores a 500 bases.

Tabla 14. Predicción de secuencias de reconocimiento por factores sigma y reguladores transcripcionales.

Genes reguladores	# de Sec (120,000 pb)	# de Sec (500 pb)	Genes reguladores	# de Sec (120,000 pb)	# de Sec (500 pb)
ada	28	17	*marR*	23	14
crp	860	468	*melR*	19	14
fis	623	416	*metJ*	149	78
fnr	500	316	*metR*	369	245
fur	272	160	*modE*	22	15
hns	218	138	*nagC*	403	268
ihf	1223	895	*narL*	82	50
lrp	1068	802	*narP*	24	15
tus	81	59	*ntrC*	16	6
araC	91	76	*ompR*	308	182
arcA	691	502	*oxyR*	91	59
argR	1097	810	*pdhR*	57	37
carP	120	80	*phoB*	543	343
cpxR	266	167	*purR*	461	283
cspA	5	4	*rhaS*	2	0
cynR	67	25	*rpoE*	15	13
cysB	53	29	*rpoN*	36	26
cytR	32	20	*soxS*	321	199
deoR	208	132	*torR*	10	7
dnaA	87	46	*trpR*	24	20
fadR	16	12	*tyrR*	382	265
farR	31	16	*argR2*	768	549
fhlA	57	36	*flhCD*	44	30
fruR	2	0	*phoB3*	6	4
galR	15	6	*rpoH2*	245	156
gcvA	126	72	*rpoH3*	154	100
glpR	156	86	*rpoD15*	606	376
hipB	68	46	*rpoD16*	907	551

ilvY	11	9	*rpoD17*	2033	971
lacI	15	12	*rpoD18*	579	357
lexA	698	466	*rpoD19*	297	199
malT	5	3	*rpoS17*	135	103
			rpoS18	18	6

Las probables unidades mono o poilicistrónicas deducidas con BPROM se comprobaron con la organización de los genes de todo el registro CP058307, y el último método fue utilizando la plataforma FGENESB de Softberry (80).

Como resultado de la predicción basándose en la orientación de los genes, se dedujo que, el registro CP058307 está conformado en su mayor parte por unidades monocistrónicas (318 de 823 = 38.63%), bicistrónicas (170 de 823 = 20.65%), tricistrónicas (109 de 823 = 13.24%) y tetracistrónicas (86 de 823 = 10.44%), los genes restantes se encuentran organizados en unidades policistrónicas.

Como resultado de la búsqueda basada en la información predictiva de 500 pb, se dedujo que, el registro CP058307 del mismo modo, está conformado en su mayor parte por unidades monocistrónicas (720 de 874 = 82.37%), bicistrónicas (301 de 874 = 34.43%), tricistrónicas (155 de 874 = 17.73%) y tetracistrónicas (98 de 874 = 11.21%), los genes restantes se encuentran organizados en unidades policistrónicas en cantidades menores.

Para finalizar, los resultados pertenecientes a la predicción basada en el programa FGENESB, se dedujo que, el registro CP058307, está conformado por unidades monocistrónicas (394 de 977 = 40%), bicistrónicas (218 de 977 = 22%), tricistrónicas (109 de 977 = 15%) y tetracistrónicas (86 de 977 = 11%), el resto de los genes se encuentran organizados en unidades policistrónicas, los resultados obtenidos de todos los métodos utilizados se pueden analizar en la tabla 15 y gráfico 3.

Tabla 15. Organización transcripcional de los genes de AVPG2015

Arreglos Transcripcionales	Cantidad de unidades transcripcionales por mapa físico	Cantidad de unidades transcripcionales de acuerdo a las secuencias reguladoras deducidas	Cantidad de unidades transcripcionales de acuerdo a la predicción con FGENESB
Monocistrónicos	318	720	394
Bicistrónicos	170	301	218
Tricistrónicos	109	155	146
Tetracistrónicos	86	98	101
Pentacistrónicos	47	52	65
Hexacistrónicos	17	1	8
Heptacistrónicos	15	0	13
Octacistrónicos	12	2	17
Nonacistrónicos	8	0	3
Decacistrónicos	11	1	3
Undecacistrónicos	7	0	2
Dodecacistrónicos	6	0	2
Policistrónicos	17	0	5

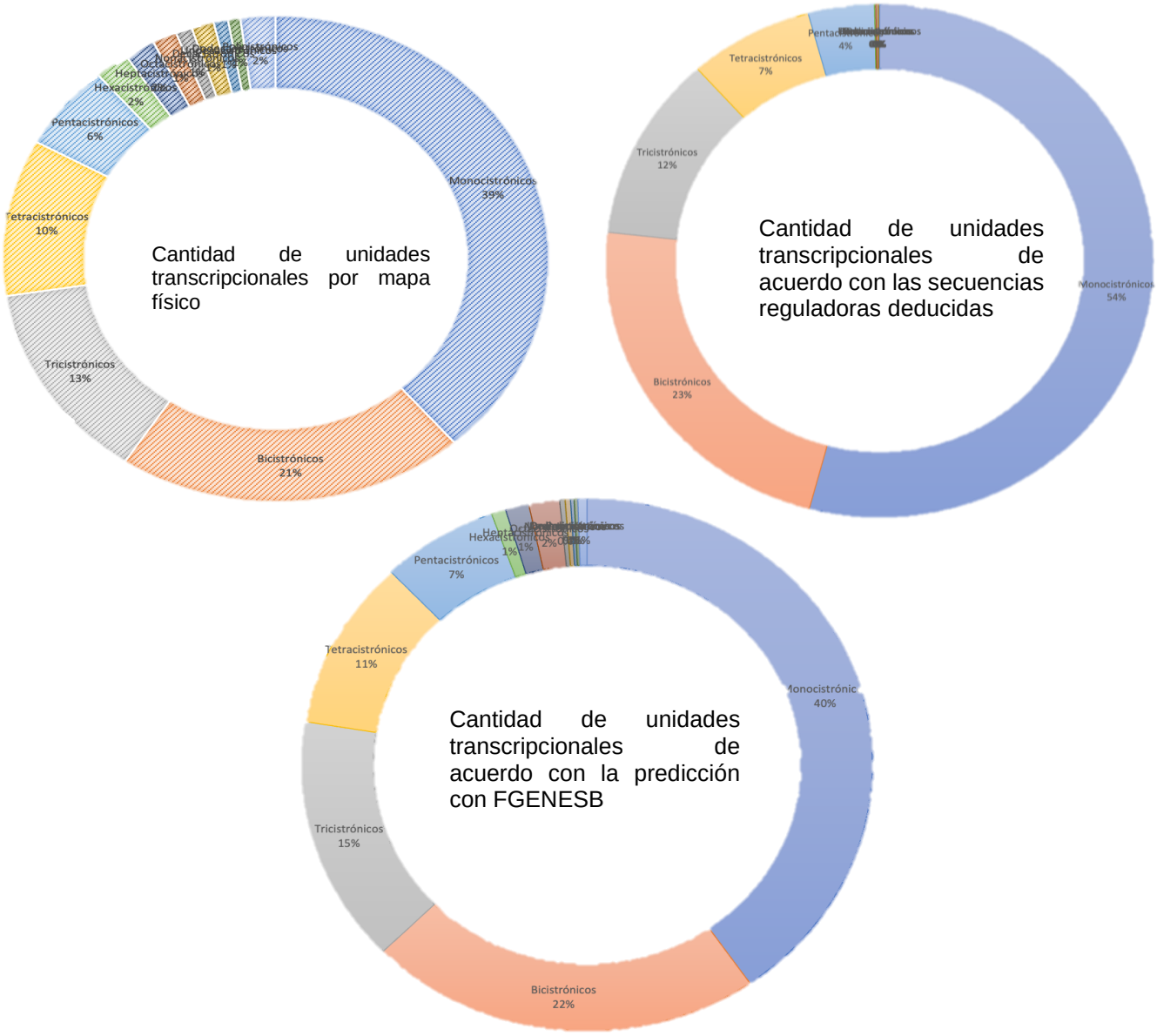


Gráfico 3. Frecuencia de unidades transcripcionales en el genoma de AVPG2015. Cada información representada en el gráfico de pastel fue obtenida por diferentes métodos: a) Arreglo cromosomal, b) Tipo de promotor, c) FGENESB.

Todo este análisis que representa un panorama general se realizó para todo el genoma de AVPG2015, aplicado para los 2490 genes que contiene el registro CP058407, pero del mismo modo se puede aplicar un análisis específico para cada uno de los genes antes mencionados,

y, partiendo del uso del análisis de los genes de condroitin liasas además del uso de los genes del proyecto qRT-PCR, podemos concluir lo siguiente:

Tabla 16. Organización de las probables unidades transcripcionales de 13 genes de AVPG2015.

Gen	Longitud de la organización-genes con igual sentido	Posición del gen	Regulador principal probable
*chl*1	2	2	RpoD
*chl*2	6	3	Lrp
*rpo*B	2	2	RpoD
*gyr*A	2	1	RpoD
*rec*N	2	2	RpoD
*atp*D	19-(9)	10-(7)	RpoD
fur	4	4	Fur
*fha*B	2	2	RpoD
*pep*N	6	1	RpoD
*omp*A	5	1	RpoD
*cdt*A	3	3	RpoD
*cdt*B	3	2	RpoD
*cdt*C	3	1	RpoD

Se puede deducir si un determinado gen puede formar parte de una probable unidad monocistrónica o policistrónica, así como predecir la posición de dicho gen en esa unidad, y el regulador causante de la expresión de los genes de dicha unidad.

Para el caso de los genes de Condroitin Liasa, primero hablando del gen *chl1,* este forma parte de una probable unidad bicistrónica, en la cual, ocupa la segunda posición de dicha unidad y Sigma 70 es el probable regulador responsable de su expresión. El gen *chl2,* forma parte de una probable unidad hexacistrónica, en la cual, ocupa la tercera posición de la unidad transcripcional y la proteína Lrp es el posible regulador responsable de su expresión.

9.8 Potenciales secuencias reguladoras en cis de los genes *chl1 y chl2*

9.8.1 Análisis de las secuencias cis en la región líder de *chl*1

Al haber predicho con el 80% de confianza los sitios de reconocimiento de factores sigma y de proteínas reguladoras de la transcripción que pudieran controlar la expresión de 2490 genes en *A. paragallinarum* de forma homóloga y en *E. coli* de forma heteróloga, sabemos que existen diferencias por el fondo genético usado en la investigación. También es conocido que la predicción de potenciales promotores pudiera variar dependiendo de la secuencia usada para llevar a cabo la búsqueda de secuencias *cis* controladoras de la expresión a nivel transcripcional. Para los genes *chl*1 y *chl*2 se tomaron dos tamaños de secuencia 5´ líder de los ORF para la búsqueda: 500 bases cercanas al ORF de cada gen y 3000 bases rio arriba incluyendo al ORF de los genes *chl* (6000 en total).

En la secuencia de 500 bases de *chl*1, se predijeron 2 posibles secuencias promotoras, en el promotor número uno, existe un blanco de reconocimiento por Sigma 70 y 2 blancos de reconocimiento para el regulador FNR. En el segundo promotor, se observan 3 potenciales sitios de reconocimiento, para 3 reguladores transcripcionales TyrR, Ihf y ArcA (imagen 5).

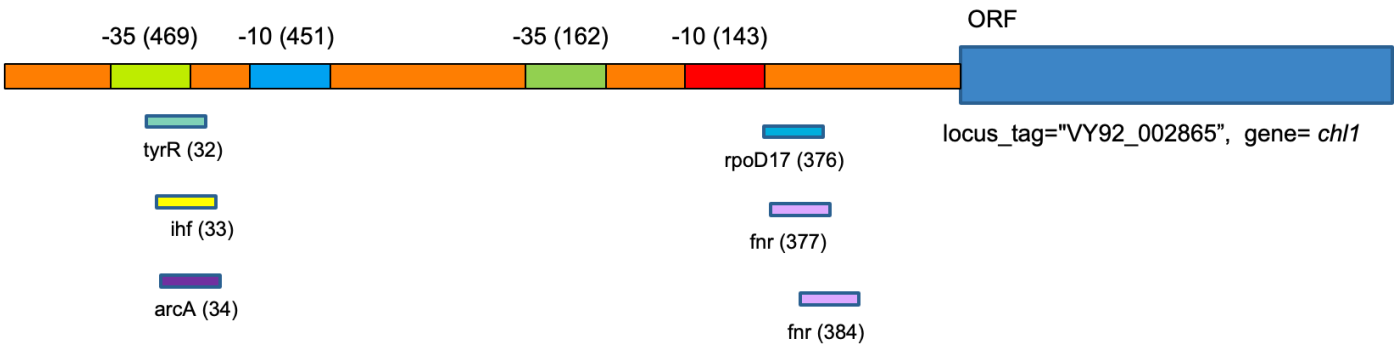


Imagen 5. Región 5´ rio arriba del marco de lectura abierto del gen *chl*1. Secuencia de 500 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción.

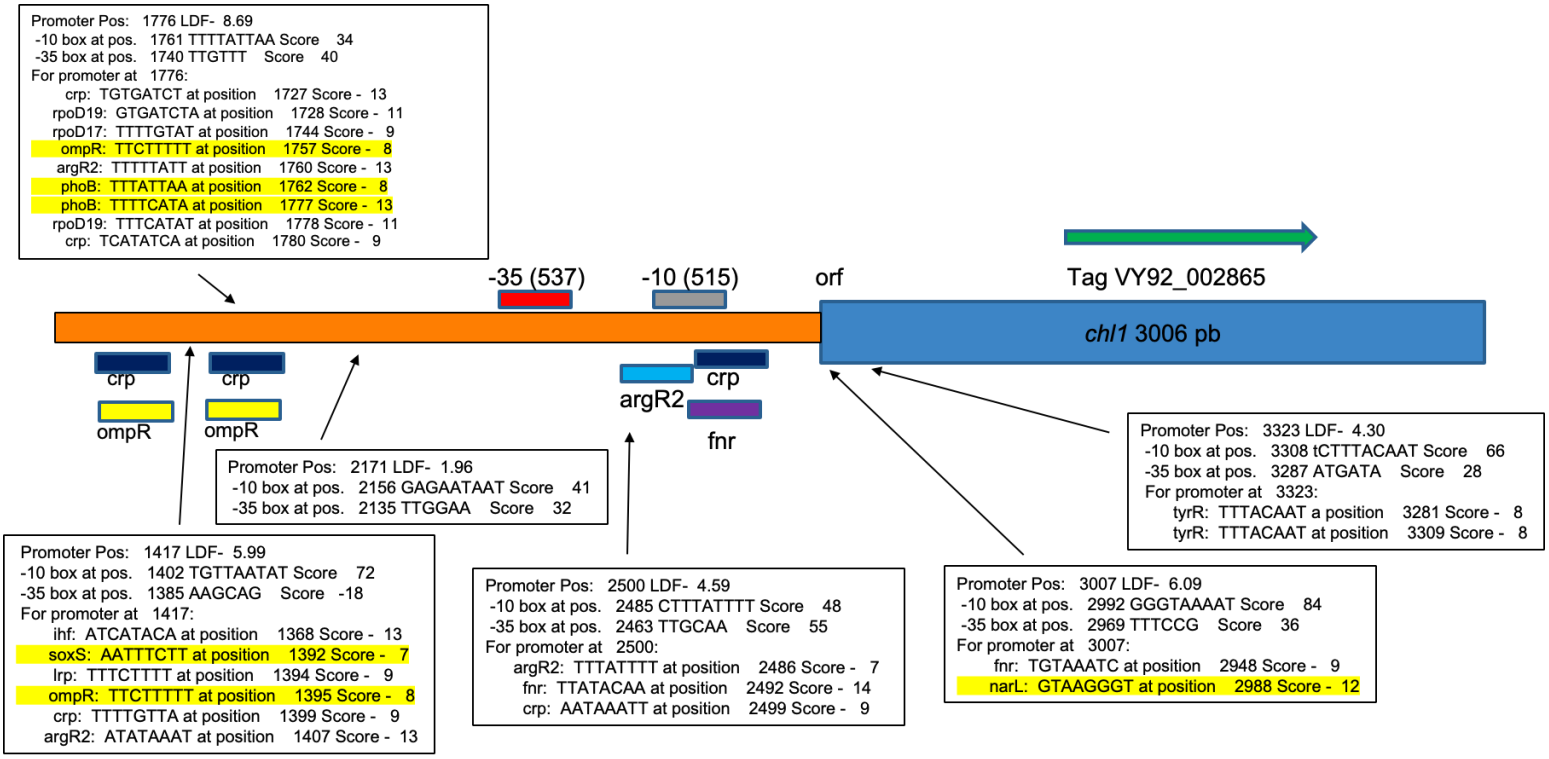


Imagen 6: Región 5´ rio arriba del ORF del gen *chl*1. Secuencia de 3000 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción. Los numeros en las cajas representan las coordenadas desde el inicio de la secuencia en el lado 5´, en tanto que las coordenadas para las secuencias marcadas como -515 y -537 representan las coordenadas rio arriba del ORF. Los genes resaltados en color amarillo corresponden a las secuencias blanco que pueden ser reconocidas en *E. coli*.

La búsqueda sobre la secuencia de 3000 bases de la región líder 5´mostró que hay un probable promotor cercano al ORF más allá de las 500 bases, en 515 y 537 bases. Se observaron 6 probables secuencias promotoras, 2 de las cuales son descartables poque se encuentran dentro del ORF del gen *chl*1 (imagen 6). Se muestra el promotor más cercano al ORF entre las bases 515 y 537 y los 7 sitios de unión de los reguladores transcripcionales; y para cada secuencia promotora, se muestran sus respectivos sitios probables de reconocimiento para reguladores transcripcionales, entre los que destacan reguladores como OmpR, PhoB, Crp, ArgR2, SoxS, Ihf, Lrp y Fnr. Los genes resaltados en color amarillo corresponden a las secuencias blanco que pueden ser reconocidas en *E. coli* pero no en *A. paragallinarum*.

9.8.2 Análisis de las secuencias *cis* en la región líder de *chl*2

Los promotores predichos en la región de 500 bases del gen *chl*2 potencialmente son dos, uno entre las bases -48 y -66 y otro entre las bases -355 y -378 y un sitio blanco para Lrp en -91 bases (imagen 7).

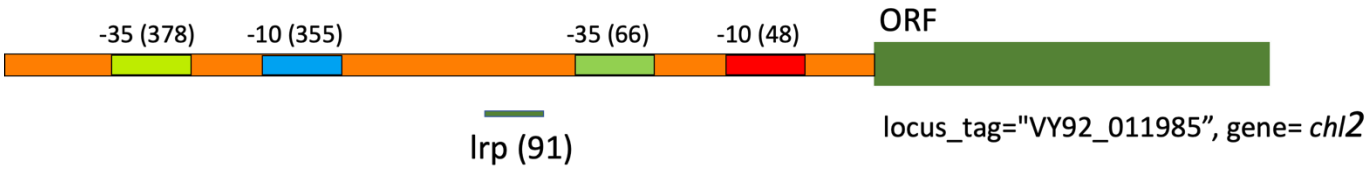


Imagen 7: Región 5´ rio arriba del ORF del gen *chl*2. Secuencia de 500 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción.

La predicción de promotores y secuencias reguladoras en *cis* usando el segmento de 300 bases cercanos al ORF mostró cinco promotores, dos de ellos descartables por estar dentro del ORF (Imagen 8). Se muestra el promotor más próximo a las 500 bases del ORF, localizado entre las posiciones -613 a -631 y los sitios blanco de reconocimiento por los reguladores transcripcionales Ihf, ArgR2 así como GcvA. En color amarillo se resaltan las secuencias blanco para ser reconocidas en *E. coli* pero no en *A. paragallinarum.*

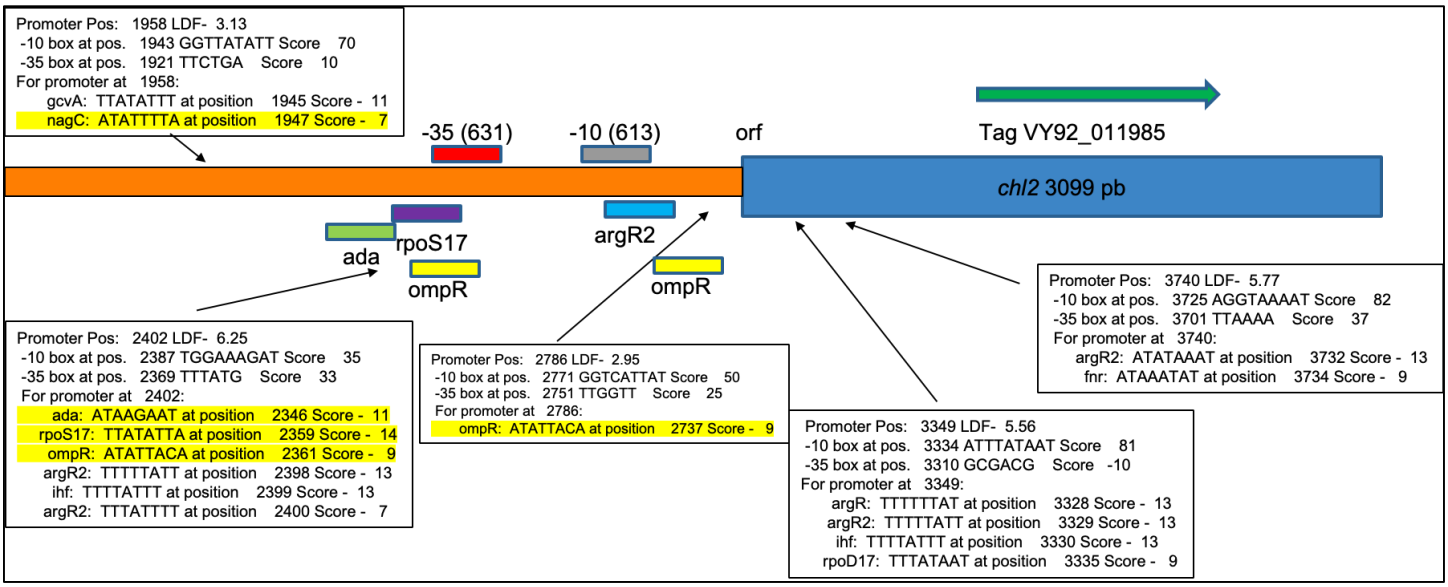


Imagen 8: Región 5´ rio arriba del ORF del gen *chl*2. Secuencia de 3000 pb en la que se muestran los posibles sitios de unión de proteínas reguladoras de la transcripción. Los numeros en las cajas representan las coordenadas desde el inicio de la secuencia en el lado 5´, en tanto que las coordenadas para las secuencias marcadas como -515 y -537 representan las coordenadas rio arriba del ORF. En color amarillo se resaltan las secuencias blanco que pueden ser reconocidas en *E. coli*.

9.9 Redes de regulación transcripcionales predichas para los genes de Condroitín Liasas

Analizando toda la información redactada hasta el momento, los genes que codifican para las enzimas Chl están regulados por sigma 70 en regiones proximales (500 pb) y distales (3000 pb), entonces, para poder representar esto en un diagrama, se realizó la predicción de una red de regulación transcripcional del gen *rpo*D (el cual codifica para Sigma 70) basándose en la información predictiva obtenida de las bases de datos computacionales del programa String.

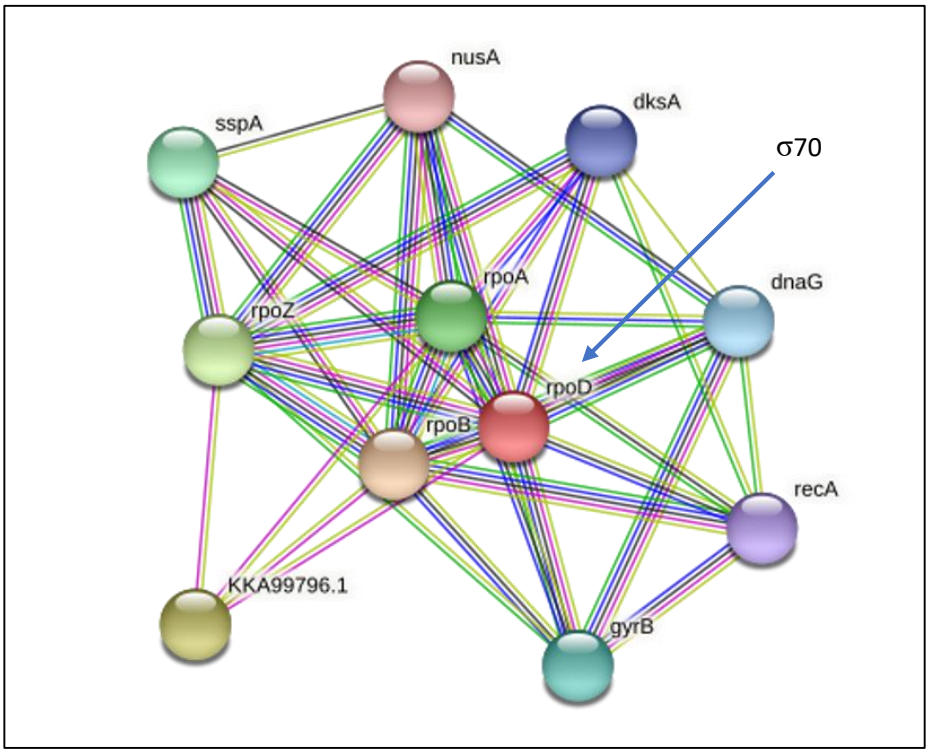


Imagen 9: Red de regulación transcripcional predicha de la proteína Sigma 70 por el programa bioinformático String

En la red de regulación de Sigma 70, esta interacciona con proteínas que van desde las subunidades del core de la ARN polimerasa hasta proteínas que puedan ser receptoras de AMP cíclico, u otras proteínas como la subunidad β de la DNA girasa; Muchas de estas interacciones se predicen con base a extracción de información experimental o extracción de bases de datos, o en algunos casos, algunos genes se co-expresan como el caso de RpoD y DnaG como se observa en la imagen 9.

Añadiendo la información de la proteína Chl1 a la red de regulación transcripcional de Sigma 70, como se observa en la imagen 10, se observa que Sigma 70 tiene un efecto directo en *chl1* ya que se favorece la expresión del gen, y a su vez interactúan directamente sobre *chl1*, *ihf*, *arcA*, *fnr* y *tyrR* actuando ya sea como inhibidores o activadores de la transcripción.

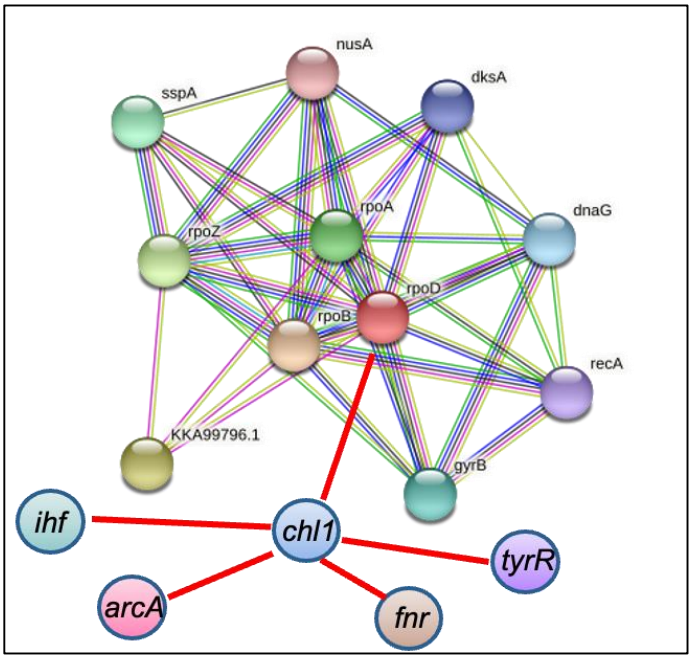


Imagen 10: Modelo propuesto de la red de regulación transcripcional que controla la expresion de la proteína Chl1. El modelo fue realizado superponiendo la información obtenida de la red de regulación de σ70 con la información predictiva obtenida de Chl1.

Para el caso de la proteína Chl2, no es regulada por Sigma 70 para el rango de 500 pb, con base en la predicción, ésta interacciona únicamente con la proteína Lrp, posiblemente como un inhibidor transcripcional. Esta información deberá ser confirmada experimentalmente,

siendo importante aclarar que esto es una predicción bioinformática, no una caracterización funcional experimental; en la imagen 11 se puede observar la red de regulación transcripcional correspondiente a la proteína Chl2.

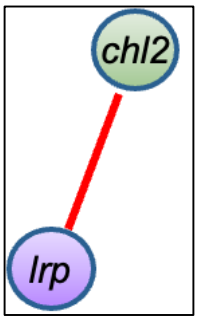


Imagen 11: Modelo propuesto de la red de regulación transcripcional que controla la expresión de la proteína Chl2. El modelo fue realizado superponiendo la información obtenida de la red de regulación de σ70 con la información predictiva obtenida de Chl1.

9.10 Predicción de estructura tridimensional de las proteínas Chl1 y Chl2

Como parte final de esta investigación, mediante el uso de tres programas bioinformáticos: Swiss Model, ITasser y PyMol, se predijeron estructuras tridimensionales probables para ambas proteínas Condroitin Liasas, obteniendo diversos resultados predictivos que le dan validez al modelo presentado comparado con modelos de proteínas cristalizados de Condroitin Liasas de otros microorganismos como *Proteus Vulgaris* y *Bacteroides thetaiotaomicron.*

Diferentes modelos generados con valores estadísticos se obtuvieron para cada proteína Chl, mediante el uso del programa Swiss model, el cual requiere un formato fasta para su utilización. En el caso de la proteína Chl1, dos posibles modelos tridimensionales se predijeron (imagen 12).

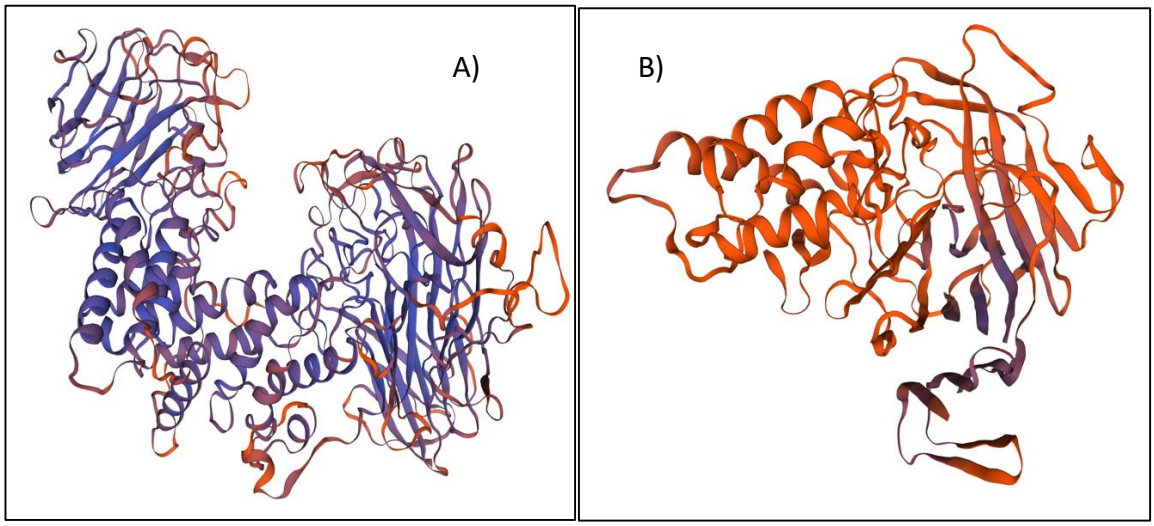


Imagen 12: Estructuras tridimensionales de la proteína Chl1. Modelos generados con el programa SwissModel. A) Chl1.M1, B) Chl1.M2.

Como se puede apreciar, el modelo 1 (Chl1.M1) es más exacto y con mayor calidad que el modelo 2 (Chl1.M2) debido a que la composición de aminoácidos del modelo 1 es más semejante a la de modelos cristalizados de proteínas Chl reportados en otras bases de datos. En caso del modelo 2, su estructura tridimensional no comparte semejanza con otras proteínas Chl reportadas. Esto es corroborado por los diferentes valores estadísticos obtenidos para los diferentes modelos, los cuales se muestran en la tabla 17.

De acuerdo a los diferentes valores de QMEAN y GMQE, podemos analizar que el modelo 1 predicho de la proteína Chl1 es más exacto y con una calidad mayor comparado con el modelo 2 de la misma proteína. No obstante, debido al valor negativo alejado de cero en ambos valores de QMEAN, podemos inferir que ambas estructuras difieren mucho en su estructura tridimensional con las estructuras cristalizadas almacenadas en la base de datos de SwissModel.

Tabla 17. Resultados estadísticos predictivos para los diferentes modelos propuestos de la proteína Chl1.

Valores estadísticos	Chl1.M1	Chl1.M2
QMEAN	-4.27	-7.59
GMQE	0.69	0.19
QMEANDisCo	0.69 ± 0.5	0.40 ± 0.05
Seq. Identidy	30.73	29.14

En el caso de la proteína Chl2, únicamente un posible modelo de estructura tridimensional se predijo. Sin embargo, basándonos en los resultados estadísticos de SwissModel y en comparación con los modelos tridimensionales predichos para la proteína Chl1, el modelo de la proteína Chl2 de AVPG 2015 sugiere que su estructura tridimensional es muy semejante a la de modelos cristalizados de proteínas Chl reportados en otras bases de datos. Lo cual es corroborado por los diferentes valores estadísticos obtenidos para los diferentes modelos (Tabla 18).

El modelo tridimensional posee resultados más confiables y de mayor calidad en el modelado, debido a que los resultados de QMEAN y GMQE abarcan valores aproximados o mayores de 1, lo que genera una mayor certeza que posee homología a otras proteínas Condroitin Liasas que se encuentran almacenadas en la base de datos de SwissModel. Los resultados se pueden observar en la tabla 18.

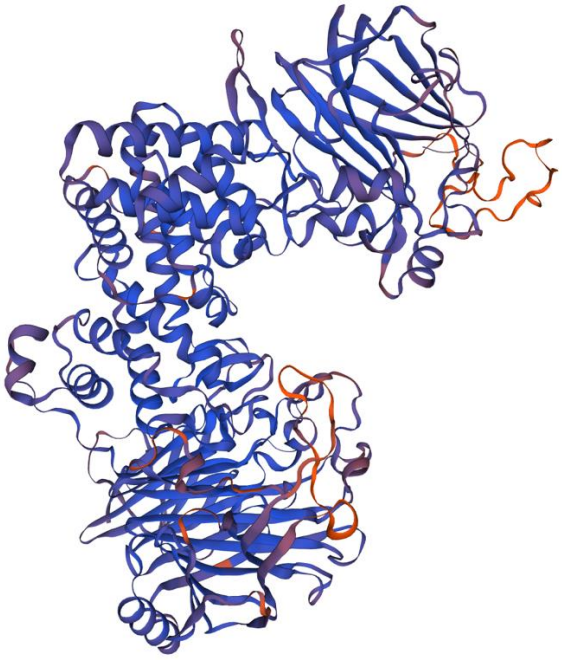


Imagen 13. Estructura tridimensional de la proteína Chl2. Modelo generado con el programa SwissModel.

Tabla 18. Resultado estadístico predictivo para el modelo propuesto de la proteína Chl2.

Valores estadísticos	Chl1.M1
QMEAN	-4.27
GMQE	0.69
QMEANDisCo	0.69 ± 0.5
Seq. Identidy	30.73

Con el uso de las secuencias de aminoácidos de ambas proteínas Chl del registro de Genbank CP058307, y de un segundo programa denominado I-Tasser, se obtuvieron múltiples modelos tridimensionales predichos con diferentes valores de C-Score, siendo el modelo resultante con un valor C-score más alto el más acertado a homología con otras proteínas cristalizadas reportadas previamente.

Para la proteína Chl1, se predijeron tres modelos tridimensionales los cuales varían en valor C-Score como se pueden ver a continuación en la imagen 14.

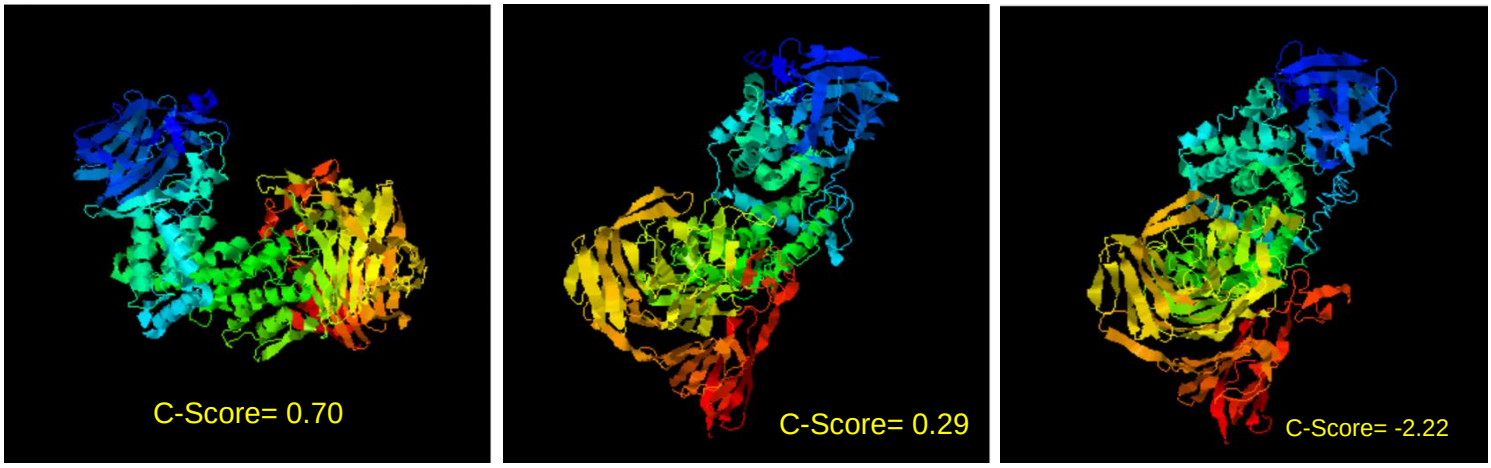


Imagen 14. Estructura de la enzima Chl1 de AVPG 2015. Modelos tridimensionales generados mediante el programa I-Tasser.

La estructura tridimensional, en este caso, más confiable basándonos en el valor C-Score, es el modelo cuyo valor es de 0.70 mientras que los demás, debido a su cercanía con el valor 0, resultan poco confiables de compartir homología con proteínas condroitin liasas cristalizadas y reportadas en otras bases de datos.

Para la proteína Chl2 se predijeron utilizando el mismo procedimiento, cuatro probables estructuras tridimensionales, con valores muy diferentes de c-score entre ellos. Se observan los diferentes modelos a continuación en la imagen 15.

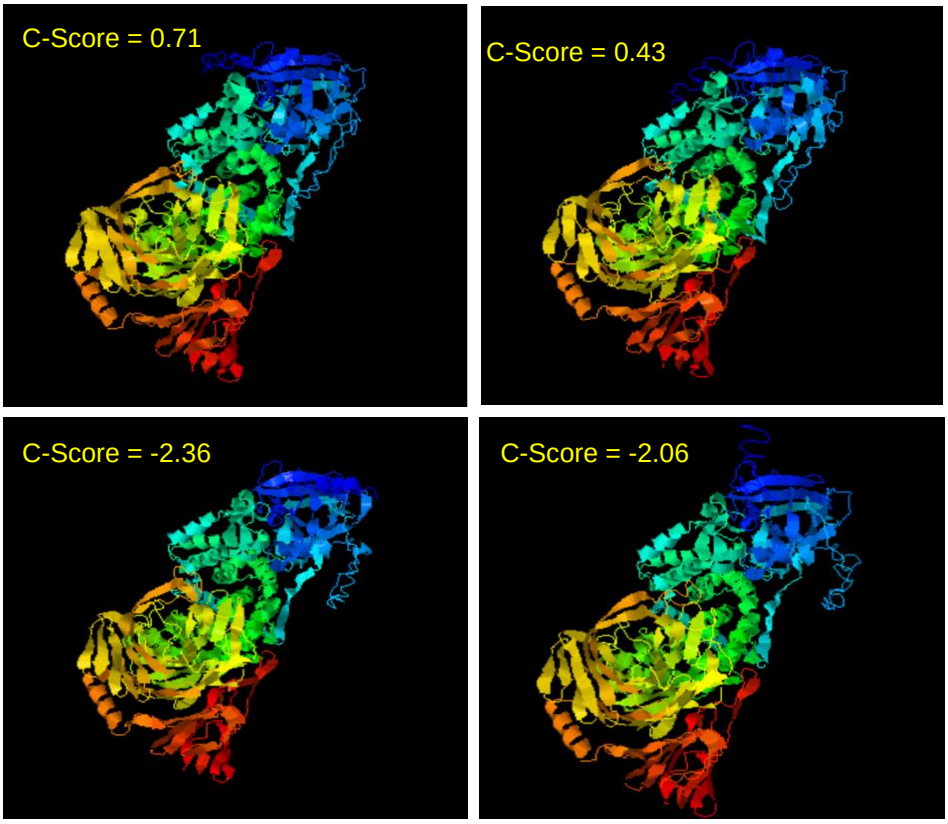


Imagen 15. Estructura de la enzima Chl2 de AVPG. Modelos tridimensionales generados mediante el programa I-Tasser.

Basándonos en el valor C-Score, el modelo tridimensional más confiable tiene un C-score de 0.71 y 0.43, los demás resultan poco confiables de compartir homología con proteínas condroitin liasas cristalizadas y reportadas en otras bases de datos debido a su aproximidad al valor 0.

Finalmente, utilizando un tercer programa bioinformático, PyMOL, se superpusieron los modelos cristalizados de proteínas condroitin liasas de *Proteus vulgaris* (Modelo 1HN0), *Bacteroides thetaiotaomicron* (Modelo 2Q1F) en la imagen 16.

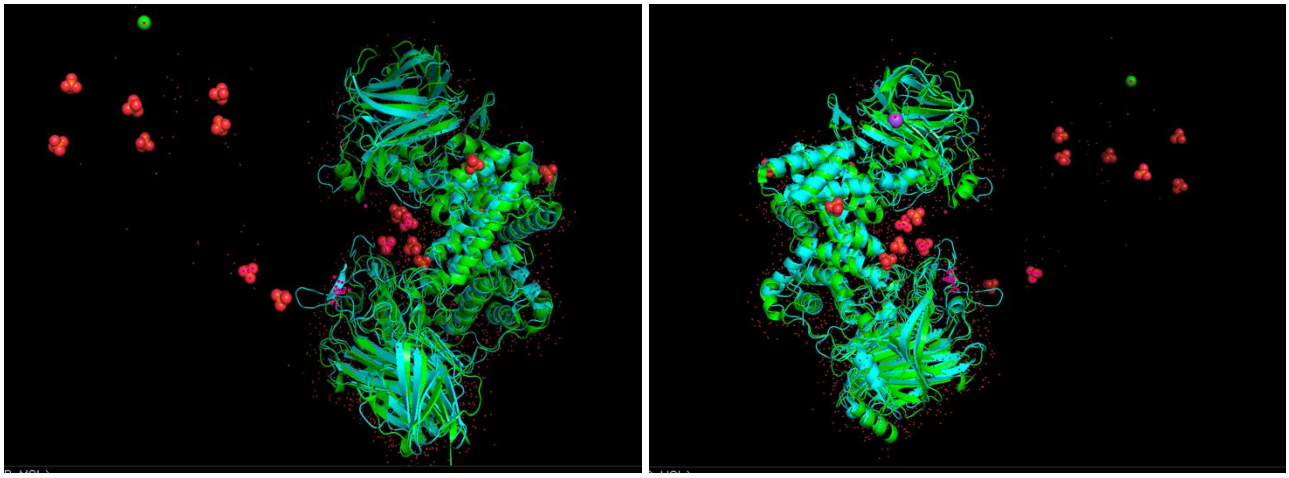


Imagen 16. Superposición de dos estructuras tridimensionales, obtenidas de dos cristales de condroitin liasas, la estructura verde corresponde a la proteína Chl de *Proteus vulgaris* y la estructura azul corresponde a la proteína Chl de *Bacteroides.sp* con programa PyMol.

Con base en esta suposición, notamos que ambas proteínas Chl comparten muchas similitudes entre si mismas en múltiples dominios, aunque con algunas diferencias en la estructura secundaria de la proteína variando entre hélices alfa y laminas beta. Utilizando la estructura de la misma proteína de *Proteus vulgaris* pero superponiendo la del modelo tridimensional de Chl1 predicho con I-Tasser se aprecia una mayor similitud entre ambas estructuras terciarias con respecto a las de la imagen 17.

Imagen 17. Superposición de dos estructuras tridimensionales utilizando el programa PyMol, la estructura verde corresponde a la proteína Chl de *Proteus vulgaris* y la estructura azul corresponde a la proteína Chl1 de AVPG 2015 predicha por medio de I-Tasser.

10. DISCUSIÓN

Avibacterium paragallinarum es un microorganismo patógeno que contiene factores de virulencia codificados en su genoma y que es de relevancia para la industria avícola debido a las grandes pérdidas económicas que ocasiona, debido a la disminución en la producción de huevo, y al alto consumo de medicamentos, entre otros. Los análisis bioinformáticos nos han permitido predecir nuevos elementos de regulación genética acerca de cómo se expresan diferentes genes y que elementos interactúan y afectan la expresión de un gen. *Avibacterium paragallinarum* 2015 contiene codificados en su genoma genes que contribuyen a la regulación transcripcional global de todos sus genes, en una exploración física del registro de Genbank CP058307 fueron encontrados siete segmentos de DNA que contienen genes de profago que codifican proteínas propias de fago y que no contribuyen a la regulación transcripcional global de AVPG 2015. Esto es apoyado igualmente por lo analizado por Horta-Valerdi *et al.* (82), los cuales a través de una secuenciación del genoma completo de la cepa CL de AVPG determinaron la presencia de varias regiones ricas en secuencias de inserción, de las cuales, 142 proteínas putativas fueron determinadas como transposasas e integrasas y 117 relacionadas con fagos, probablemente siendo parte del fago Mu con funciones vestigiales.

Además, fueron anotados todas las secuencias de AVPG2015 que actúan como reguladores transcripcionales, de los cuales se encuentran codificados tres factores Sigma (σ70, σ32 y σ24) así como cuatro proteínas que regulan los factores σ antes mencionados. De acuerdo con Paget *et., al* (33) la función de σ70 es regular la transcripción general de la mayoría de los genes mientras que los factores σ accesorios tienen la función de activar genes temporales o específicos en presencia de una señal característica, entre las cuales destacan 3 categorías: respuesta al estrés, desarrollo y metabolismo auxiliar; por ejemplo, se cree que σ32 (σH) regula la expresión de genes en respuesta a choque térmico mientras que σ24 (σE) modula la expresión de genes de desarrollo bacteriano.

A su vez, con base en lo establecido por Boyce *et al.,* (2002), los cuales confirman la presencia de los factores σ antes mencionados (σ70, σ32 y σ24) en *Pasteurella multocida*, se

predijo que las funciones en que participan los factores σ32 y σ24 estaban relacionadas en las respuestas de calor y choque extracitoplasmático respectivamente, podría dar sustento a los resultados obtenidos aquí, debido a que tanto *Pasterella multocida* como *Avibacterium paragallinarum,* pertenecen a la familia *Pasteurellaceae*.

Este resultado es apoyado por Craig *et al.,* (2002), quienes establecen que en el genoma de *Haemophilus influenzae* NTHi se encuentra un operón bacteriano que codifica para un factor σE, el cual controla la respuesta al estrés extracitoplasmático. Como se describió anteriormente, en diferentes miembros de la familia *Pasteurellaceae*, probablemente se encuentren codificados diversos factores σ, los cuales pudieran realizar la misma función en cada microorganismo, incluido *Avibacterium paragallinarum*. Esta suposición necesita ser confirmada ya que no se han realizado más estudios comparativos y que incluyan además a otras cepas de *Avibacterium paragallinarum.*

Taverner *et al.,* (2004) describió que para que un gen pueda expresarse adecuadamente, además de la correcta unión a la RNA polimerasa, se necesita de un control en el cual se puede llevar a cabo una activación o represión génica bajo diferentes condiciones, tarea que es llevada a cabo por proteínas que actúan como reguladores transcripcionales, las cuales se unen al ADN cumpliendo la función de activadores o inhibidores de genes, en este contexto, es posible identificar tanto *in-vitro* como *in-silico*. La determinación de los sitios de unión al DNA de dichos reguladores transcripcionales, *in-vitro*, incluyen metodologías como inmunoprecipitación de cromatina (ChIP) o identificación de DAM metilasa (DamID), lo que permite identificar dónde se une el regulador en el genoma. De manera *in-silico*, existen programas cuyos algoritmos de predicción han sido mejorados y, por lo tanto, son una buena herramienta que cumple su objetivo principal, aunque cabe destacar que siempre existirán los falsos positivos o negativos en este tipo de programas, debido a que los reguladores transcripcionales no siempre se vinculan a su sitio de enlace consenso, por lo que estas predicciones se deben corroborar de una manera experimental.

Thieffry *et al.,* (1998), realizaron un análisis para predecir sitios reguladores *cis* transcripcionales en el genoma de *E.coli* (cepa K-12), utilizando como fuente la información recopilada en RegulonDB. Ellos confirmaron la presencia de 56 reguladores transcripcionales codificados en el genoma de *E. coli*; el 75% de los sitios blanco de dichos reguladores fue confirmado siguiendo la información recopilada en la base de datos. Siendo

E. coli la bacteria más estudiada actualmente, y además un microorganismo Gram negativo como lo es *Avibacterium paragallinarum,* es probable que compartan en su genoma genes que codifiquen para los mismos reguladores transcripcionales. En el genoma de AVPG2015 se determinó la presencia de 27 reguladores transcripcionales, aunque al momento se desconoce si los sitios blanco de dichos reguladores son semejantes a lo establecido en *E. coli* K-12.

Para poder identificar los sitios y secuencias potencialmente reguladoras de la transcripción de los genes anotados en AVPG2015 se hizo uso del programa bioinformático BPROM. Esta metodología predictiva ha sido también empleada por Chantes-Guerra *et al.,* (2022) quienes identificaron un putativo receptor de sideróforos que se encuentra regulado por Fur, en presencia de un agente quelante de hierro en *Gallibacterium anatis;* lo que podría respaldar la utilidad/fortaleza de este tipo de herramientas bioinformáticas. BPROM fue también utilizada como herramienta predictiva para determinar las secuencias regulatorias de las secuencias de DNA del receptor de captación de hierro, además de los genes *fur.* En base a esos resultados, en AVPG, con el uso de esta herramienta, fueron predichos todos los sitios de reconocimiento de DNA que reconocen cada regulador transcripcional mencionado anteriormente.

En segmentos previos, se ha descrito acerca de los potenciales factores σ, así como reguladores transcripcionales que regulan no solo a los genes *chl1* y *chl2,* aplicando la predicción a dos regiones tanto proximal (500 pb) como distal (6000 pb). Con el fin de corroborar y validar las predicciones realizadas en BPROM, se realizaron las mismas predicciones en genes específicos de AVPG2015, específicamente en los genes de interés del proyecto de Ramos (Ramos Salas, manuscrito en preparación), quien realizó experimentos de expresión diferencial (qRT-PCR), de genes de relevancia para *Avibacterium paragallinarum*. Estos genes incluyeron 4 genes housekeeping (*gyrA, atpD*, *rpoB* y *recN*) y 9 genes que codifican factores de virulencia (*fur*, *pepN*, *cdtA*, *cdtB*, *cdtC*, *ompA*, *fhaB*, *avxA* y *hmtp210*). Para todos estos genes se realizó la misma predicción en BPROM utilizando los mismos rangos de búsqueda; es importante mencionar que, debido al gran volumen de información predicho por cada gen, se generó una tabla comparativa (tablas 10, 11) en la que se ve resumida toda la información mencionada hasta el momento y en donde además se incluye la información resumida de los genes *chl1* y *chl2* en la tabla 9. Para poder ejemplificar

lo descrito anteriormente, se tomó el gen *pep*N como referencia; los factores σ predichos para este gen, fueron *rpo*D, y *rpo*H. Para esta proteína, otros posibles reguladores identificados son *fis*, *crp*, *arc*A, en los cuales existen dos sistemas de regulación, uno donde se va a acelerar la velocidad de transcripción (Crp) y otro que puede inhibir la transcripción (ArcA).
Un trabajo similar al descrito aquí fue realizado por Perez-Rueda, (2000) quien realizó una búsqueda detallada en la literatura y usando análisis predictivos en diferentes medios informáticos, definió la presencia de un total de 314 proteínas reguladoras de la transcripción en *Escherichia coli* (cepa K-12), pero sin haberlo confirmado experimentalmente. El también determino la posible función de un regulador (activador o inhibidor) mediante el criterio de posición del dominio HTH en la secuencia de la proteína. La predicción de la probable función de reguladores predichos para un gen genera un mayor panorama del entorno regulatorio de genes específicos con funciones variables y podría facilitar el estudio/caracterización de esos genes.

Además de haber podido predecir los posibles elementos reguladores de control genético con las que cuenta AVPG y los sitios de DNA predichos reconocidos por dichos elementos regulatorios, también se dedujeron las secuencias consenso y se obtuvo el logo canónico de las mismas, siendo en la mayoría de los casos, secuencias con un contenido rico en A-T. Los resultados del logo canónico de AVPG concuerdan con lo reportado por Weindl *et al.,* (2007) quienes realizaron un modelo computacional de *E. coli* evaluando la energía de unión entre σ70 y la región promotora. Ellos determinaron que las secuencias del DNA reconocidas por σ70 están conformados regiones ricas en A-T. También, utilizando un algoritmo de la matriz derivada de los promotores de σ70, determinaron una cantidad limitada de secuencias blanco en el DNA que son reconocidas por factores σ accesorios: para σ19 fueron reconocidas 2 secuencias, para σ24 fueron reconocidas 61 secuencias, para σ28 fueron reconocidas 20 secuencias blanco y σ32 se identificaron 33 secuencias, y, se espera que compartan una semejanza con respecto a las secuencias blanco que son reconocidas por σ70. Es probable que AVPG2015 comparta semejanzas con las secuencias blanco-reconocidas por los factores σ de *E. coli.*
En base a los resultados obtenidos en el presente trabajo, podemos concluir que el genoma de AVPG2015 se encuentra organizado transcripcionalmente en su mayoría por unidades de transcripción monocistrónicas, bicistrónicas y tricistrónicas. Esto es semejante a lo

establecido por Cho *et al.*, (2009) quienes determinaron la arquitectura del genoma de *Escherichia coli* (cepa K-12 MG1655) identificando regiones de unión a ARN polimerasa, nivel de expresión de transcripciones de ARNm, sitios de inicio de transcripción, así como cuantos ORF codificantes de proteínas potenciales (pORF) hay dentro de cada segmento de transcripción guiado por RNAP y finalmente determinaron la arquitectura de unidades de transcripción de todo el genoma. Ellos determinaron que el genoma de *E. coli* K-12 MG1655 está constituido de 4,661 unidades de transcripción, siendo 3,010 unidades catalogadas como monocistrónicas y 1,652 unidades catalogadas como policistrónicas.

En base en lo anterior, y debido a la carencia de información detallada de otros microorganismos, se puede establecer la hipótesis de que, en microorganismos Gram negativos, gamma proteobacterias, como es el caso de *E. coli* y *Avibacterium paragallinarum*, su genoma está organizado en su mayoría por unidades monocistrónicas, representado esto el 40% o más de todo el genoma, mientras que las unidades policistrónicas representan una menor proporción en el genoma de los microorganismos.

11. CONCLUSIÓN

El genoma de *A. paragallinarum* 2015 tiene una estructura con distribución polar, en la que se observa en un lado del genoma la presencia de profagos y codificación de factores σ. El DNA viral corresponde al 12.19% y esta información no contribuye con los genes que codifican los factores σ y factores transcripcionales convencionales.

En el genoma de AVPG2015 existen secuencias que codifican tres factores Sigma (σ70, σ32 y σ24) siendo σ70 el que regula gran parte de los genes que constituyen el genoma. Adicionalmente, cuatro proteínas regulan a los factores σ (tres proteínas de regulación de σ24 y un regulador transcripcional de σ54). Hay vestigios de que σ54 se perdió en este microorganismo.

En el genoma de AVPG existen 90.1% de secuencias blanco para RpoD (σ^{70}), 9.4% para RpoH (σ^{32}) y 0.5% para σ^{24}. La principal variación de estas secuencias depende del contenido de T y A (%), la amplitud de la secuencia rica en A y T consiste de 4 a 8 bases.

En el genoma de *A. paragallinarum* 2015 se identificaron 27 reguladores transcripcionales, mientras que en el sistema heterólogo de *E. coli* pueden reconocerse hasta 54 reguladores. Las secuencias más frecuentes, en el reconocimiento transcripcional, corresponden a las de σ^{70}, ArgR y ArcA con el 25.7% del total de secuencias identificadas.

El gen *chl1* en regiones cortas (500 pb) y largas (3000 pb) es regulado por σ70 (*rpoD17*). Por la cercanía de las secuencias regulatorias del gen *chl*1, al inicio del orf, parece estar este más estrechamente controlado que *chl2*, de este último no se pudo predecir consistentemente alguna regulación transcripcional.

Basándose en el análisis de unidades transcripcionales de AVPG2015, su genoma está organizado en su mayoría por unidades monocistrónicas, bicistrónicas y tricistrónicas, pudiendo formar parte de probables operones o genes con regulación génica individual.
El gen *chl1* probablemente forme parte de una unidad bicistrónica, probablemente regulada por σ70, mientras que el gen *chl2* pudiese formar parte de una unidad hexacistrónica, siendo probablemente regulada por la proteína Lrp.

12. PERSPECTIVAS

- Predicción de probables interacciones de las proteinas reguladoras de AVPG.
- Predicción de las proteínas reguladoras mas importantes para otros genes diana de AVPG2015.
- Explorar modelos tridimensionales de las proteínas reguladoras presentes en el genoma de AVPG2015.
- Realizar la amplificación de redes de regulación con el programa String
- Identificar de manerae experimental las proteínas reguladoras de la transcripción de los genes *chl1 y chl2*.

REFERENCIAS BIBLIOGRÁFICAS

1. Wahyuni, Agnesia Endang, *et al*. “Isolation, Identification, and Serotyping of *Avibacterium Paragallinarum* from Quails in Indonesia with Typical Infectious

Coryza Disease Symptoms." *Veterinary World*, vol. 11, no. 4, 2018, pp. 519–524., https://doi.org/10.14202/vetworld.2018.519-524.

2. Soriano Vargas, Edgardo. "*Haemophilus paragallinarum*: Etiology of Infectious Coryza." *Veterinaria México*, 3 July 2004, https://www.redalyc.org/pdf/423/42335308.pdf.
3. Colas Chavez, M., Lamazares, M. del C., Pérez García, L., Sosa Testé, I. M., Abeledo, M. A., Merino López, A., Fuente, D., & Gómez Álvarez, E. (2011). Evaluación epidemiológica de procesos respiratorios bacterianos en reemplazos de ponedoras. Revista De Salud Animal, 33(3). Recuperado a partir de http://revistas.censa.edu.cu/index.php/RSA/article/view/266.
4. Han, Moo-Sung, *et al.* "The Current Epidemiological Status of Infectious Coryza and Efficacy of Poulshot Coryza in Specific Pathogen-Free Chickens." *Journal of Veterinary Science*, vol. 17, no. 3, 2016, p. 323., https://doi.org/10.4142/jvs.2016.17.3.323.
5. Corney, B.G., *et al.* "Rapid and sensitive detection of *Avibacterium Paragallinarum* in the presence of other bacteria using a 5′ Taq nuclease assay: A New Tool for diagnosing Infectious coryza." Avian Pathology, vol. 37, no. 6, 2008, pp. 599–604., https://doi.org/10.1080/03079450802449139.
6. Guerrero Gonzáles Zaira. 2018. Anotación del genoma de *Avibacterium paragallinarum* cepa 5. Tesis de Maestría. Benemérita Universidad Autónoma de Puebla, Puebla, México.
7. Bennett, D.C., *et al*. 2015. Enzimas condroitin liasa. Google Patents: Biomarin Pharmaceutical Inc. (Consultado en 23/09/2020). Disponible en: https://patents.google.com/patent/ES2255062T3/es
8. Murtadha Nabeel, Abdul-Gani, and Bahaa Abdullah Laftaah. "Purification and characterization of Chondroitinase ABC from *Proteus Vulgaris*, an Iraqi clinically isolate." *Current Science*, vol. 113, no. 11, 2017, p. 2134., https://doi.org/10.18520/cs/v113/i11/2134-2140.
9. Allan, Nick D, *et al*. "Putative virulence factors are released in association with membrane vesicles from *Burkholderia Cepacia*." *Canadian Journal of Microbiology*, vol. 49, no. 10, 2003, pp. 613–624., https://doi.org/10.1139/w03-078.
10. Bisgaard, *et al*. 2014. The Family *Pasteurellaceae.* En E. R. al., The Prokaryotes – Gammaproteobacteria (págs. 536-564). Berlin Heidelberg: Springer-Verlag.
11. Blackall, Patrick J., *et al.* "Reclassification of *Pasteurella gallinarum*, [*Haemophilus*] *paragallinarum*, *Pasteurella Avium* and *Pasteurella Volantium* as *Avibacterium Gallinarum* Gen. Nov., Comb. Nov., *Avibacterium Paragallinarum* Comb. Nov., *Avibacterium Avium* Comb. Nov. and *Avibacterium Volantium* Comb. Nov.." *International Journal of Systematic and Evolutionary Microbiology*, vol. 55, no. 1, 2005, pp. 353–362., https://doi.org/10.1099/ijs.0.63357-0.

12. Blackall, Pat J., and Edgardo Soriano-Vargas. “Infectious Coryza and Related Bacterial Infections.” *Diseases of Poultry*, 2019, pp. 890–906., https://doi.org/10.1002/9781119371199.ch20.
13. Blackall, P. J. “Infectious coryza: Overview of the Disease and New Diagnostic Options.” *Clinical Microbiology Reviews*, vol. 12, no. 4, 1999, pp. 627–632., https://doi.org/10.1128/cmr.12.4.627.
14. Fang, Yangtao, *et al*. “Expression, purification and characterization of Chondroitinase AC II from marine bacterium *Arthrobacter Sp.*. CS01.” *Marine drugs*, vol. 17, no. 3, 2019, p. 185., https://doi.org/10.3390/md17030185.
15. Frey, Joachim, and Peter Kuhnert. “RTX toxins in *Pasteurellaceae*.” *International Journal of Medical Microbiology*, vol. 292, no. 3-4, 2002, pp. 149–158., https://doi.org/10.1078/1438-4221-00200.
16. Wu, Jin-Ru, *et al*. “Analysis of the biosynthesis genes and chemical components of the capsule of *Avibacterium paragallinarum*.” *Veterinary Microbiology*, vol. 145, no. 1-2, 2010, pp. 90–99., https://doi.org/10.1016/j.vetmic.2010.03.002.
17. Kadurugamuwa, J L, and T J Beveridge. “Virulence factors are released from *Pseudomonas aeruginosa* in association with membrane vesicles during normal Growth and Exposure to Gentamicin: A Novel Mechanism of Enzyme secretion.” *Journal of Bacteriology*, vol. 177, no. 14, 1995, pp. 3998–4008., https://doi.org/10.1128/jb.177.14.3998-4008.1995.
18. Liu, C.-C., *et al*. “The fimbrial protein is a virulence factor and potential vaccine antigen of *Avibacterium paragallinarum*.” *Avian Diseases*, vol. 60, no. 3, 2016, pp. 649–655., https://doi.org/10.1637/11410-031316-reg.1.
19. Pan, Yen-Chiang, *et al.* “Identification and characterization of an RTX toxin–like gene and its operon from *Avibacterium paragallinarum*.” *Avian Diseases*, vol. 56, no. 3, 2012, pp. 537–544., https://doi.org/10.1637/10047-122211-reg.1.
20. Persson, Gry, and Anders M Bojesen. “Bacterial determinants of importance in the virulence of *Gallibacterium anatis* in poultry.” *Veterinary Research*, vol. 46, no. 1, 2015, https://doi.org/10.1186/s13567-015-0206-z.
21. Ramón Rocha, Marcela. O., *et al.* “Membrane vesicles released by *Avibacterium paragallinarum* contain putative virulence factors.” *FEMS Microbiology Letters*, vol. 257, no. 1, 2006, pp. 63–68., https://doi.org/10.1111/j.1574-6968.2006.00154.x.
22. Rivero-García, PC, *et al.* “*Haemophilus paragallinarum* secretes metalloproteases.” *Canadian Journal of Microbiology*, vol. 51, no. 10, 2005, pp. 893–896., https://doi.org/10.1139/w05-067.
23. Sakamoto, Ryuichi, *et al.* “Development of a recombinant vaccine against infectious coryza in chickens.” *Research in Veterinary Science*, vol. 94, no. 3, 2013, pp. 504–509., https://doi.org/10.1016/j.rvsc.2012.10.027.
24. Soriano-Vargas, Edgardo, Sánchez-Morales, Josué, Morales-Erasto, Vladimir, & García-Sánchez, Alejandra. (2013). Aislamiento e identificación de *Avibacterium paragallinarum* NAD independiente (serovariedad C-1) en gallinas de reemplazo con

coriza infecciosa. Veterinaria México, 44(1), 57-62. Recuperado en 13 de diciembre de 2022, de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0301-50922013000100005&lng=es&tlng=es.

25. Takagi, M, *et al.* "Purification of hemagglutinin from *Haemophilus paragallinarum* using monoclonal antibody." *Veterinary Microbiology*, vol. 34, no. 2, 1993, pp. 191 197., https://doi.org/10.1016/0378-1135(93)90172-4.
26. Wang, Yi-Ping, *et al.* "The haemagglutinin of *Avibacterium paragallinarum* is a trimeric autotransporter adhesin that confers haemagglutination, cell adherence and biofilm formation activities." *Veterinary Microbiology*, vol. 174, no. 3-4, 2014, pp. 474–482., https://doi.org/10.1016/j.vetmic.2014.10.013.
27. Luna Trinidad Angel.(2010). Evaluación de dos bacterinas en gallinas de postura, mediante detección de anticuerpos inhibidores de la hemoaglutinación y protección ante un desafío con *Avibacterium paragallinarum*, en Tehuacán, Puebla (Tesis de Licenciatura). Universidad Veracruzana. Tehuacán, Puebla.
28. Chen, Y.-C., *et al.* "Identification and functional analysis of the cytolethal distending toxin gene from *Avibacterium paragallinarum*." *Avian Pathology*, vol. 43, no. 1, 2013, pp. 43–50., https://doi.org/10.1080/03079457.2013.861895.
29. Xu, Fuzhou, et al. "Draft genome sequence of *Avibacterium paragallinarum* strain 221." Genome Announcements, vol. 1, no. 3, 2013, https://doi.org/10.1128/genomea.00290-13.
30. Hamai, Akio, *et al.* "Two distinct chondroitin sulfate ABC lyases." *Journal of Biological Chemistry*, vol. 272, no. 14, 1997, pp. 9123–9130., https://doi.org/10.1074/jbc.272.14.9123.
31. https://www.kegg.jp/dbget-bin/www_bget?ec:4.2.2.20.
32. "Information on EC 4.2.2.5 - Chondroitin AC Lyase." *Information on EC 4.2.2.5 - Chondroitin AC Lyase - BRENDA Enzyme Database*, https://www.brenda enzymes.org/enzyme.php?ecno=4.2.2.5.
33. Paget, Mark SB, and John D Helmann. "The sigma70 family of sigma factors." *Genome Biology*, vol. 4, no. 1, 2003, p. 203., https://doi.org/10.1186/gb 2003-4-1-203.
34. Boyce, John D., *et al.* "Genomic scale analysis of *Pasteurella multocida* gene expression during growth within the natural chicken host." *Infection and Immunity*, vol. 70, no. 12, 2002, pp. 6871–6879., https://doi.org/10.1128/iai.70.12.6871 6879.2002.
35. Craig, Jane E., *et al.* "The extracytoplasmic sigma factor, ςE , is required for intracellular survival of non typeable *Haemophilus influenzae* in J774 macrophages." *Infection and Immunity*, vol. 70, no. 2, 2002, pp. 708–715., https://doi.org/10.1128/iai.70.2.708-715.2002.

36. Chantes-Guerra, Alberto, *et al.* "A putative siderophore receptor of *Gallibacterium anatis* 12656-12 under Fur control also binds hemoglobin." *Frontiers in Microbiology*, vol. 13, 2022, https://doi.org/10.3389/fmicb.2022.951173.
37. Kristensen, Bodil M., *et al.* "Natural transformation of *Gallibacterium anatis*." *Applied and Environmental Microbiology*, vol. 78, no. 14, 2012, pp. 4914 4922., https://doi.org/10.1128/aem.00412-12.
38. Cho, Byung-Kwan, *et al.* "The transcription unit architecture of the *Escherichia coli* genome." *Nature Biotechnology*, vol. 27, no. 11, 2009, pp. 1043–1049., https://doi.org/10.1038/nbt.1582.
39. Thieffry, D., *et al.* "Prediction of transcriptional regulatory sites in the complete genome sequence of *Escherichia coli* K-12." *Bioinformatics*, vol. 14, no. 5, 1998, pp. 391–400., https://doi.org/10.1093/bioinformatics/14.5.391.
40. Taverner, Nicola V, *et al.* "Identifying transcriptional targets." *Genome Biology*, vol. 5, no. 3, 2004, p. 210., https://doi.org/10.1186/gb-2004-5-3-210.
41. Perez-Rueda, E. "The repertoire of DNA-Binding transcriptional regulators in *Escherichia coli* K-12." *Nucleic Acids Research*, vol. 28, no. 8, 2000, pp. 1838–1847., https://doi.org/10.1093/nar/28.8.1838.
42. Weindl, Johanna, *et al.* "Modeling DNA-binding of *Escherichia coli* σ70 exhibits a characteristic energy landscape around strongs promoters." *Nucleic Acids Research*, vol. 35, no. 20, 2007, pp. 7003–7010., https://doi.org/10.1093/nar/gkm720.
43. Gioia, Jason, *et al.* "The genome sequence of *Mannheimia haemolytica* A1: Insights into virulence, natural competence, and *Pasteurellaceae* phylogeny." *Journal of Bacteriology*, vol. 188, no. 20, 2006, pp. 7257–7266., https://doi.org/10.1128/jb.00675-06.
44. Bossé, Janine T., et al. "Complete genome for *Actinobacillus pleuropneumoniae* serovar 8 Reference Strain 405: Comparative Analysis with Draft Genomes for Different Laboratory Stock Cultures Indicates Little Genetic Variation." *Microbial Genomics*, vol. 7, no. 11, 2021, https://doi.org/10.1099/mgen.0.000687.
45. "*Avibacterium paragallinarum* strain AVPG2015 chromosome, complete Genome - Nucleotide - NCBI." *National Center for Biotechnology Information*, U.S. National Library of Medicine, https://www.ncbi.nlm.nih.gov/nuccore/CP058307.1.
46. Morales-Erasto, V, *et al.* "Eric-PCR genotyping of emergent serovar C-1 isolates of *Avibacterium paragallinarum* from Mexico." *Avian Diseases*, vol. 55, no. 4, 2011, pp. 686–688., https://doi.org/10.1637/9670-012711-resnote.1.
47. Expasy is operated by the SIB Swiss Institute of Bioinformatics | Terms of Use. "Chondroitin AC Lyase." *4.2.2.5 Chondroitin AC Lyase*, https://enzyme.expasy.org/EC/4.2.2.5.
48. Takagaki,K., and I. Kakizaki. "Degradation of glycosaminoglycans." *Comprehensive Glycoscience*, 2007, pp. 171–192., https://doi.org/10.1016/b978-0444519672/00043-x.

49. Henrotin, Yves, *et al.* "Chondroitin sulfate in the treatment of osteoarthritis: From *in Vitro*Studies to clinical recommendations." *Therapeutic Advances in Musculoskeletal Disease*, vol. 2, no. 6, 2010, pp. 335–348., https://doi.org/10.1177/1759720x10383076.
50. Valcarcel, Jesus, *et al.* "Glycosaminoglycans from marine sources as therapeutic agents." *Biotechnology Advances*, vol. 35, no. 6, 2017, pp. 711–725., https://doi.org/10.1016/j.biotechadv.2017.07.008.
51. Valcarcel, Jesus, et al. "Marine chondroitin sulfate of defined molecular weight by enzymatic depolymerization." *Carbohydrate Polymers*, vol. 229, 2020, p. 115450., https://doi.org/10.1016/j.carbpol.2019.115450.
52. Ramón Rocha, Marcela. O., *et al.* "Membrane vesicles released by *Avibacterium paragallinarum* contain putative virulence factors." *FEMS Microbiology Letters*, vol. 257, no. 1, 2006, pp. 63–68., https://doi.org/10.1111/j.1574-6968.2006.00154.x.
53. Mei, Chen, *et al.* "Component identification and functional analysis of outer membrane vesicles released by *Avibacterium paragallinarum*." *Frontiers in Microbiology*, vol. 11, 2020, https://doi.org/10.3389/fmicb.2020.518060.
54. Araya-hidalgo, Edward, *et al.* "Sequence analysis of the hypervariable region in *hmtp210* of *Avibacterium paragallinarum*." *Journal of Veterinary Medical Science*, vol. 79, no. 7, 2017, pp. 1210–1214., https://doi.org/10.1292/jvms.16-0530.
55. Küng, Eliane, and Joachim Frey. "AvxA, a composite serine-protease-RTX toxin of *Avibacterium paragallinarum*." *Veterinary Microbiology*, vol. 163, no. 3-4, 2013, pp. 290–298., https://doi.org/10.1016/j.vetmic.2012.12.029.
56. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410. PubMed
57. The RAST Server: Rapid Annotations using Subsystems Technology. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.
58. Tu, Tzu-Yi, *et al.* "Loss of the capsule increases the adherence activity but decreases the virulence of *Avibacterium paragallinarum*." *Avian Diseases*, vol. 59, no. 1, 2015, pp. 87–93., https://doi.org/10.1637/10937-091414-reg.
59. Pérez-Márquez, Víctor, *et al.* "Secreted proteins of *Avibacterium paragallinarum* are lethal for chicken embryo." *Annals of the New York Academy of Sciences*, vol. 1149, no. 1, 2008, pp. 380–383., https://doi.org/10.1196/annals.1428.007.
60. Jinadasa, Rasika N., *et al.* "Cytolethal distending toxin: A conserved bacterial genotoxin that blocks cell cycle progression, leading to apoptosis of a broad range of mammalian cell lineages." *Microbiology*, vol. 157, no. 7, 2011, pp. 1851–1875., https://doi.org/10.1099/mic.0.049536-0.

61. Beiranvand, Sheida, *et al.* "Novel NAD-Independent *Avibacterium paragallinarum* : Isolation, characterization and molecular identification in Iran." *Veterinary Medicine and Science*, vol. 8, no. 3, 2022, pp. 1157–1165., https://doi.org/10.1002/vms3.754.
62. Fernández, R. P., *et al.* "Protection conferred by bivalent and trivalent infectious coryza bacterins against prevalent serovars of *Avibacterium* (*Haemophilus*) *paragallinarum* in Mexico." *Avian Diseases*, vol. 49, no. 4, 2005, pp. 585–587., https://doi.org/10.1637/7355-031405r1.1.
63. Soriano Vargas, E., Terzolo, H.R., (2004). Epizootiologia, prevención y control de la coriza infecciosa. Veterinaria México, 35(3): 261-279.
64. Jeong, Ok-Mi, *et al.* "Isolation and characterization of *Avibacterium paragallinarum* with different Nicotinamide Adenine Dinucleotide requirements." *Veterinary Microbiology*, vol. 205, 2017, pp. 62–65., https://doi.org/10.1016/j.vetmic.2017.05.005.
65. V. Solovyev, A Salamov (2011) Automatic Annotation of Microbial Genomes and Metagenomic Sequences. In Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies (Ed. R.W. Li), Nova Science Publishers, p. 61-78
66. Crooks GE, Hon G, Chandonia JM, Brenner SE WebLogo: A sequence logo generator, Genome Research, 14:1188-1190, (2004) [Full Text].
67. Predicción de redes de regulación transcripcional, programa String (Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ, von Mering C‡. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021 Jan 8;49(D1):D605-12.
68. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296-W303 (2018).
69. Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., Schwede, T. The SWISS-MODEL Repository - new features and functionality. Nucleic Acids Res. 45, D313-D319 (2017).
70. Guex, N., Peitsch, M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30, S162-S173 (2009).
71. Studer, G., Tauriello, G., Bienert, S., Biasini, M., Johner, N., Schwede, T. ProMod3 - A versatile homology modelling toolbox. PLOS Comp. Biol. 17(1), e1008667 (2021).
72. Studer, G., Rempfer, C., Waterhouse, A.M., Gumienny, R., Haas, J., Schwede, T. QMEANDisCo - distance constraints applied on model quality estimation. Bioinformatics 36, 1765-1771 (2020).

73. Studer, G., Biasini, M., Schwede, T. Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane), Bioinformatics 30, i505–i511 (2014).
74. Benkert, P., Biasini, M., Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27, 343-350 (2011).
75. Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Scientific Reports 7 (2017).
76. Mariani, V., Biasini, M., Barbato, A., Schwede, T. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. Bioinformatics 29, 2722-2728 (2013).
77. W Zheng, C Zhang, Y Li, R Pearce, EW Bell, Y Zhang. Folding non-homology proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. Cell Reports Methods, 1: 100014 (2021))
78. PyMOL (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC).
79. Barcenas. Gabriela (2020). Identificación funcional de dos operones relacionados con la utilización de carbohidratos específicos de tejidos de origen animal, EN *Avibacterium paragallinarum.* Benemérita Universidad Autónoma de Puebla.
80. V. Solovyev, A Salamov (2011) Automatic Annotation of Microbial Genomes and Metagenomic Sequences. In Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies (Ed. R.W. Li), Nova Science Publishers, p. 61-78.
81. Tremolo, H. R. (2000). Coriza lnfecciosa: una revisión. Propuestas de investigación para su diagnóstico y control. Revista de Medicina Veterinaria, 81, 262-269.https://www.researchgate.net/publication/280741318_Coriza_Infecciosa_una_revision_Propuestas_de_investigacion_para_su_diagnostico_y_control.
82. Horta-Valerdi, Guillermo, *et al.* "The genome sequence of *Avibacterium paragallinarum* strain CL has a large repertoire of insertion sequence elements." *Genome Announcements*, vol. 5, no. 15, 2017, https://doi.org/10.1128/genomea.00152-17.
83. "VetBact." *Start Page*, https://www.vetbact.org/?artid=136.
84. Cobos-Justo, M.E.; (Benemérita Universidad Autónoma de Puebla, Puebla, Puebla, México). Comunicación Personal, 2021.