



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA



Instituto de Ciencias

**Centro de Investigaciones en Ciencias Microbiológicas
(CICM-ICUAP)**

Posgrado en Ciencias Microbiológicas

**CARACTERIZACIÓN MOLECULAR DE RASGOS DE VIRULENCIA EN AISLADOS
DE *Avibacterium paragallinarum*, CON EL APOYO DE LA COMPARACIÓN DE
GENOMAS**

TESIS

PARA OBTENER EL GRADO DE DOCTORA EN CIENCIAS (MICROBIOLOGÍA)

PRESENTA:

M.C. María Elena Cobos Justo

DIRECTOR DE TESIS:

D.C. Candelario Vázquez Cruz

JUNIO 2023



BUAP

Puebla, Pue. a 2 de mayo 2023.

**A LA ACADEMIA DEL POSGRADO
EN MICROBIOLOGÍA
CICM-ICUAP
PRESENTE.**

Por este conducto los abajo firmantes, integrantes del Comité revisor de Tesis de Doctorado de la alumna **María Elena Cobos Justo**, les informamos que hemos revisado el escrito titulado:

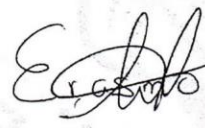
“Caracterización molecular de rasgos de virulencia en aislados de *Avibacterium paragallinarum*, con el apoyo de la comparación de genomas”

A nuestro juicio, la alumna **María Elena Cobos Justo** 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”


D.C. María Patricia Sánchez Alonso


D.C. Erasmo Negrete Abascal


D.C. Leticia Vega Alvarado


D.C. Gustavo Hernández Guzmán

AGRADECIMIENTOS

A mi director de tesis:

D.C. Candelario Vázquez Cruz, por abrirme las puertas de su laboratorio y por todo el apoyo que me ha dado para la realización de este proyecto.

A mi comité revisor:

- D.C. Patricia Sánchez Alonso
- D.C. Leticia Vega Alvarado
- D.C. Erasmo Negrete Abascal
- D.C. Gustavo Hernández Guzmán

Gracias por su apoyo a lo largo de todo este tiempo, por enriquecer este proyecto con todas sus observaciones y sugerencias.

A CONACYT por beca a María Elena Cobos Justo No. 297565

Al proyecto CONACYT CB-2015-01-259209

Al proyecto BUAP 100103133-VIEP 2022

Proyecto apoyado por UNAM-DGAPA-PAPIIT IN204122

INDICE

ABREVIATURAS	8
RESUMEN	9
I. INTRODUCCIÓN.....	10
II. ANTECEDENTES	12
2.1. Factores de virulencia	12
2.1.1. Toxinas RTX (Avx)	12
2.1.2. Toxina Distensora Citoletal	14
2.1.3. Fimbrias.....	15
2.1.4. Hemaglutinina	16
2.1.5. Cápsula	17
2.1.6. Proteasas	17
2.1.7. Condroitin liasas	18
2.2. Tecnologías de secuenciación de DNA genómico.....	19
2.2.1. Illumina	20
2.2.2. Pacific Biosciences.....	20
2.3. Análisis y comparación de genomas	21
III. JUSTIFICACIÓN	23
IV. OBJETIVOS.....	24
V. MATERIALES Y MÉTODOS.....	25
5.1. Cepas bacterianas	25
5.2. Condiciones de crecimiento y extracción de DNA genómico	26
5.2.1. Cuantificación de DNA genómico	27
5.2.2. Relación 260/280 y 260/230 de DNA	27
5.3. Datos generales	27
5.4. Análisis y mejoramiento de la calidad de secuencias de Illumina MiSeq.....	28
5.5. Ensamble de genomas	28
5.5.1. Illumina MiSeq.....	28
5.5.2. PacBio	28
5.6. Anotación en Rapid Annotation using Subsystem Technology (RAST)	29
5.7. Resistencia antimicrobiana.....	29
5.7.1. Gel de lisis <i>in situ</i> para la visualización de plásmidos de alto peso molecular.....	29

5.7.2. Antibiógramas por difusión en disco.....	30
VI. RESULTADOS	31
6.1. Extracción de DNA genómico	31
6.1.1. Cuantificación de DNA genómico	32
6.1.2. Relación 260/280 y 260/230 de DNA	32
6.2. Análisis de calidad en los datos de secuencias de Illumina MiSeq.....	33
6.3. Ensamble de genomas con secuencias de Illumina MiSeq	34
6.4. Ensamble de genomas con secuencias PacBio	36
6.4.1. Ensamblados de novo	36
6.4.2. Ensamblados híbridos	37
6.5. Anotación en Rapid Annotation using Subsystem Technology (RAST)	38
6.6. Comparación lineal y filogenética de cepas de <i>A. paragallinarum</i>	43
6.7. Búsqueda de nuevos factores de virulencia	46
6.8. Otros elementos	52
6.9. Determinantes genéticos relacionados con Resistencia antimicrobiana	56
6.10. Antibiógramas por difusión en disco.....	58
VII. DISCUSIÓN.....	60
VIII. CONCLUSIONES	66
IX. PERSPECTIVAS	67
X. REFERENCIAS.....	68
XII. MANUSCRITO DEL ARTICULO PUBLICADO.....	84

INDICE DE FIGURAS

Figura 1. DNA genómico obtenido para secuenciar.	31
Figura 2. Gráfico de barras y bigotes. Representación de parámetros del análisis de calidad de secuencias de Illumina.	33
Figura 3. Gráfico de barras y bigotes de secuencias editadas.	33
Figura 4. Clasificación de los genes anotados en RAST.....	38
Figura 5. Comparación filogenética y porcentaje de identidad del gen <i>cdtB</i> en cepas de <i>A. paragallinarum</i>	39
Figura 6. Comparación filogenética y porcentaje de identidad del gen <i>avxA</i> en cepas de <i>A. paragallinarum</i>	40
Figura 7. Comparación filogenética y porcentaje de identidad de la secuencia del gen <i>ushI</i> en cepas de <i>A. paragallinarum</i>	41
Figura 8. Comparación filogenética y porcentaje de identidad de la secuencia del gen <i>ushII</i> en cepas de <i>A. paragallinarum</i>	42
Figura 9. Comparación lineal de tres cepas de <i>A. paragallinarum</i>	43
Figura 10. Comparación lineal de cepas de <i>A. paragallinarum</i>	44
Figura 11. Agrupación filogenética de cepas de <i>A. paragallinarum</i> basado en MLST.	45
Figura 12. Organización del genoma de <i>A. paragallinarum</i> cepa 2015.....	53
Figura 13. Heat map correspondiente a reguladores transcripcionales en cepas de diferentes géneros bacterianos.....	55
Figura 14. Gel <i>in situ</i> de Eckhardt.....	56
Figura 15. Estructura secundaria del 16S rRNA.	57

INDICE DE TABLAS

Tabla 1. Cepas de <i>A. paragallinarum</i> secuenciadas en este proyecto.	25
Tabla 2. Cepas de <i>A. paragallinarum</i> secuenciadas con diferentes tecnologías de secuenciación.	27
Tabla 3. Datos generales de ensamble de genoma de 11 cepas de <i>A. paragallinarum</i> con ABySS.	34
Tabla 4. Datos generales de ensamble de genoma de 11 cepas de <i>A. paragallinarum</i> con SPAdes.	35
Tabla 5. Datos generales de ensamble de genoma de 11 cepas de <i>A. paragallinarum</i> con MIRA.	35
Tabla 6. Datos generales del ensamble de novo de 10 cepas de <i>A. paragallinarum</i> con secuencias de PacBio.	36
Tabla 7. Datos generales del ensamble híbrido de 8 cepas de <i>A. paragallinarum</i>	37
Tabla 8. Genes asociados a virulencia presentes en 16 cepas de <i>A. paragallinarum</i>	47
Tabla 9. Función de probables genes de <i>A. paragallinarum</i> relacionados con virulencia. .	49
Tabla 10. Proteasas encontradas en el genoma de las diferentes cepas.	51
Tabla 11. Número de proteínas hipotéticas, fagos y proteínas de elementos móviles presentes en las 16 cepas de esta investigación y cepas de diferentes países.	52
Tabla 12. Resistencia antimicrobiana de cepas de <i>A. paragallinarum</i> , R=Resistente, I=Intermedio, S=sensible.	58

ABREVIATURAS

CI	Coriza Infecciosa
FV	Factor de Virulencia
RTX	Repetidos en Toxina Estructural
CDT	Toxina Distensora Citoletal
CU	Chaperona-Usher
HA	Hemaglutinina
GAG	Glucosaminoglicanos
CS	Sulfato de condroitina
DS	Sulfato de dermatán
SMRT	Single Molecule Real Time
PacBio	Pacific Biosciences
ZMW	Zero Mode Waveguide
IPD	Inter Pulse Duration
FD	Factor de dilución
DO	Densidad óptica
°C	Grados centígrados
BHI	Infusión Cerebro Corazón
CO₂	Dióxido de carbono
EDTA	Ácido etilen-diamino-tetra-acético
rpm	Revoluciones por minuto
pb	Pares de bases
SI	Sistema inmune
MV's	Vesículas de membrana
IgG	Inmunoglobulina G
TA	Toxina/Antitoxina
MLST	Multi Locus Sequence Typing

RESUMEN

La coriza infecciosa (CI) es una enfermedad cosmopolita, que puede causar pérdidas económicas importantes en países dedicados a la producción de huevo y carne de pollo. México es uno de ellos y para hacer frente a la enfermedad es importante combatir eficientemente al agente causal: la bacteria *Avibacterium paragallinarum*, modelo de estudio en la presente investigación. Para prevenir la CI, están disponibles las vacunas elaboradas con bacterias atenuadas. Sin embargo, se ha observado protección incompleta, pues existen diferentes serovariedades de *A. paragallinarum* que aparecen subsecuentemente en la cadena productiva de las granjas, con frecuentes reinfecciones de CI por falta de memoria inmunológica. Por tanto, es necesario recurrir a la antibioterapia y vigilar el comportamiento de resistencia y sensibilidad al tratamiento.

El objetivo de este trabajo fue identificar los factores de virulencia de cepas de *A. paragallinarum* de reciente aislamiento, para comprender mejor el proceso patogénico de la CI. Para ello se hizo la secuenciación de los genomas con tecnología de nueva generación y la anotación pormenorizada de los genes, de acuerdo a la tendencia tecnológica actual. Con información actualizada se puede coadyuvar al desarrollo de nuevas herramientas inmunogénicas para combatir la CI, como bacterinas más eficientes. En el presente trabajo se reporta la secuencia de 16 genomas de cepas de *A. paragallinarum* utilizando las tecnologías Illumina y PacBio, además de la comparación entre estos. De estos genomas cinco fueron circularizados y dos registrados en GenBank. El análisis genómico comparativo, permitió identificar genes asociados con la virulencia, previamente descritos, tales como *avxA*, *cdtB* y un operón fimbrial, así como genes sin reporte previo o asociación con virulencia.

Los genes anotados en este trabajo relacionados con la virulencia corresponden a procesos como: la adherencia, producción de bombas de expulsión, resistencia a fluoroquinolonas, producción de bacteriocinas, formación de cápsula, metabolismo de carbohidratos, proteasas, y sistemas toxina-antitoxina. Asimismo, se reporta el perfil de resistencia a antimicrobianos de las cepas investigadas.

I. INTRODUCCIÓN

La familia *Pasteurellaceae* está constituida por bacterias patógenas de gran importancia para la producción animal y salud humana, poseen genomas relativamente pequeños y son capaces de sobrevivir a ambientes “extremos” como agua de bebederos de granja y temperaturas de refrigeración; de ese grupo un integrante importante en la microbiología veterinaria es *Avibacterium paragallinarum* (Christensen *et al.*, 2014).

Avibacterium paragallinarum es una bacteria Gram negativa pleomórfica, inmóvil, sin esporas, con dimensiones de 1 a 3 μm de longitud por 0.4 a 0.8 μm de ancho. Crece como gotas de rocío a 37°C en condiciones de microaerofilia en medios nutritivos como Infusión Cerebro-Corazón (BHI), medio Casman, medio Columbia, agar sangre y agar chocolate que deben ser enriquecidos con NADH (Terzolo & Moreira, 2005). De acuerdo con el esquema de antígenos y clasificación serológica de *A. paragallinarum*, Page diferenció tres serogrupos designados como A, B y C (Page., 1962); posteriormente Kume identificó siete serotipos (HA-1 a HA-7) (Kume *et al.*, 1983) y finalmente Blackall propuso nueve serovariedades distribuidas en tres serogrupos A-1, A-2, A-3, A-4, B-1, C-1, C-2, C-3 y C-4 (Blackall *et al.*, 1990). *A. paragallinarum* es: catalasa (-), pigmento (-), oxidasa (-), ureasa (-), indol (-), microaerofilia (+), β -galactosidasa (+), arabinosa (-), galactosa (-), maltosa (+), manitol (+), sorbitol (V), sacarosa (V), trealosa (-), lisina descarboxilasa (-) y ornitina descarboxilasa (-) (Terzolo & Moreira, 2005).

La enfermedad causada por *A. paragallinarum* es la Coriza Infecciosa (CI) un padecimiento avícola altamente contagioso. Esta enfermedad puede causar una disminución en la producción de huevo que va del 10 al 40%. Se encuentra distribuida en países de todo el mundo, la raza y la edad son considerados factores de riesgo (Swayne *et al.*, 2013; Dereja *et al.*, 2017), y tiene mayor incidencia en ciertos meses del año, quizás por el cambio de las condiciones climáticas (Crispo *et al.*, 2019).

La infección con *A. paragallinarum* puede coincidir con infecciones por otras bacterias patógenas como *Mycoplasma gallisepticum*, *Escherichia coli*, *Pasteurella multocida*, *Gallibacterium anatis*, *Ornithobacterium rhinotracheale* y virus como Newcastle, dando origen a la CI complicada (Morales *et al.*, 2016; Van *et al.*, 2019). Las aves con CI presentan signos y síntomas como decaimiento, estertores, estornudos, inflamación periorbitaria, senos inflamatorios hinchados, edema palpebral, ojos llorosos, secreción ocular y nasal, desorientación, torticollis y bajo peso (Crispo *et al.*, 2019), aunque algunas aves infectadas no presentan signos clínicos (Dereja *et al.*, 2017).

Actualmente las técnicas tradicionales de cultivo, caracterización bioquímica y serología de *A. paragallinarum* son vigentes. Sin embargo, al ser un microorganismo “problemático o fastidioso” para su cultivo *in vitro* llega a ser difícil su caracterización de virulencia, por lo que los métodos basados en secuenciación de DAN o de genomas han ido ganando terreno, pues se obtiene información simultánea sobre su genotipo y grupo taxonómico, contenido de factores de virulencia y posibles mecanismos de interacción entre la bacteria y el hospedero, proporcionando un adelanto significativo en su caracterización (Comas *et al.*, 2020; Alföldi & Lindblad, 2013).

II. ANTECEDENTES

2.1. Factores de virulencia

La virulencia es la habilidad que posee un organismo para poder infectar a su hospedero y causar enfermedad, en esta habilidad subyacen los llamados factores de virulencia (FV). Los FV son moléculas de la bacteria que le ayudan a evadir las defensas del huésped para dañarlo y colonizarlo. Dependiendo del microorganismo patogénico podemos encontrar diferentes clases de FV. Unos pueden ser secretados, por ejemplo, las toxinas, enzimas y exopolisacáridos, otros pueden tener una función estructural en la superficie celular bacteriana como la cápsula, lipopolisacáridos, glicoproteínas o vesículas de membrana. Existen los que son apéndices de colonización, como fimbrias o adhesinas. O bien, pueden tener funciones regulatorias sobre redes de control de la expresión génica, mediadas por proteínas y RNAs no-codificantes (Sharma *et al.*, 2017; Leitao, 2020). A continuación, se describirán brevemente los FV investigados en *A. paragallinarum*.

2.1.1. Toxinas RTX (Avx)

Las toxinas RTX (Repetidos en la Toxina estructural) son toxinas de naturaleza proteica que forman poros en las células blanco, pueden tener actividad citotóxica y/o hemolítica. Estas toxinas son producidas principalmente por las bacterias Gram negativas patógenas incluyendo varias especies de la familia *Pasteurellaceae* (Frey & Kuhnert, 2002). La principal característica de las toxinas RTX es la presencia de repeticiones ricas en glicina y aspartato que unen calcio en la región C-terminal, estas repeticiones mantienen desplegada a la proteína RTX, para que sea reconocida su señal de secreción C-terminal por un sistema basado en transportadores de casete de unión a ATP (ABC), que permite la traslocación de los polipéptidos de RTX en un solo paso a través de un canal de secreción oligomérico que se extiende por la membrana, citoplasma, espacio periplásmico y la membrana externa sin un intermediario de secreción periplásmico (Koronakis *et al.*, 1989). Cuando las repeticiones van emergiendo se

unen a los iones de calcio extracelulares y se activa la estructuración de las repeticiones, sacando la proteína de la bacteria e impulsando el plegamiento del resto de la proteína. Las repeticiones de RTX contribuyen a la adherencia, exportación, interacción y penetración del huésped, además de provocar lisis y muerte celular (Linhartová *et al.*, 2010; Benz, 2016).

Aunque las toxinas RTX pueden provocar hemólisis por la formación de poros en la membrana celular, sus objetivos específicos son los leucocitos que expresan integrinas $\beta 2$ y que actúan como receptores específicos para las toxinas. Se ha demostrado que la unión a la fracción CD18 de las integrinas $\beta 2$ con las toxinas RTX es específica del huésped (Frey, 2019), y después de ser absorbida por la célula diana se transportan a la mitocondria e inducen apoptosis (Atapattu & Czuprynski, 2005; Kisiela *et al.*, 2010; Frey, 2011).

El análisis del genoma completo de *A. paragallinarum* 221 mostró la presencia de un operón similar al de la toxina RTX, este operón está formado por cuatro genes, *avxC*, *avxA*, *avxB* y *avxD* (Pan *et al.*, 2012). *avxC* es el gen activador, *avxA* el gen estructural y *avxB* y *avxD* son los genes que codifican para un sistema de secreción tipo I. *AvxA* tiene dos dominios, el N-terminal es homólogo a una serin proteasa y el C-terminal homólogo a una citotoxina, este último dominio está formado por 11 repeticiones ricas en glicina que se unen a iones de Ca^{2+} . Ambos dominios se separan entre sí, la actividad de serin-proteasa no se ha evaluado a diferencia de la RTX, *AvxA* pertenece a las toxinas RTX de autoprocésamiento multifuncional (Küng & Frey, 2013). *AvxA* es un factor de virulencia común en todos los serotipos de *A. paragallinarum* y tiene un rol clave en la patogénesis de la bacteria, por su capacidad para inducir una fuerte respuesta inmunológica podría ayudar al desarrollo de vacunas eficientes (Küng & Frey 2013, Frey, 2019).

2.1.2. Toxina Distensora Citoletal

La toxina distensora citoletal (CDT) ha sido estudiada en bacterias Gram negativas patógenas, se le cataloga como genotoxina por actuar sobre el material genético. Su actividad es mediada por tres polipéptidos, CdtA, CdtB y CdtC, codificados por los respectivos genes *cdtA*, *cdtB* y *cdtC*. Sin embargo, algunos estudios han demostrado que CdtB sola es suficiente para causar un arresto del ciclo celular en las fases G2/M pues tiene actividad DNasa, pero es más potente en presencia de CdtA y CdtC (Shenker *et al.*, 2004; Shima *et al.*, 2012). CDT detiene el ciclo celular dando como resultado la distensión del núcleo y el citoplasma, seguido de la muerte celular apoptótica (Jinadasa *et al.*, 2011). En el genoma de *A. paragallinarum* de las cepas 221, H8 y TW07 se identificaron los genes toxigénicos *cdtABC*. El extracto del lisado libre de células de TW07 mostró actividad citolítica en líneas celulares HeLa y de fibroblastos de embrión de pollo y la proteína recombinante CdtA-AVPG mostró actividad de DNasa (Chen *et al.*, 2014).

CDT es considerada un factor de virulencia capaz de atravesar la capa mucosa o epitelial del huésped, además suprime el sistema inmune y promueve la respuesta proinflamatoria, por lo que se sugiere que es clave en las primeras etapas de la infección. En infecciones por *Aggregatibacter actinomycetemcomitans* la CDT se propone que contribuye a la persistencia de la enfermedad (DiRienzo, 2014); es probable que la respuesta proinflamatoria proporcione fragmentos de proteína tisular y hemina que cubren las necesidades nutricionales de la bacteria (Shenker *et al.*, 2015).

La investigación del efecto citotóxico de CDT de bacterias de diferentes géneros ha mostrado que cada una de estas toxinas se incorpora a las células blanco por vías diferentes, después de reconocer receptores específicos, por ejemplo, carbohidratos, contenido de colesterol y el linaje celular del huésped infectado (McSweeney & Dreyfus, 2005; Eshragui *et al.*, 2010; Gargi *et al.*, 2013).

2.1.3. Fimbrias

En bacterias Gram negativas y en *A. paragallinarum* se han identificado fimbrias como factores de virulencia. Las fimbrias son orgánulos formados de largos polímeros de proteínas que se localizan en la superficie celular y juegan un papel importante durante la infección pues contribuyen en la adherencia, colonización, tropismo tisular y formación de biopelícula (Wurpel *et al.*, 2013; Kudirkiene *et al.*, 2014). Algunas fimbrias son ensambladas y secretadas por un sistema Chaperona-Usher (CU) (Gaibel & Waksman, 2014; Liu *et al.*, 2016). Los genes fimbriales generalmente están agrupados en operones y codifican cuatro componentes; una chaperona periplásmica que ayuda al plegamiento de las subunidades fimbriales, la llamada usher, que es proteína integral de membrana externa para facilitar la liberación y secreción de dos proteínas fimbrias, la subunidad fimbrial principal y la adhesina (Nuccio & Bäumler, 2007).

En *A. paragallinarum* se identificó un gen que codifica para la proteína fimbrial FlfA, el gen se clonó en un vector de expresión, la proteína recombinante se expresó y purificó. En ensayos realizados en pollos se observó que la proteína recombinante inducía protección contra la coriza infecciosa. Además, las cepas mutantes nulas (knockout) del gen *flfA* mostraron ser menos virulentas que las cepas silvestres (Liu *et al.*, 2016). Resultados similares han sido reportados con *Gallibacterium anatis*. Esta bacteria que produce fimbrias F17 y las utiliza como importante factor de virulencia para colonizar y hacerle daño al hospedero. En este microorganismo se han identificado hasta tres operones fimbriales CU por cepa, observando que los genes que codifican las subunidades fimbriales y las adhesinas son más diversos en secuencia. Esto le representa a la bacteria mayor versatilidad para expresar fimbrias en diferentes condiciones y poder interactuar con varios tipos de células del huésped (Kudirkiene *et al.*, 2014). Respecto a la secuencia de usher, esta es más conservada y al ser la plataforma de secreción es un potencial blanco terapéutico, pues se ha propuesto el uso de inhibidores de moléculas pequeñas que se dirijan al proceso de secreción con el fin de evitar la

expresión de las fimbrias en la superficie de la bacteria para evitar la adhesión, colonización y formación de biopelícula (Werneburg & Thanassi, 2018).

2.1.4. Hemaglutinina

Con frecuencia se observa que los microorganismos colonizan diferentes células del huésped por adherencia superficial y formación de biopelículas a través de proteínas con actividad de hemaglutininas (Wang *et al.*, 2014). En *A. paragallinarum* se determinó que la proteína HMTp210 es el principal antígeno hemaglutinante (HA). La cepa mutante en HA mostró disminución en su capacidad de adhesión a cultivos de células HeLa y formación de biopelícula, y resultó ser menos virulentas que la cepa silvestre (Wang *et al.*, 2014). También se determinó que la actividad de HMTp210 varía entre las cepas analizadas, y se propone que esto se debe al diferente contenido de la proteína HA en la superficie de las bacterias (Wang *et al.*, 2014; Wu *et al.*, 2011).

En la estructura de las HA de *A. paragallinarum* se han reportado regiones hipervariables cuya antigenicidad es específica de cada serovariedad. En ensayos de inmunizaron de pollos con proteínas HA recombinantes que contenían la región hipervariable se observó una tasa de protección contra *A. paragallinarum* que varió del 83 al 100%, esto sugiere que la región hipervariable de las HA es útil como vacuna (Wu *et al.*, 2011; Sakamoto *et al.*, 2013).

En otro estudio se identificó y purificó un antígeno de HA mediante un anticuerpo monoclonal inhibidor de la hemaglutinación. La secuencia N-terminal del antígeno se identificó como una hemaglutinina de la familia de proteínas de membrana externa (OmpA) con funciones de adhesina (Hobb *et al.*, 2002). Con base en varios experimentos de hemaglutinación se propone que esta prueba depende de la expresión de las hemaglutininas y de que estas interactúen exitosamente con los receptores de los eritrocitos y que la hemaglutinación variará en dependencia del tipo de eritrocitos utilizados, por ejemplo, de aves, mamíferos o ambos (Montes *et al.*, 2016, Zepeda *et al.*, 2009).

2.1.5. Cápsula

La cápsula es uno de los componentes superficiales de algunas bacterias, la cual les permite interactuar con su entorno, es una barrera y un mecanismo de protección ambiental (Harper *et al.*, 2012). Esta cubierta celular está constituida por unidades repetidas de glucosaminoglicanos (GAG), unidas por varios tipos de enlaces glucosídicos y eso la hace compleja (Taylor & Roberts, 2005), esta diversidad permite que cada tipo de célula capsulada diferente pueda relacionarse con un huésped específico. En cepas de *A. paragallinarum* se encontraron dos grupos de genes que codifican para cápsula, uno que contiene genes para la síntesis de condroitina y otro que contiene genes para la síntesis de heparosano (Wu *et al.*, 2010).

La cápsula le proporciona a las bacterias tolerancia a la desecación, adherencia, resistencia a la fagocitosis, protección contra la muerte mediada por el complemento, resistencia a la actividad bactericida entre otras, estas características hacen que tenga un papel relevante en la virulencia. (Taylor & Roberts, 2005; Harper *et al.*, 2012). En *A. paragallinarum* se demostró que las cepas carentes de cápsula tuvieron mayor capacidad de adhesión y formación de biopelículas, pero también fueron menos virulentas (Tu *et al.*, 2015).

2.1.6. Proteasas

Las bacterias patógenas pasan por situaciones de estrés dentro del hospedero, es por ello que deben mantener un control de sus proteínas y de las proteínas del huésped; para poder lograrlo necesitan de las proteasas (Ingmer & Brondsted, 2009; Aljghami *et al.*, 2022). Las proteasas son importantes para la adquisición de nutrientes, división celular, procesamiento de proteínas, señalización, degradación de proteínas, liberación de aminoácidos, y evitar el sistema inmune del hospedero; la función que desarrollen dependerá del entorno en el que la bacteria se encuentre (Huston, 2010). De acuerdo a la composición de su sitio activo y el mecanismo de catálisis, las proteasas se clasifican en: aspártico, asparagina,

treonina, glutamato, serina, cisteína, y metaloproteasas, siendo las últimas tres las más abundantes en bacterias (Figaj *et al.*, 2019).

En bacterias de la familia *Pasteurellaceae* se han identificado metaloproteasas (García *et al.*, 2005). En *A. paragallinarum* se identificó la secreción de metaloproteasas en cultivos que carecían de suero; estas proteasas degradaron parcialmente IgG de pollo, por lo que podrían ser consideradas como un factor de virulencia, que permite a la bacteria evadir el sistema inmune del hospedero y obtener nutrientes, contribuyendo así al desarrollo de la coriza infecciosa (Rivero *et al.*, 2005).

2.1.7. Condroitin liasas

Los glucosaminoglicanos (GAG) son polisacáridos lineales complejos que se clasifican en: hialuronato, sulfato de condroitina (CS)/sulfato de dermatán (DS), sulfato de heparán/heparina y sulfato de queratán (Hernáiz & Linhardt, 2001). El sulfato de condroitina es el sustrato de las condroitin liasas, y es un heteropolisacárido compuesto por unidades repetidas de N-acetil-D-galactosamina y ácido D-glucurónico, estos disacáridos se pueden sulfatar en la posición 4 o 6 del residuo de N-acetilgalactosamina. Las clases principales de CS son CS-A (4-sulfato de condroitina), DS (CS-B) y CS-C (6-sulfato de condroitina) (Volpi, 2019). Este polisacárido se encuentra en tejidos conectivos de mamíferos, especialmente en cartílago, piel, vasos sanguíneos, ligamentos y tendones; su función es dar estructura a los órganos y tejidos, además de tener participación en funciones biológicas como la inflamación y la reparación de heridas (Volpi, 2006).

Las condroitin liasas son enzimas capaces de despolimerizar polisacáridos, y son importantes para los microorganismos patógenos en procesos de adhesión, migración y proliferación. Se han caracterizado cuatro clases de condroitina liasas, cada una de ellas actúa sobre un sustrato específico como se muestra a continuación (Hernáiz & Linhardt, 2001; Hong *et al.*, 2002).

Sustrato	Enzima
Condroitina-4-0-sulfato (sulfato de condroitina A)	<ul style="list-style-type: none"> ● Condroitin liasa ABC ● Condroitin liasa AC
Dermatán sulfato (sulfato de condroitina B)	<ul style="list-style-type: none"> ● Condroitin liasa ABC ● Condroitin liasa B
Condroitina-6-0-sulfato (sulfato de condroitina C)	<ul style="list-style-type: none"> ● Condroitin liasa ABC ● Condroitin liasa AC ● Condroitin liasa C

En investigaciones previas se determinó que las condroitin liasas ABC tienen actividad óptima a pH=7,0 y 40°C, y las condroitin liasas AC a pH=5.7-6.0 y 45-50°C. Las condroitin liasas ABC puede ser ABC endoliasa y ABC exoliasa, de igual manera las condroitin liasas AC pueden ser AC-I (endoliasa) y ACII (exoliasa) (Hong *et al.*, 2002). Estas enzimas se han descrito como factores de virulencia en microorganismos patógenos de peces como *Edwardsiella piscicida* y *Flavobacterium columnare*, entre otras (Xie *et al.*, 2005; Leung *et al.*, 2019).

2.2. Tecnologías de secuenciación de DNA genómico

Las tecnologías de secuenciación de ácidos nucleicos avanzan constantemente, dando lugar a la recopilación masiva de datos del genoma, esto ha permitido avanzar en la investigación biológica y molecular, pues brinda varias ventajas como el costo experimental bajo, rapidez y mayor alcance y profundidad de investigación a nivel de laboratorio. Las tecnologías de secuenciación se clasifican de acuerdo a la longitud de lecturas que generan, concomitantemente con la marca registrada de la tecnología utilizada (Mardis, 2013; Hu *et al.*, 2021), en este proyecto nos enfocaremos en: tecnologías de lectura corta o de “segunda generación” y tecnologías de lectura larga o de “tercera generación”, representadas por Illumina y Pacific Biosciences, respectivamente.

2.2.1. Illumina

Las tecnologías de segunda generación como es Illumina se basan en una técnica conocida como “amplificación puente”, que comienza con la fragmentación del DNA para posteriormente ligar adaptadores en cada extremo de estos fragmentos, estos adaptadores se usan como sustrato para reacciones de síntesis de amplificación repetida, en un soporte sólido que contiene secuencias de oligonucleótidos complementarias al adaptador ligado (Mardis, 2013). El DNA se somete a rondas de amplificación y durante la reacción de síntesis se incorporan y se detectan nucleótidos con una etiqueta fluorescente diferente para cada una de las cuatro bases, son estos mismos nucleótidos los que terminarán la síntesis de cada reacción, pues tienen acoplado un terminador reversible; cada que un dNTP marcado con fluorescencia se incorpora a la cadena de ácido nucleico se obtienen imágenes de la señal fluorescente resultante, las reacciones se repiten durante 300 o más rondas (Slatko *et al.*, 2018). Una ventaja de Illumina es que se pueden secuenciar ambos extremos de un fragmento de DNA, aumentando la calidad y cobertura de las lecturas-secuencias. Esta es la tecnología base por base más precisa del mercado pues solo tiene una tasa de error del 0,1%. Una desventaja es que, al generar lecturas cortas, la tarea de reconstrucción del genoma se vuelve compleja (Goodwin *et al.*, 2016; Hu *et al.*, 2021).

2.2.2. Pacific Biosciences

La secuenciación de tercera generación es capaz de secuenciar moléculas largas de DNA o RNA. Pacific Biosciences (PacBio) también conocida como SMRT (Single Molecule Real Time), permite la secuenciación de fragmentos de 30-50kb. En el método SMRT se une una DNA polimerasa con el DNA que se va a secuenciar, la reacción ocurre en un chip “SMRT Cell” que tiene muchos poros pequeños llamados ZMW (Zero Mode Waveguide). ZMW guía la energía de la luz hacia un área cuyas dimensiones son pequeñas, en comparación con la longitud de onda de la luz que ilumina, esto permite que la imagen ocurra solo en la parte inferior de la ZMW donde la DNA polimerasa, unida al DNA, incorpora cada base en una cadena creciente. (Slatko *et al.*, 2018; Hu *et al.*, 2021). Los cuatro

nucleótidos están marcados con diferentes fluoróforos enlazados con fósforo para la detección diferencial. Al unirse cada nucleótido a la polimerasa se produce un pulso de luz característico, este pulso es capturado en una “película”. Los pulsos de luz se interpretan como secuencia de nucleótidos y esta secuencia es llamada “Lectura larga continua” (Hu *et al.*, 2021).

Dentro de las ventajas de la secuenciación PacBio es que puede medir la tasa de adición de cada nucleótido durante la síntesis, denominada Inter Pulse Duration (IPD). Los nucleótidos con metilaciones de adenina y citosina cambian la IPD y se pueden identificar como bases modificadas, dando así ventajas en estudios epigenéticos. Otra ventaja es que proporciona lecturas largas que facilitan el ensamble del genoma. Sin embargo, una desventaja es que tiene una tasa de error alta, aunque esto se puede corregir si se combinan lecturas PacBio con lecturas de Illumina para aumentar la precisión (Rhoads & Fai 2015; Slatko *et al.*, 2018).

2.3. Análisis y comparación de genomas

De forma breve se puede decir que un genoma es la información genética total de un organismo o célula, que se encuentra almacenada en los ácidos nucleicos. Para poder conocer de manera precisa esta secuencia, se han desarrollado tecnologías de secuenciación que cada día son más eficientes y precisas, algunas de ellas mencionada previamente (Goldman & Landweber, 2016; Giani *et al.*, 2020).

El resultado producido por los aparatos de secuenciación son secuencias conocidas como lecturas. Estas deben ser ordenadas por solapamientos de similitud para deducir la secuencia completa del genoma, a este proceso se le conoce como ensamble del genoma. La reconstrucción o ensamble debe anotarse apropiadamente para conocer la función de los genes (Giani *et al.*, 2020). El análisis de la anotación brinda información bioquímica, molecular y funcional de las células, y en la microbiología, sobre cómo las bacterias se comportan metabólicamente, evolucionan e interactúan con el ambiente, entre sí y con sus

anfitriones (Loman & Pallen, 2015; Heather & Chain, 2016). Debido a la relevancia de la información que quiere obtener, es importante que la secuencia se reconstruya lo más completa y libre de errores posibles, actualmente esto se logra, pues las tecnologías de secuenciación y sus resultados son cada vez de mejor calidad (Loman & Pallen, 2015; Giani *et al.*, 2020).

Con genomas bien reconstruidos como contigs, scaffolds o totalmente completos, se puede hacer análisis de genómica comparativa y conocer las variaciones entre grupos de genes o subsistemas moleculares, por ejemplo, de patogénesis. La genómica comparativa es un campo de investigación que se ha desarrollado con herramientas computacionales que permiten conocer genes que se conservan o son comunes entre especies, linajes o cepas, y de esta manera comprender los cambios evolutivos entre microorganismos, también nos permite conocer los genes únicos de cada organismo o sus características moleculares particulares (Sivashankari & Shanmughavel, 2007; Touchman, 2010).

El conocimiento molecular de rasgos de virulencia, asociado a la disponibilidad de genomas ayuda al diseño de nuevas herramientas, medicamentos y vacunas para combatir enfermedades infecciosas como la CI (Loman & Pallen, 2015).

III. JUSTIFICACIÓN

Avibacterium paragallinarum es un microorganismo con baja proliferación en medios de cultivos sintéticos a pesar de proporcionarle nutrimentos especiales que esta bacteria no puede sintetizar; esto hace largo y costoso su cultivo in vitro. Forma colonias muy pequeñas que en primoaislamiento pueden ser fácilmente cubiertas por otras bacterias del microbioma presente en la muestra. Esta bacteria produce coriza infecciosa, y puede afectar a la industria avícola produciendo importantes pérdidas económicas. Por ello es de suma importancia conocer los mecanismos de patogenicidad y virulencia, y los genes involucrados, que permitan controlar de forma eficiente la coriza infecciosa utilizándolos como nuevos inmunógenos en las vacunas o bacterinas. Actualmente se conocen un grupo de estos genes y es probable que existan genes no estudiados o no descubiertos que pudieran ser un blanco para combatir la enfermedad. Es poco en trabajo genético que se ha hecho hasta la fecha porque *A. paragallinarum* es un organismo recalcitrante a la transformación y muy sensible a los procedimientos convencionales de transformación genética. Actualmente la metodología de secuenciación masiva de ácidos nucleicos representa una herramienta poderosa para solventar los problemas de investigación genética de *A. paragallinarum*, así podremos conocer todo el contenido de genes en corto tiempo y posiblemente el papel que estos juegan en la patogenicidad producida por la bacteria. Consecuentemente, la interpretación de esta información implica un importante reto que debe combinar métodos computacionales y experimentales.

IV. OBJETIVOS

Objetivo general

Caracterizar molecularmente nuevos rasgos y genes de virulencia a partir de aislados recientes de *Avibacterium paragallinarum* obtenidos de aves de granja.

Objetivos específicos

- ❖ Obtener cepas de *Avibacterium paragallinarum* de aves de granja.
- ❖ Secuenciar los genomas de 18 cepas, ensamblar, anotar y hacer la comparación de los genomas.
- ❖ Predecir genes involucrados en la virulencia y genes nuevos en *A. paragallinarum*.
- ❖ Evaluar los genes predichos potencialmente relacionados con virulencia (la utilización de diferentes fuentes de carbono y la resistencia a antibióticos de *A. paragallinarum*
- ❖ Proponer primers que sean útiles para hacer estudios con PCR de genes potencialmente relacionados con virulencia, que codifican para las toxinas AvxA, CDT y genes que codifican para fimbrias.

V. MATERIALES Y MÉTODOS

5.1. Cepas bacterianas

Tabla 1. Cepas de *A. paragallinarum* secuenciadas en este proyecto.

Cepas	Fenotipo	Medios de cultivo	Origen
MIMOCE-ICUAP			
Aavg1	virulento	Agar Sangre (BHI agar	BIOVETSA
Aavg2	virulento	enriquecido con 5% sangre	BIOVETSA
Aavg3	virulento	de carnero) con estría	BIOVETSA
Aavg4	virulento	nodriza de <i>Staphylococcus</i>	BIOVETSA
Aavg5	virulento	<i>sp.</i>	BIOVETSA
Aavg6	virulento	Agar Infusión Cerebro	BIOVETSA
Aavg7	virulento	Corazón (BHI BD Bioxon)	BIOVETSA
Aavg8	virulento	enriquecido con suero de	BIOVETSA
Aavg9	virulento	caballo (Biowest) 5%,	BIOVETSA
Aavg10	virulento	NADH (Sigma Aldrich) 10	BIOVETSA
ESV-135	virulento	µg/µl y sulfato de	<i>FES Iztacala UNAM</i>
AVPG2015	virulento	condroitina 4 mg/ml.	<i>FES Iztacala UNAM</i>
CL	virulento	Medio Líquido	<i>FES Iztacala UNAM</i>
CT	virulento	Infusión Cerebro Corazón	<i>FES Iztacala UNAM</i>
5lab	virulento	(BHI BD Bioxon)	<i>FES Iztacala UNAM</i>
6lab	virulento	enriquecido con suero de	<i>FES Iztacala UNAM</i>
	virulento	caballo (Biowest) 5% y	
		NADH (Sigma Aldrich) 10	
		µg/µl. Opcionalmente	
		enriquecido con sulfato de	
		condroitina.	

5.2. Condiciones de crecimiento y extracción de DNA genómico

Las cepas de *A. paragallinarum* se sembraron en agar sangre con estría nodriza de *Staphylococcus sp.* he incubaron a 37°C por 24 h, de estos cultivos se tomó una asada y se inoculó en 3 ml de medio BHI enriquecido (suero de caballo 5% y NADH 10 µg/µl), se incubó a 37°C por 18 h, de este precultivo se agregaron 200 µl a 20 ml de medio BHI enriquecido para posteriormente incubar a 37°C por 14 h.

Los cultivos se separaron en dos alícuotas por cepa (10 ml cada una), ambas alícuotas se centrifugaron a 11000 rpm durante 1 min y se desechó el sobrenadante, a las pastillas se agregaron 400 µl de buffer de lisis y con ayuda de una punta se resuspendieron. Con un palillo estéril se agregó una pizca de lisozima (Sigma Aldrich®) y se incubó a 37°C por 30 min, transcurrido el tiempo de incubación se añadieron 50 µl de sarcosil (Sigma Aldrich®) al 20% y se agitó vigorosamente hasta completar la lisis, a esta solución se le añadió un volumen de fenol-tris (Sigma Aldrich) y se mezcló por inversión varias veces. La suspensión se centrifugó a 11000 rpm durante 10 min y la fase acuosa se transfirió a un microtubo de 1.5 ml, a la fase acuosa se añadió un volumen de cloroformo (J.T.Baker®) y se agitó varias veces, se centrifugó a 11000 rpm por 10 min (este paso se repitió dos veces), el sobrenadante se pasó a un microtubo de 1.5 ml y se agregó un volumen de etanol absoluto (J.T.Baker®) y 0.5 volúmenes de acetato de amonio 7.5 M, en DNA se dejó precipitar a -20°C por 1 h y posteriormente se centrifugó a 11000 rpm durante 10 min. Se decantó el sobrenadante y la pastilla se lavó con 200 µl de etanol al 70%, se desechó el etanol y la pastilla se resuspendió en 200 µl de agua inyectable estéril (PiSA®).

Al DNA se agregaron 15 µg de RNasa (Merck) y se incubó a 37°C por 2 h, transcurrido el tiempo se añadió 1 volumen de Fenol-Cloroformo-Alcohol isoamílico, se mezcló hasta homogeneizar y se centrifugó a 11000 rpm durante 10 min, se separó la fase superior y a esta se añadió 1 volumen de etanol absoluto, se dejó precipitando a -20°C por 1 h, transcurrido el tiempo se centrifugó a 11000 rpm por 10 min, se decantó el sobrenadante y a la pastilla se lavó con 200 µl de etanol al 70%, se desechó el etanol y la pastilla se resuspendió en 100 µl de agua

inyectable estéril (PiSA®). El DNA se guardó a -20°C hasta su uso.

5.2.1. Cuantificación de DNA genómico

En una celda de cuarzo se agregaron 990 µl de agua destilada y 10 µl de DNA genómico, con ayuda de una punta se homogeneizó y se midió la DO a 260 nm. Para calcular la concentración de DNA en ng/µl usamos la siguiente fórmula:

$$(FC) (A) (FD) = \text{ng}/\mu\text{l} \quad (\text{FC} = \text{factor de conversión (50)}, A = \text{Absorbancia (260 nm)}, FD = \text{Factor de dilución 1:100})$$

5.2.2. Relación 260/280 y 260/230 de DNA

Para descartar que el DNA estuviera contaminado con fenol o proteínas, se determinó la relación 260/280 y 260/230. En una celda de cuarzo se agregaron 990 µl de agua destilada y 10 µl de DNA genómico, con ayuda de una punta se homogeneizó y se midió la DO a 230 nm, 260 nm y 280 nm, se sumaron las DO de la lectura a 260 nm y 230 nm y se dividen entre dos, de la misma manera se hizo con las lecturas 260 nm y 280 nm.

5.3. Datos generales

Tabla 2. Cepas de *A. paragallinarum* secuenciadas con diferentes tecnologías de secuenciación.

Cepas	Tecnología de secuenciación		
Avg1	Ilumina MiSeq		
Avg2	Ilumina MiSeq		
Avg3	Ilumina MiSeq	PacBio	
Avg4	Ilumina MiSeq		
Avg5	Ilumina MiSeq		
Avg6	Ilumina MiSeq	PacBio	
Avg7	Ilumina MiSeq	PacBio	
Avg8	Ilumina MiSeq	PacBio	
Avg9	Ilumina MiSeq	PacBio	
Avg10	Ilumina MiSeq	PacBio	
ESV-135	Ilumina MiSeq	PacBio	
AVPG 2015		PacBio	454
CL	Ilumina MiSeq	PacBio	454
CT	Ilumina MiSeq	PacBio	454
5lab	Ilumina MiSeq	PacBio	454
6lab	Ilumina MiSeq	PacBio	454

5.4. Análisis y mejoramiento de la calidad de secuencias de Illumina MiSeq

Inicialmente los archivos de las secuencias Illumina y PacBio en formato sff y bam respectivamente, se convirtieron a formato fasta o fastq para poder usarse con diferentes herramientas bioinformáticas. Para poder hacer la conversión se utilizó la herramienta Bamtools (Barnett *et al.*, 2011).

Posteriormente se usó la herramienta fastQC versión 0.11.3 (Andrews, 2010) para tener una visión general y rápida de la calidad de los datos de secuenciación del genoma de las cepas de *A. paragallinarum*. Evaluados los datos, se removieron las secuencias que de acuerdo a fastQC eran de mala calidad, y también los adaptadores, para esto se utilizó Trimomatic versión 0.36 (Bolger *et al.*, 2014).

5.5. Ensamble de genomas

5.5.1. Illumina MiSeq

Para el ensamble de genomas con secuencias de Illumina se usaron tres herramientas: MIRA Mimicking Intelligent Read Assembly (Chevreux *et al.*, 1999), SPAdes 3.12.0 (Nurk *et al.*, 2013) y ABySS 2.1.0 (Jackman *et al.*, 2013). Para poder visualizar información adicional como N50 y porcentaje de GC fue necesario usar la herramienta Quast (Gurevich *et al.*, 2013).

5.5.2. PacBio

Para ensamblar de novo secuencias PacBio se usó la herramienta Canu v1.5 (Koren *et al.*, 2017), esta herramienta lleva a cabo tres procesos: corrección, trimming y ensamble.

El ensamble híbrido (secuencias Illumina y PacBio) de los 12 genomas de *A. paragallinarum* se hizo usando la herramienta SPAdes 3.12.0 (Antipov *et al.*, 2016).

5.6. Anotación en Rapid Annotation using Subsystem Technology (RAST)

Se creó una cuenta en RAST (Aziz *et al.*, 2008, Overbeek *et al.*, 2014, Brettin *et al.*, 2015), se subieron los contigs o scaffolds que resultaron de los ensambles de secuencias de Illumina y PacBio para hacer una anotación genómica de alta calidad. Con los datos obtenidos se hizo una búsqueda de factores de virulencia nuevos y previamente reportados en *A. paragallinarum*. Para la búsqueda de factores de virulencia también se hizo una búsqueda en VFDB (Virulence Factor Database), usando como referencia *Escherichia coli*, pues también es una bacteria Gram negativa y *Haemophilus influenzae*, que pertenece a la familia Pasteurellaceae.

5.7. Resistencia antimicrobiana

5.7.1. Gel de lisis *in situ* para la visualización de plásmidos de alto peso molecular

Para determinar si *A. paragallinarum* tenía plásmidos de alto peso molecular utilizamos la técnica de lisis *in situ* (Eckhardt, 1978) modificada como se indica a continuación:

De cultivos frescos de 18 h se tomaron 200 μ l y se centrifugaron a 11000 rpm durante 3 min, se retiró el sobrenadante y a la pastilla se agregaron 20 μ l de Tampón I frío y se resuspendió con ayuda de una punta. Se agregaron 20 μ l de solución de lisis recién preparada y 5 mg/ml de proteasa, con ayuda de una punta se resuspendió y se incubó a 37°C por 10 min, transcurrido el tiempo de incubación se agregaron 10 μ l de la mezcla en un pozo de gel de agarosa al 0.8%, en el mismo pozo se agregaron 10 μ l de Tampón II. Se agregó TBE 1X sin que cubriera el pozo y se dejó correr unos minutos a 30 volts, posteriormente se agregó TBE 1X hasta que cubriera el gel y se dejó correr a 90 volts hasta que el colorante llegó al borde del gel. El gen se tiñó con bromuro de etidio por 10 min y posteriormente el exceso se eliminó con agua destilada.

5.7.2. Antibiogramas por difusión en disco

De cultivos de 14 horas se tomaron 200 µl y se sembraron en placas de medio BHI con NADH (10 µg/ml), el cultivo se distribuyó por toda la placa con ayuda de un asa bacteriológica; la placa se dejó secar por 10 min y posteriormente con ayuda de pinzas estériles se colocaron discos individuales con antibiótico a probar en el centro de la placa. Las placas se incubaron a 37°C por 24 h en microaerofilia.

Los antibióticos que se usaron fueron: Ampicilina (10 µg/ml), Tetraciclina (30 µg/ml), Amikacina (30 µg/ml), Cloranfenicol (30 µg/ml), Gentamicina (10 µg/ml) y Estreptomina (30 µg/ml), las concentraciones usadas fueron de acuerdo a BIO RAD.

VI. RESULTADOS

6.1. Extracción de DNA genómico

Se extrajo DNA genómico de buena calidad de todas las cepas seleccionadas para ser secuenciadas con la plataforma de Illumina MiSeq y PacBio. La buena calidad del DNA era indispensable para una secuenciación eficiente (**Figura 1**).

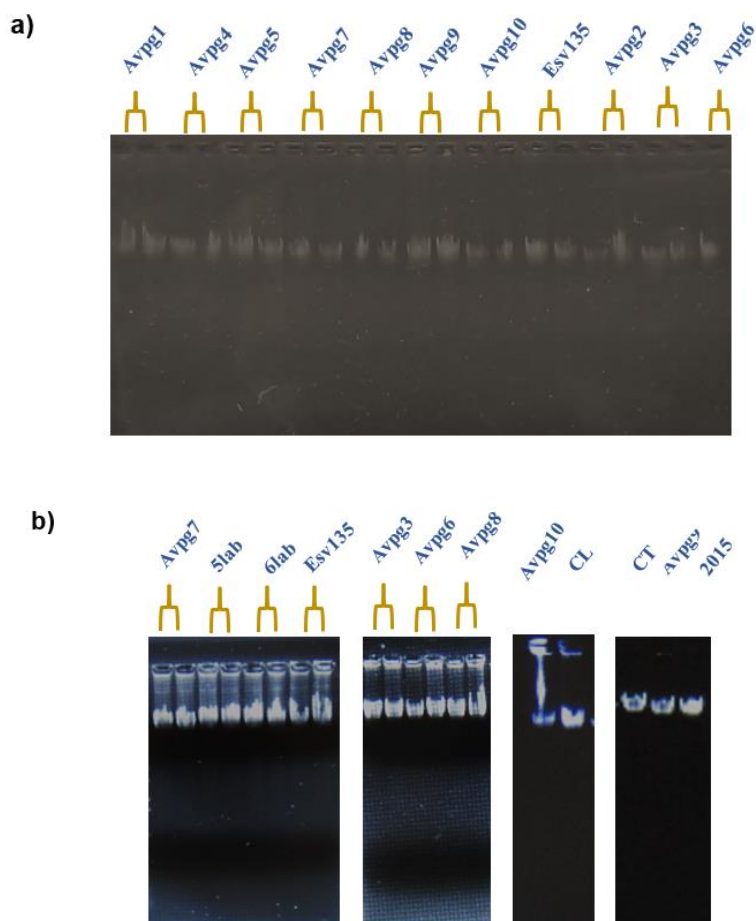


Figura 1. DNA genómico obtenido para secuenciar. Gel de agarosa 0.8%, a) DNA de 11 cepas *A. paragallinarum* elegidas para secuenciar con la plataforma Illumina, b) DNA de 12 cepas de *A. paragallinarum* elegidas para secuenciar con la plataforma PacBio.

6.1.1. Cuantificación de DNA genómico

Se determinó la concentración de DNA genómico obtenido para asegurar que la muestra fuera suficiente para la secuenciación, 100-500ng para Illumina y 25µg para PacBio. Para la secuenciación con Illumina MiSeq se obtuvieron entre 20ng/µl y 30 ng/µl para cada cepa, para la secuenciación con PacBio se obtuvieron entre 475ng/µl y 2685ng/µl.

6.1.2. Relación 260/280 y 260/230 de DNA

Para conocer la pureza del DNA, se midió la relación 260/280 para descartar la presencia de restos de proteínas y la relación 260/230 para descartar la presencia de sales o fenol. El DNA secuenciado con Illumina tuvo relación 260/230 de 1.3 a 1.7 y 260/280 de 1.6 a 1.9, estos valores sugieren que algunas muestras pudieron tener restos de proteínas o fenol, por otro lado, el DNA secuenciado con PacBio tuvo una relación 260/230 de 2.1-2.4 y una relación 260/280 de 1.7-1.9, la calidad de estas muestras fue optima.

6.2. Análisis de calidad en los datos de secuencias de Illumina MiSeq

Con la finalidad de conocer la calidad de las secuencias de Illumina, se hizo un análisis en fastQC y se identificaron los parámetros que debían ser mejorados y que porcentaje de lecturas eran de mala calidad, un ejemplo de ello se muestra en la **Figura 2**.

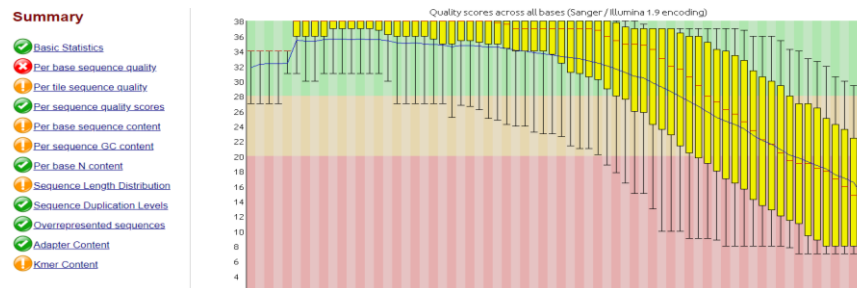


Figura 2. Gráfico de barras y bigotes. Representación de parámetros del análisis de calidad de secuencias de Illumina. Antes de mejorar la calidad se observaban barras dentro del área roja y naranja, lo ideal es que la mayoría se encuentren en el área verde.

Con la información que se obtuvo del análisis de calidad se pudieron hacer cambios mediante Trimmomatic, herramienta que nos permite el recorte y filtrado de calidad de las secuencias, estos cambios ayudaron a que el ensamble pudiera ser más eficiente y que se generara menor número de contigs. Se evaluó nuevamente la calidad de las secuencias como se muestra en la **Figura 3**.

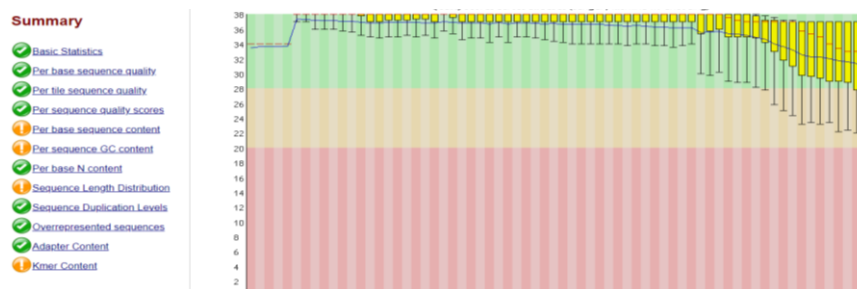


Figura 3. Gráfico de barras y bigotes de secuencias editadas. Representación de parámetros del análisis de calidad de secuencias de Illumina corregidos con Trimmomatic. Se observa que la mayoría de las barras se encuentran dentro del área verde que indica una buena calidad.

6.3. Ensamble de genomas con secuencias de Illumina MiSeq

Con el objetivo de tener ensamblajes de genomas con el menor número de contigs, se usaron tres herramientas para hacer el ensamblaje de novo de genomas de *A. paragallinarum*. Se hizo una comparación de los resultados generados por las tres herramientas de ensamblaje, y se encontró que la herramienta más eficiente que permitía tener un genoma menos fragmentado era el de ABySS (Tabla 4), seguido por SPAdes (Tabla 5), mientras que con MIRA (Tabla 6) el genoma resultó más fragmentado, en todos los casos el porcentaje de GC fue entre 40.79 y 40.91, y el tamaño del genoma fue de 2.3Mb a 2.6Mb.

Tabla 3. Datos generales de ensamblaje de genoma de 11 cepas de *A. paragallinarum* con ABySS.

ABySS				
Cepas	Scaffolds	Tamaño	%GC	N50
Avpg1	69	2586466	40.88	171707
Avpg2	65	2431607	40.89	193862
Avpg3	56	2429169	40.91	193744
Avpg4	56	2435002	40.89	193727
Avpg5	54	2435999	40.90	193735
Avpg6	55	2432557	40.90	193379
Avpg7	63	2434395	40.90	193728
Avpg8	60	2433705	40.89	193739
Avpg9	58	2434129	40.90	188934
Avpg10	58	2436329	40.90	146207
Esv135	68	2542387	40.88	193683

Tabla 4. Datos generales de ensamble de genoma de 11 cepas de *A. paragallinarum* con SPAdes.

SPAdes				
Cepas	Scaffolds	Tamaño	%GC	N50
Avgp1	76	2522445	40.79	192521
Avgp2	93	2397413	40.82	193737
Avgp3	76	2385463	40.84	193714
Avgp4	63	2383587	40.83	193438
Avgp5	62	2382882	40.83	193442
Avgp6	82	2391771	40.83	193437
Avgp7	62	2382121	40.83	193437
Avgp8	69	2385431	40.83	193442
Avgp9	65	2383314	40.84	193894
Avgp10	72	2388404	40.83	193892
Esv135	74	2479040	40.80	193445

Tabla 5. Datos generales de ensamble de genoma de 11 cepas de *A. paragallinarum* con MIRA.

MIRA				
Cepas	Contigs	Tamaño	%GC	N50
Avgp1	144	2667203	40.8	146122
Avgp2	151	2514890	40.8	137037
Avgp3	144	2506523	40.8	122065
Avgp4	133	2495890	40.8	137057
Avgp5	141	2504276	40.8	137095
Avgp6	147	2509110	40.8	113994
Avgp7	146	2507149	40.8	130104
Avgp8	149	2515932	40.8	137132
Avgp9	150	2516756	40.8	137013
Avgp10	129	2501256	40.8	146169
Esv135	160	2624156	40.8	124542

6.4. Ensamble de genomas con secuencias PacBio

6.4.1. Ensamblados de novo

Los ensamblados que se hicieron con secuencias de PacBio generaron menor número de Scaffolds respecto a los ensamblados con secuencias de Illumina, como se puede ver en la **Tabla 7**, en algunos casos como el de las cepas Esv135, Avpg6, Avpg7 y 5lab el total del genoma está en un solo Scaffold. El contenido de GC y el tamaño del genoma están entre los valores esperados para *A. paragallinarum*.

Tabla 6. Datos generales del ensamblado de novo de 10 cepas de *A. paragallinarum* con secuencias de PacBio.

Canu				
Cepas	Scaffolds	Tamaño	%GC	N50
Avpg3	5	2591072	40.9	1335489
Avpg6	1	2427866	40.9	2427866
Avpg7	1	2428694	40.9	2428694
Avpg8	4	2493260	40.9	9931
Avpg9	14	2552673	40.9	716773
Avpg10	5	2455042	40.9	1968046
Esv135	1	2521134	40.9	2521134
2015	1	2532404	40.9	2532404
CL				
CT	Genomas no reconstruidos			
5lab	1	2429472	40.9	2429472
6lab	4	2416389	40.9	1087917

Los genomas circularizados de las cepas 2015 y Esv135, se registraron en la base de datos GenBank, los números de acceso son CP058307.1 y CP050316.1 respectivamente.

6.4.2. Ensamblés híbridos

Los ensamblés híbridos generaron un mayor número de scaffolds, aunque con la mitad de las cepas se obtuvo un scaffold que contiene casi la totalidad del genoma (Tabla 8). El contenido de GC y el tamaño del genoma oscilan alrededor de 40.85% y 2.43 Mb respectivamente, valores esperados para *A. paragallinarum*. Estos ensamblés se hicieron con resultados de secuenciación provenientes de tecnologías Illumina y PacBio.

Tabla 7. Datos generales del ensamble híbrido de 8 cepas de *A. paragallinarum*.

SPAdes				
Cepas	Scaffolds	Tamaño	%GC	N50
Avg3	14	2406550	40.9	1275021
Avg6	31	2434094	40.92	2418337
Avg7	13	2414679	40.92	2409304
Avg8	14	2416945	40.90	2411887
Avg9	15	2395799	40.86	716773
Avg10	28	2443785	40.85	2417978
Esv135	16	2509783	40.86	811295

6.5. Anotación en Rapid Annotation using Subsystem Technology (RAST)

Para poder conocer los genes presentes en los genomas reconstruidos, los scaffolds que resultaron de cada uno de los ensamblajes se analizaron en la plataforma de RAST, se buscaron marcos de lectura abiertos a lo largo de toda la secuencia, y se enviaron a las bases de datos para identificar y anotar a que genes correspondían. Cuando esta información estuvo disponible nos enfocamos a buscar genes asociados a metabolismo y a virulencia entre otros. Esta información fue utilizada para hacer comparaciones de secuencias y de contenido de genes en el genoma de las diferentes cepas.

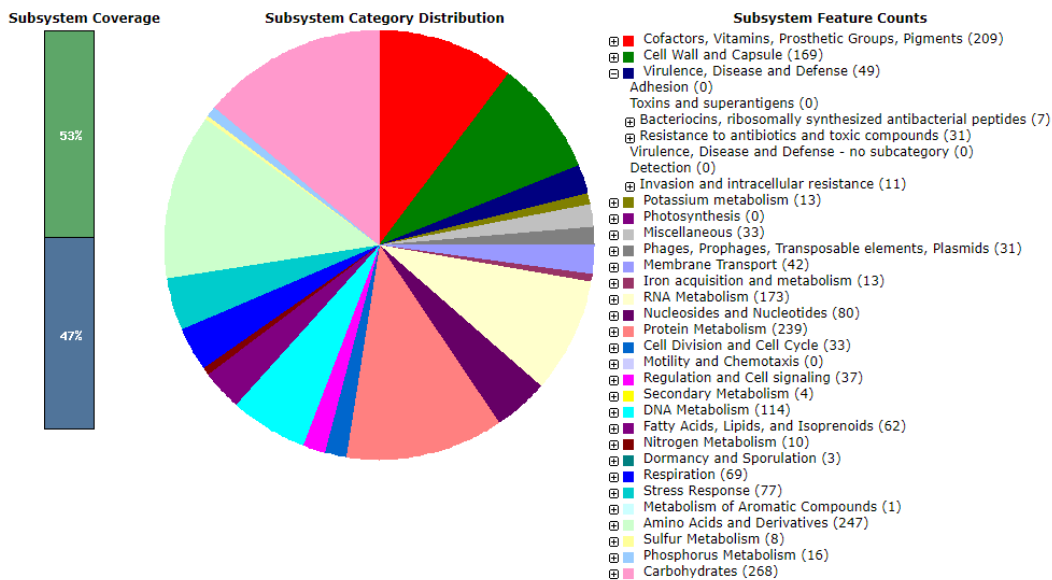


Figura 4. Clasificación de los genes anotados en RAST. Representación gráfica del contenido y distribución de genes en diferentes categorías de una cepa de *A. paragallinarum* (Esv135).

En las cepas ya anotadas se buscaron genes asociados a la virulencia ya descritos anteriormente en *A. paragallinarum*, se encontró que en todas las cepas estaban presentes los genes *cdtB* (Toxina Distensora Citoletal), *avxA* (Toxina formadora de poros) y dos grupos de genes involucrados en la formación de fimbrias (Usher). Una comparación filogenética se hizo en Clustal Omega.

cdtB

En todas las cepas de *A. paragallinarum* se encontró el operón *cdtABC*; con la secuencia correspondiente al gen estructural *cdtB* se hizo una comparación, usando secuencias de todas las cepas de esta investigación y de algunas secuencias de las cepas reportadas en otros países, 72 (Perú), 221 (China) y JF4211 (Suiza), se encontró un porcentaje de identidad del 99.88% al 100% (Figura 5). El gen *cdtB* se encuentra altamente conservado en cepas mexicanas y cepas de otros países.

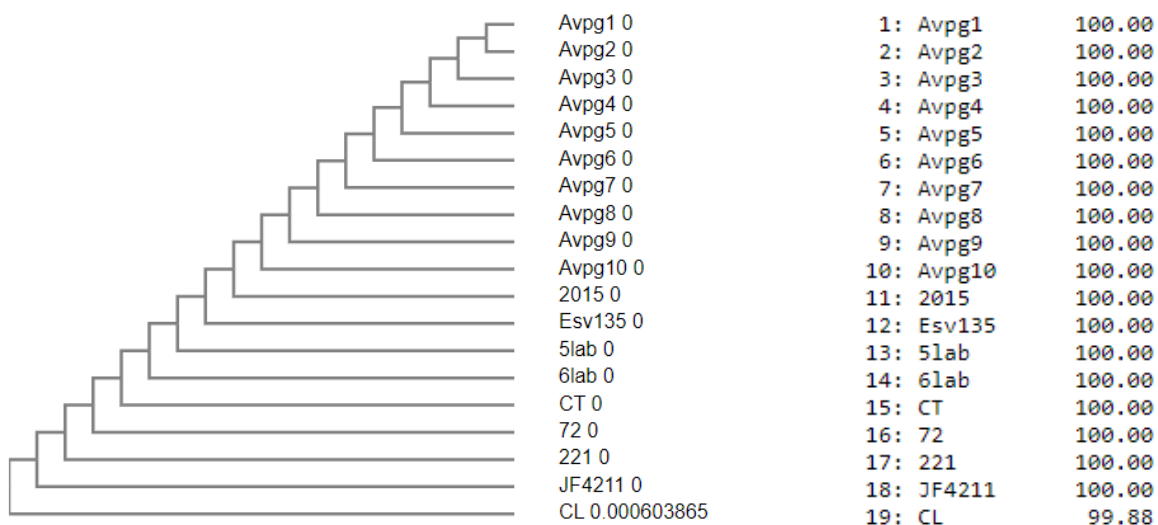


Figura 5. Comparación filogenética y porcentaje de identidad del gen *cdtB* en cepas de *A. paragallinarum*.

avxA

El operón *avxCABD* se encontró en todas las cepas de *A. paragallinarum* usadas en esta investigación, sin embargo, solo se hizo la comparación filogenética usando la secuencia del gen *avxA*, pues este es el gen estructural, se hizo una comparación incluyendo las secuencias de los genes de las cepas 72 (Perú), 221 (China) y JF4211 (Suiza), se encontró que es un gen con un porcentaje de identidad del 93.06% al 100% (Figura 6). Es un gen conservado entre cepas mexicanas y cepas de otros países, sin embargo, muestra menos identidad que *cdtB*.

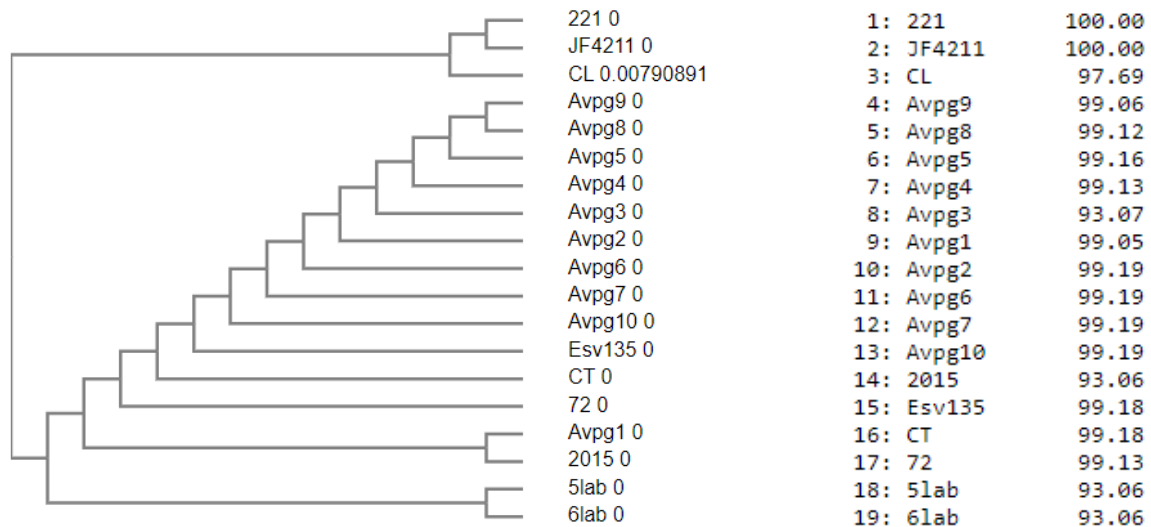


Figura 6. Comparación filogenética y porcentaje de identidad del gen *avxA* en cepas de *A. paragallinarum*.

Genes que codifican para proteínas fimbriales (*ushI* y *ushII*)

Se encontró que en todas las cepas de *A. paragallinarum* están presentes dos operones fimbriales, ambos integrado por los genes que codifican para una proteína estructural, una adhesina y un sistema chaperona-usher, las secuencias entre ambos operones son diferentes. A continuación, se muestra la comparación entre secuencias de ambos genes que codifican para la proteína integral de membrana usher (*ushI* y *ushII*).

La comparación también se hizo con secuencias de los genes usher de la cepa 72 (Perú), en el gen *ushI* se encontró un porcentaje de identidad del 98.56% al 100% (Figura 7) y en el gen *ushII* de 64.97% al 100% (Figura 8). Los genes del mismo grupo se mantienen conservados, sin embargo, son diferentes entre sí, a pesar de encontrarse en la misma cepa y tener la misma función. Los genes de las cepas 221 (China) y JF4211 (Suiza) se excluyeron de este alineamiento pues se encuentran fragmentados.

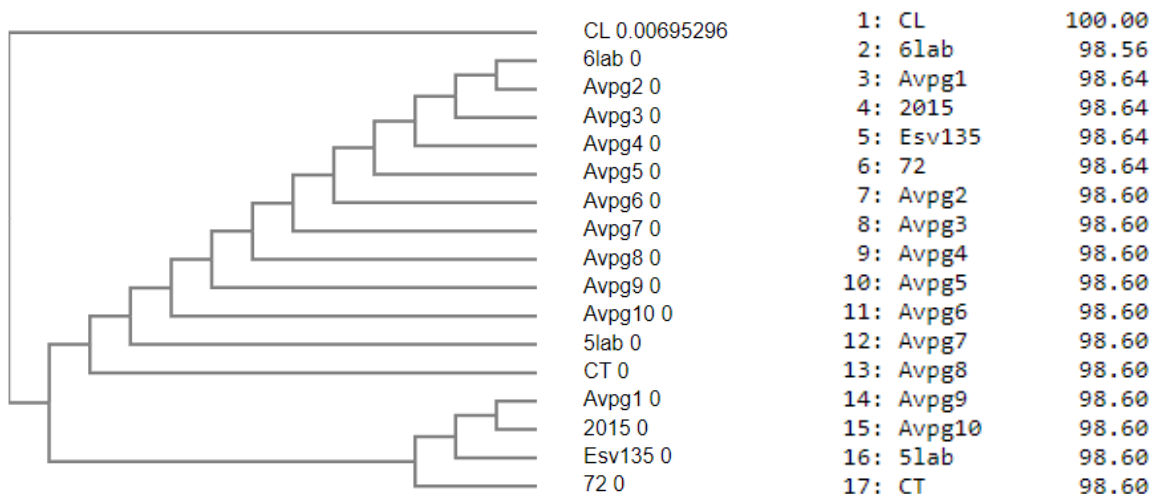


Figura 7. Comparación filogenética y porcentaje de identidad de la secuencia del gen *ushI* en cepas de *A. paragallinarum*.

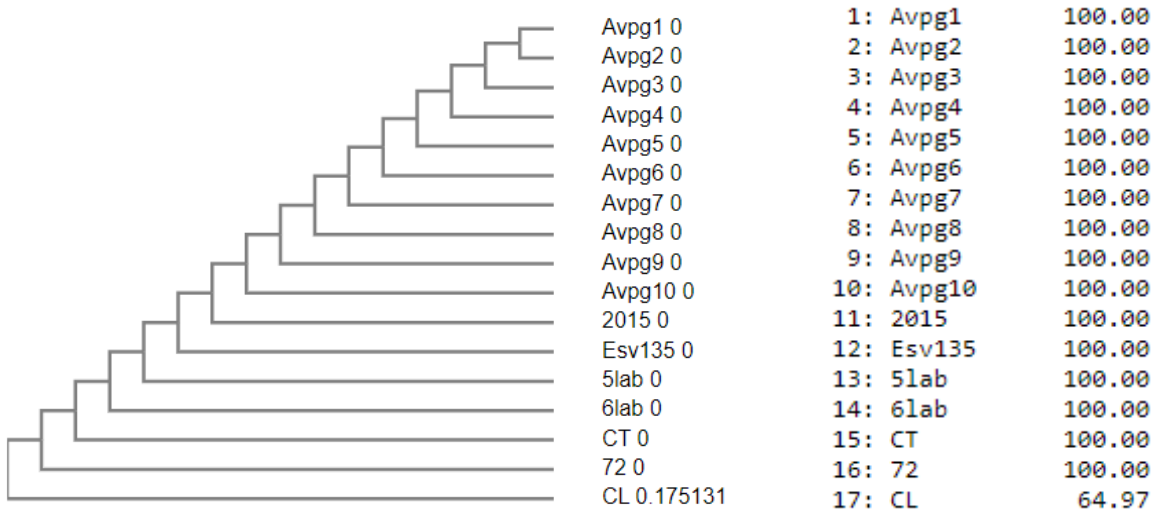


Figura 8. Comparación filogenética y porcentaje de identidad de la secuencia del gen *ushII* en cepas de *A. paragallinarum*.

6.6. Comparación lineal y filogenética de cepas de *A. paragallinarum*

Las comparaciones filogenéticas hechas con las secuencias de los genes *cdtB*, *avxA*, *ushI* y *ushII*, mostraron que las cepas CL, JF4211 y 221 se encuentran separadas del resto de las cepas, por esta razón se hizo una comparación lineal entre estas tres cepas (Figura 9). Se muestra con líneas color gris oscuro algunas regiones conservadas entre estos genomas.

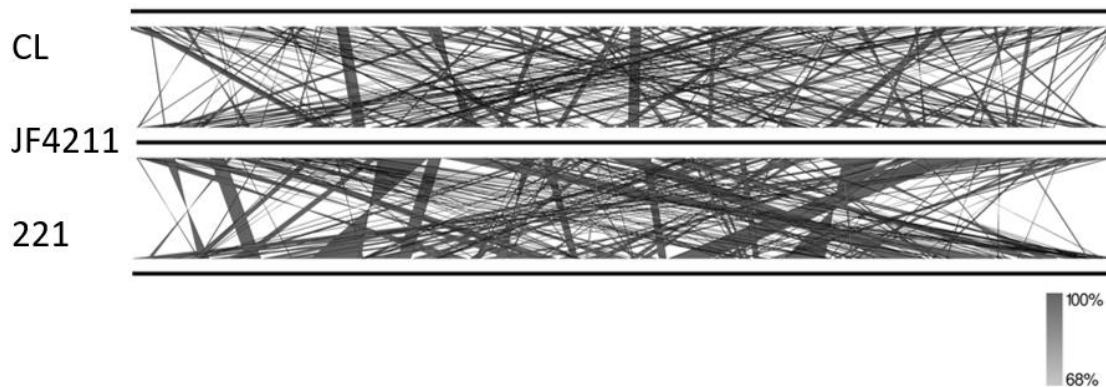


Figura 9. Comparación lineal de tres cepas de *A. paragallinarum*. La imagen muestra las regiones de identidad entre los genomas de cepas CL, JF4211 y 221 de *A. paragallinarum*. Las líneas horizontales negras muestran la comparación a lo largo del genoma y las líneas cruzadas entre los genomas indican tramos de secuencias homólogas. Imagen obtenida con el software Easyfig.

También se muestra una comparación lineal de los genomas de cepas de *A. paragallinarum* representativas de diferentes países, las cepas comparadas fueron: FARPER (Perú), AVPG2015 (México), Esv135 (México), ADLAP (USA), p4chr1 (China), AP02 (USA), AG21 (USA), CL (México) y la cepa GAESV200 de *Gallibacterium anatis* (México). La comparación mostró un número significativo de reordenamientos cromosómicos (**Figura 10**). Identificamos que las dos cepas mexicanas AVPG2015 y Esv135 son las que tienen más homología entre sí, contrario a lo que se observó al hacer la comparación con la cepa CL y GAESV200 donde se observaron el mayor número de reordenamientos.

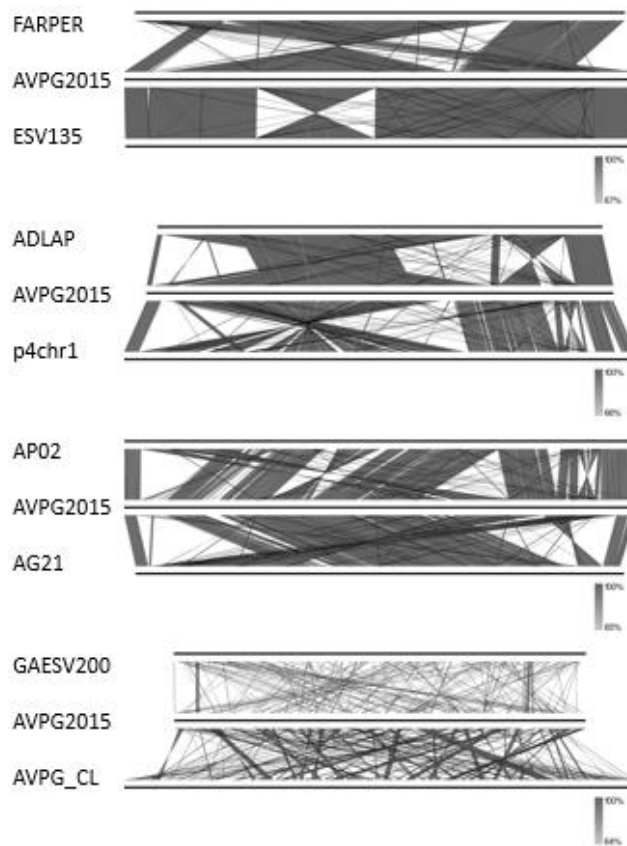


Figura 10. Comparación lineal de cepas de *A. paragallinarum*. La imagen muestra las regiones de identidad entre los genomas de cepas de *A. paragallinarum* y una cepa de *G. anatis*. Las líneas horizontales negras muestran la comparación a lo largo del genoma y las líneas cruzadas entre los genomas indican tramos de secuencias homólogas. Imagen obtenida con el software Easyfig.

También se hizo una comparación filogenética de las 44 cepas disponibles en GenBank incluidas las cepas 2015 y Esv135 que resultaron de esta investigación, el análisis se hizo con ocho genes housekeeping concatenados (*adk*, *atpD*, *fumC*, *gyrB*, *if2*, *mdh*, *recN*, *gtpR*) usados para Multi Locus Sequence Typing (MLST) (Figura 11). Podemos observar que varias cepas del mismo país se agruparon en el mismo clado del árbol, lo que sugiere que son homólogas y que tienen una relación genética similar, aunque también hubo excepciones como la cepa mexicana CL que se encuentra más cercana a una cepa proveniente de Sudáfrica y una cepa proveniente de China, las cepas mexicanas se encuentran más cerca de las cepas estadounidenses ADLAP.

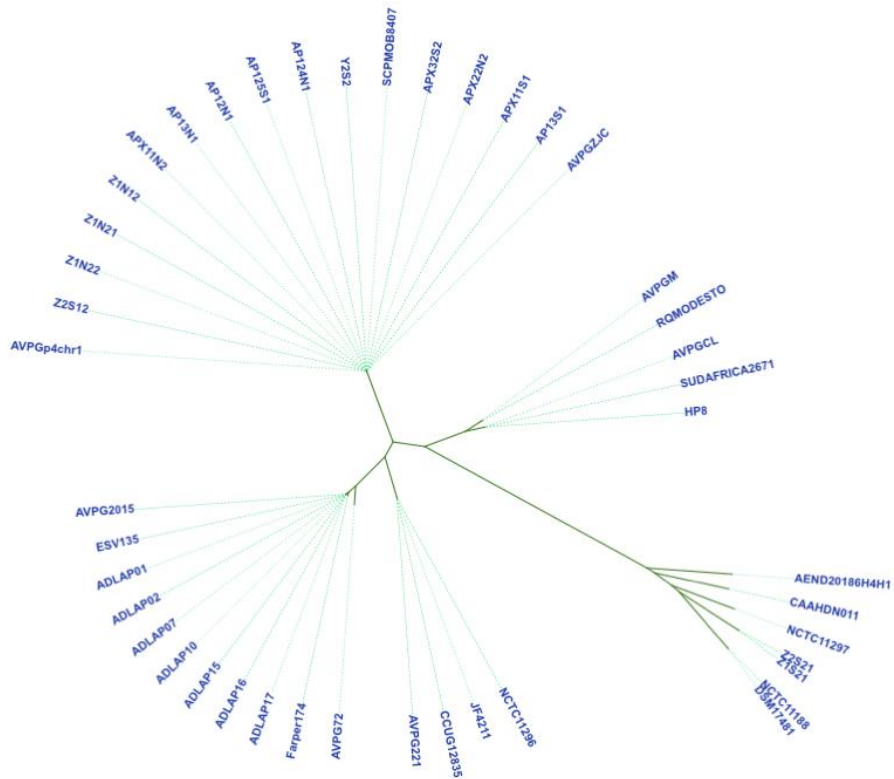


Figura 11. Agrupación filogenética de cepas de *A. paragallinarum* basado en MLST. Este filograma fue obtenido por los programas en línea kaling y Itol. Se observa la formación de varios clados, generalmente conformados por cepas del mismo país de origen, lo que indica una estrecha relación evolutiva.

6.7. Búsqueda de nuevos factores de virulencia

Además de los factores de virulencia ya descritos en *A. paragallinarum*, se hizo una búsqueda de otros posibles factores de virulencia presentes en las cepas de *A. paragallinarum* de este proyecto; usando las plataformas RAST y VFDB (Tabla 9) se encontraron genes involucrados en Quorum sensing, adherencia, evasión inmune, toxinas, resistencia a antimicrobianos, resistencia a cobre y arsénico, metaloproteasas y metabolismo de carbohidratos entre otros.

Tabla 8. Genes asociados a virulencia presentes en 16 cepas de *A. paragallinarum*.

Factor de virulencia	Avpg1	Avpg2	Avpg3	Avpg4	Avpg5	Avpg6	Avpg7	Avpg8	Avpg9	Avpg10	2015	Esv135	CL	CT	5lab	6lab
Quorum sensing																
<i>lsrKRABCDGFS</i>																
Adherencia																
<i>pilN</i>																
<i>pilM</i>																
<i>pilT</i>																
<i>pilF</i>																
<i>pilA</i>																
<i>pilC</i>																
<i>pilP</i>																
<i>pilR</i>																
Bombas de eflujo																
<i>cmeC</i>																
<i>mdtK</i>																
<i>macA</i>																
<i>macB</i>																
<i>tetA</i>																
<i>tetR</i>																
Resistencia a fluoroquinolonas																
<i>parC</i>																
<i>parE</i>																
<i>gyrA</i>																
<i>gyrB</i>																
Resistencia a cloranfenicol																
<i>cat</i>																

Producción de bacteriocinas

<i>truA</i>																
<i>dedA</i>																
<i>accD</i>																
<i>folC</i>																
<i>dedE</i>																
<i>purF</i>																

Toxina antitoxina

<i>vapC</i>																
<i>higA</i>																
<i>higB</i>																
<i>relB</i>																
<i>relE</i>																
<i>yafQ</i>																
<i>dinJ</i>																

Tabla 9. Función de probables genes de *A. paragallinarum* relacionados con virulencia.

Gen	Función	Gen	Función	Gen	Función
<i>IsrKRAB</i> <i>CDFGS</i>	Autoinductores para QS y formación de biofilm	<i>macB</i>	Proteína de unión a ATP/permeasa	<i>purF</i>	Amidofosforribosiltransferasa
<i>pilN</i>	Pilus tipo IV PilN	<i>tetA</i>	Proteína de salida de tetraciclina	<i>vapC</i>	Proteína de la toxina VapC
<i>pilM</i>	Pilus tipo IV PilM	<i>tetR</i>	Regulador transcripcional	<i>higA</i>	Proteína HigA (antitoxina de HigB)
<i>pilT</i>	Proteína de ensamblaje de pilus PilT	<i>parC</i>	Subunidad A de la topoisomerasa IV	<i>higB</i>	Proteína de toxina HigB
<i>pilF</i>	Pilus tipo IV PilF	<i>parE</i>	Subunidad B de la topoisomerasa IV	<i>relB</i>	Proteína de estabilización del replicón RelB/StbD (antitoxina para RelE/StbE)
<i>pilA</i>	Prepilina	<i>gyrA</i>	Subunidad A de la ADN girasa	<i>relE</i>	Toxina de estabilización del replicón RelE/StbE
<i>pilC</i>	Proteínas del sistema de secreción tipo II	<i>gyrB</i>	Subunidad B de la ADN girasa	<i>yafQ</i>	Proteína de toxina YafQ
<i>pilP</i>	Pilus tipo IV PilP	<i>truA</i>	ARNt pseudouridina sintasa A	<i>dinJ</i>	Proteína J inducible por daños en el ADN
<i>pilR</i>	Proteína reguladora de expresión de fimbrias tipo IV PilR	<i>dedA</i>	Proteína DedA	<i>cat</i>	Cloranfenicol acetiltransferasa
<i>cmeC</i>	Lipoproteína de membrana externa CmeC	<i>accD</i>	Acetil-coenzima A carboxil transferasa cadena beta		
<i>mdtK</i>	Proteína de extrusión. Familia MATE	<i>folC</i>	Dihidrofolato sintasa		
<i>macA</i>	Proteína de salida específica de macrólidos MacA	<i>dedE</i>	Proteína productora de colicina V		

En el genoma de *A. paragallinarum* también se buscaron los genes que codifican para proteasas, se delimitaron las secuencias codificantes y se alinearon sus secuencias de aminoácidos. En la **Tabla 11** se el análisis comparativo de las proteasas identificadas en el genoma de las 16 cepas de este proyecto y de tres cepas de otros países. Se puede observar que el valor del score es igual en la mayoría de las cepas, a excepción de las cepas CL, 221 y JF4211, en azul se señalan los valores de score que son menores respecto a los scores de otras cepas. Este procedimiento muestra que es posible hacer distinciones en algunos grupos de bacterias.

Tabla 10. Proteasas encontradas en el genoma de las diferentes cepas.

CEPAS	PROTEASAS									
	Aminopeptidasa N	ClpB	Carboxipeptidasa terminal	SppA	U32	RseP (regulador de sigma E)	SohB	DegS	Serin peptidasa	Endopeptidasa
	Score									
Avpg1	1838	1747	1398	1208	963	919	683	674	610	184
Avpg2	1835	1747	1398	1208	963	918	683	674	610	184
Avpg3	1835	1747	1398	1208	963	918	683	674	610	184
Avpg4	1835	1747	1398	1208	963	918	683	674	610	184
Avpg5	1835	1747	1398	1208	963	918	683	674	610	184
Avpg6	1835	1747	1398	1208	963	918	683	674	610	184
Avpg7	1838	1747	1398	1208	963	918	683	674	610	184
Avpg8	1838	1747	1398	1208	963	918	683	674	610	184
Avpg9	1835	1747	1398	1208	963	918	683	674	610	184
Avpg10	1835	1747	1398	1208	963	918	683	674	610	184
2015	1838	1747	1398	1208	963	919	683	674	610	184
Esv135	1838	1747	1398	1208	963	919	683	674	610	184
CL	1828	1694	1391	1208	958	915	681	674	610	184
CT	1838	1747	1398	1208	963	918	683	674	610	184
5lab	1838	1747	1398		963	919	683	674	610	184
6lab	1838	1747	1398	1208	963	918	683	674	610	184
72	1836	1747	1398	1208	963	918	683	674	610	184
221	1833	1692	1394	1207	960	910	681	674	610	184
JF4211	1833	1692	1394	1207	960	910	679	674	610	184
E-value	0	0	0	0	0	0	0	0	0	1.00E-50

6.8. Otros elementos

Otros elementos que se encontraron en las cepas de *A. paragallinarum* fueron transposasas, fagos y proteínas de elementos móviles, el número de estos fue diferente en cada cepa (Tabla 12), pero las diferencias en ellos son notables, las proteínas que no fueron identificadas por la base de datos se representaron como proteínas hipotéticas y estas también varían de acuerdo a la cepa.

Tabla 11. Número de proteínas hipotéticas, fagos y proteínas de elementos móviles presentes en las 16 cepas de esta investigación y cepas de diferentes países.

Cepa	Proteínas hipotéticas	Fagos	Proteínas de elementos móviles (SI y transposasas)
Avgp1	643	112	36
Avgp2	554	73	24
Avgp3	599	101	21
Avgp4	547	70	23
Avgp5	537	73	26
Avgp6	490	64	15
Avgp7	508	68	16
Avgp8	497	64	16
Avgp9	495	72	16
Avgp10	591	64	15
Esv135	363	92	22
2015	341	84	22
CL	507	111	129
CT	522	66	15
5lab	480	64	16
6lab	607	71	19
72	514	80	15
221	701	159	133
JF4211	852	202	188
CCUG 12835	565	143	6
p4chr1	427	82	37
SCPM-O-B-8407	485	104	8
FARPER	318	50	21
ADLAP02	309	19	17
Modesto	596	0	10
NCTC 10926	809	125	130
<i>Avibacterium volantium</i> CVC1	414	49	34

Usando como referencia el genoma de la cepa 2015 de *A. paragallinarum*, localizamos y señalamos la ubicación de las secuencias que codifican para las integrasas, transposasas y proteínas que favorecen la resistencia arsenato, nitrato, cloranfenicol y tetraciclina (Figura 12). Identificamos que los genes que codifican para las resistencias antes mencionadas y la mayoría de las secuencias que corresponden a transposasas se localizan en la región final del genoma.

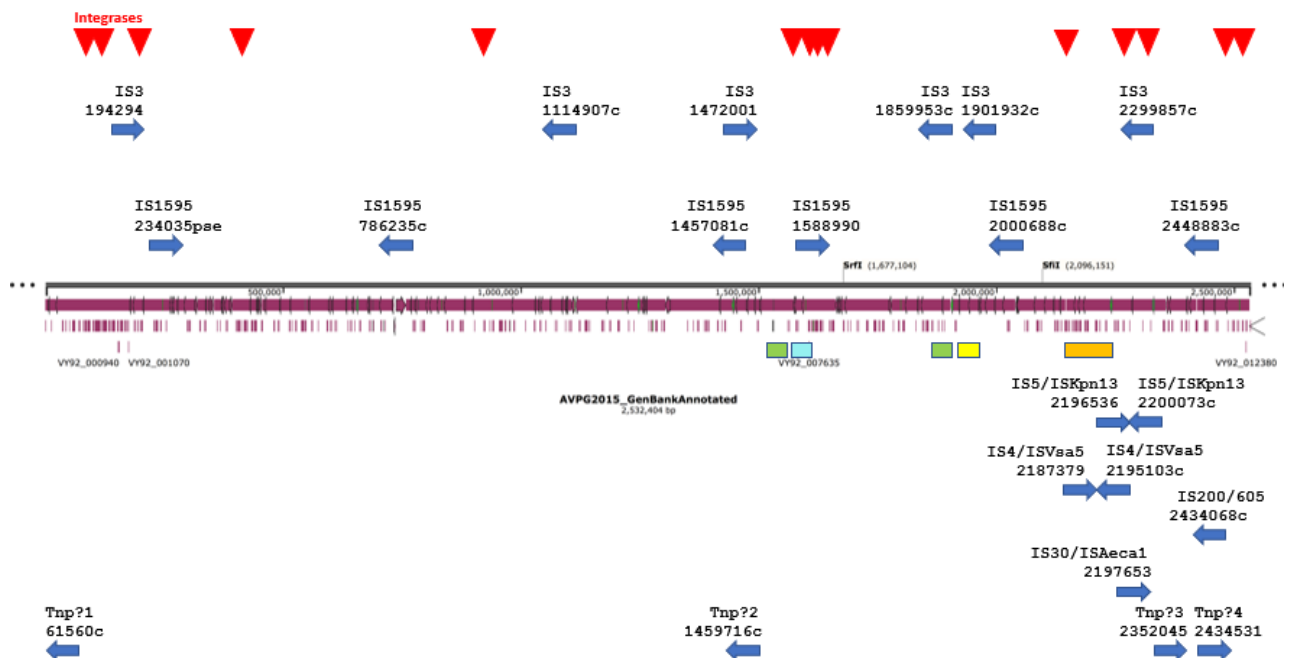


Figura 12. Organización del genoma de *A. paragallinarum* cepa 2015. En este esquema se indican las posiciones de los marcos de lectura abiertos que codifican para supuestas integrasas (rojo) y transposasas (azul). También se han identificado secuencias que codifican para proteínas relacionadas con resistencia a arsenato (verde), nitrato (amarillo), cloranfenicol y tetraciclina (naranja).

Se hizo un estudio de los reguladores transcripcionales que existen en *A. paragallinarum*, y se compararon con los reguladores presentes en otros patógenos como *Gallibacterium anatis*, *Escherichia coli* O157:H7 Sakai, *Pseudomonas aeruginosa* PAO, *Acinetobacter baumannii*, y el dizotrofo de vida libre *Azotobacter vinelandii* (Figura 13). Primero se quería conocer si la regulación de nuestro modelo de estudio era similar al de otras bacterias patógenas y segundo se quería identificar como *A. paragallinarum* podría regular la asimilación de nitrógeno, pues de acuerdo a Moreno Castillo 2023 *A. paragallinarum* carece del gen *rpoN* que codifica para el factor σ_{54} responsable de la asimilación de nitrógeno. Encontramos genes que codifican para proteínas tipo PII y GlnB, que son reguladores transcripcionales que también ayudan a la asimilación de nitrógeno. podemos observar cuales tienen mayor porcentaje de similitud y cuales son muy diferentes entre sí.

SizeScore	VarAus	VarGal	VarColi	VarPae	VarAct	VarVinc	GeneName	Query
426	0	90.516682	90.774681	90.514452	92.769998	92.769998	tsfR	QueryHit_NZ_CP050316.1_prot_WP_0005030805.1_2040 [genetansB] QueryHit_NZ_CP050316.1_prot_WP_0005030805.1_1341 [locus_tag=HL79_R006465]
13673636	61.201239	62.137462	78.024416	64.493204	64.493204	64.493204	lysR	QueryHit_NZ_CP050316.1_prot_WP_035687123.1_710 [locus_tag=HL79_R00810]
47169048	95.079996	95.151515	85.844156	85.844156	85.844156	85.844156	helicium-N	QueryHit_NZ_CP050316.1_prot_WP_035687123.1_710 [locus_tag=HL79_R00810]
64478131	9.964683	19.440727	24.502158	24.513691	24.513691	24.513691	rno	QueryHit_NZ_CP050316.1_prot_WP_035687466.1_309 [genetansB]
64571791	13.242009	88.781292	87.915624	87.915624	87.915624	87.915624	metJ	QueryHit_NZ_CP050316.1_prot_WP_035687521.1_2147 [genetansB]
8.75	22.8125	46.25	51.5625	46.19	46.19	46.19	grrA	QueryHit_NZ_CP050316.1_prot_WP_035688582.1_1010 [genetansB]
9.148265	16.73243	33.438486	38.801262	34.069421	34.069421	34.069421	irpA	QueryHit_NZ_CP050316.1_prot_WP_035688768.1_989 [genetansB]
9.375	16.071439	26.330366	26.330366	26.330366	26.330366	26.330366	tsfR	QueryHit_NZ_CP050316.1_prot_WP_035689239.1_231 [genetansB]
11.442786	16.40796	51.542239	53.23381	48.756219	48.756219	48.756219	nisA	QueryHit_NZ_CP050316.1_prot_WP_035689636.1_1954 [genetansB]
11.52315	21.883914	31.639389	38.873995	32.707776	32.707776	32.707776	nutB	QueryHit_NZ_CP050316.1_prot_WP_035689703.1_1915 [genetansB]
11.670337	11.975379	45.836165	41.66666	41.66666	41.66666	41.66666	grrB	QueryHit_NZ_CP050316.1_prot_WP_035689589.1_1491 [genetansB]
12.235294	13.276717	37.176741	40.367506	40.367506	40.367506	40.367506	grrA	QueryHit_NZ_CP050316.1_prot_WP_035690416.1_1381 [genetansB]
14.012739	35.031847	79.287516	80.732464	79.522293	79.522293	79.522293	asnC	QueryHit_NZ_CP050316.1_prot_WP_035686390.1_1943 [genetansC]
14.484127	21.829397	37.301587	43.650794	38.492063	38.492063	38.492063	YacD/PmpH	QueryHit_NZ_CP050316.1_prot_WP_035688786.1_578 [locus_tag=HL79_R002890]
14.731898	26.949667	40.207972	40.818182	40.818182	40.818182	40.818182	rpoH	QueryHit_NZ_CP050316.1_prot_WP_035690781.1_1789 [genetansB]
15.779468	22.43346	48.479087	89.60076	SO	FNHfamily	SO	FNHfamily	QueryHit_NZ_CP050316.1_prot_WP_046907381.1_856 [locus_tag=HL79_R004395]
17.424242	82.727273	94.318182	94.393939	94.793789	94.793789	94.793789	Rha family	QueryHit_NZ_CP050316.1_prot_WP_035689202.1_427 [locus_tag=HL79_R002140]
17.802119	32.620091	30.936667	29.23784	31.802262	31.802262	31.802262	P4 family	QueryHit_NZ_CP050316.1_prot_WP_035686239.1_1866 [locus_tag=HL79_R009585]
17.924258	62.735499	54.716981	63.36473	53.616352	53.616352	53.616352	Lys family	QueryHit_NZ_CP050316.1_prot_WP_035688558.1_361 [locus_tag=HL79_R001800]
18.096915	18.275156	65.974949	65.297741	70.215877	70.215877	70.215877	anxA	QueryHit_NZ_CP050316.1_prot_WP_037809993.1_321 [genetansC]
18.484048	32.42442	49.209029	59.209029	63.818181	63.818181	63.818181	rnaA	QueryHit_NZ_CP050316.1_prot_WP_035687721.1_1397 [genetansA]
18.522727	46.704545	62.693182	96.761384	95.977278	95.977278	95.977278	nadR	QueryHit_NZ_CP050316.1_prot_WP_035689069.1_2202 [genetansD]
18.687251	38.304094	68.567251	95.950292	71.671427	71.671427	71.671427	gntR	QueryHit_NZ_CP050316.1_prot_WP_035685118.1_1090 [genetansR]
18.781728	25.216034	30.071506	30.414142	52.020457	52.020457	52.020457	rpmI	QueryHit_NZ_CP050316.1_prot_WP_046907274.1_2254 [genetansE]
18.902439	21.963616	84.146841	87.112109	92.021218	92.021218	92.021218	rspD	QueryHit_NZ_CP050316.1_prot_WP_035687696.1_1407 [genetansD]
18.947368	24.912811	82.315789	87.286246	83.508772	83.508772	83.508772	aspA	QueryHit_NZ_CP050316.1_prot_WP_037805538.1_2171 [genetansA]
19.960661	20.156556	88.848401	93.722211	93.889696	93.889696	93.889696	pHA	QueryHit_NZ_CP050316.1_prot_WP_037805106.1_1440 [genetansA]
20.130213	39.77937	49.348613	47.664765	61.451551	61.451551	61.451551	grrA	QueryHit_NZ_CP050316.1_prot_WP_037805778.1_1407 [genetansA]
20.490506	30.85443	40.588443	44.541139	39.398734	39.398734	39.398734	rpoD	QueryHit_NZ_CP050316.1_prot_WP_035687518.1_2327 [genetansD]
20.97561	26.666667	68.292683	71.544715	68.292683	68.292683	68.292683	ovrR	QueryHit_NZ_CP050316.1_prot_WP_037806661.1_1929 [genetansD]
21.972322	44.01706	86.880278	80.94168	80.94168	80.94168	80.94168	grrA	QueryHit_NZ_CP050316.1_prot_WP_035689693.1_1925 [locus_tag=HL79_R002600]
22.535211	27.94272	42.957746	67.535211	42.235211	42.235211	42.235211	lexA	QueryHit_NZ_CP050316.1_prot_WP_035686020.1_291 [genetansA]
23.238831	53.532234	43.621866	59.62029	42.28636	42.28636	42.28636	evnB	QueryHit_NZ_CP050316.1_prot_WP_037806293.1_852 [genetansB]
23.46227	27.140214	39.224921	52.187029	39.215515	39.215515	39.215515	Rha-binding	QueryHit_NZ_CP050316.1_prot_WP_035689121.1_1694 [locus_tag=HL79_R008700]
23.613445	22.680676	23.613445	26.218487	22.352941	22.352941	22.352941	cvpA	QueryHit_NZ_CP050316.1_prot_WP_035686718.1_2173 [genetansA]
24.105012	34.128878	76.458711	66.587112	60.52206	60.52206	60.52206	fabR	QueryHit_NZ_CP050316.1_prot_WP_035686373.1_1930 [genetansA]
24.212126	29.19073	33.112683	43.046358	33.74384	33.74384	33.74384	vnrR	QueryHit_NZ_CP050316.1_prot_WP_037806091.1_1933 [genetansR]
24.242424	24.826927	41.464444	42.424242	42.424242	42.424242	42.424242	nraA	QueryHit_NZ_CP050316.1_prot_WP_035686684.1_633 [genetansA]
24.267534	92.410989	75.204479	92.30187	77.315068	77.315068	77.315068	pyrR	QueryHit_NZ_CP050316.1_prot_WP_035688882.1_246 [genetansC]
25	29.4038	90.379747	91.30779	90.50329	90.50329	90.50329	argB	QueryHit_NZ_CP050316.1_prot_WP_037806162.1_1105 [genetansR]
25.911819	35.90304	85.06679	81.762113	83.886828	83.886828	83.886828	transcription	QueryHit_NZ_CP050316.1_prot_WP_035686393.1_1925 [locus_tag=HL79_R009890]
26.680942	34.932548	44.625268	52.205567	45.695931	45.695931	45.695931	mtd	QueryHit_NZ_CP050316.1_prot_WP_046907219.1_1434 [genetansD]
27.210268	79.251701	77.891156	83.979592	78.061224	78.061224	78.061224	MarR/NprM	QueryHit_NZ_CP050316.1_prot_WP_046907299.1_1973 [locus_tag=HL79_R007100]
27.414573	81.445733	81.23201	77.95966	74.14466	74.14466	74.14466	response	QueryHit_NZ_CP050316.1_prot_WP_046907312.1_1892 [locus_tag=HL79_R001800]
29.023558	60.151803	69.658444	94.516129	37.001858	37.001858	37.001858	DeoR family	QueryHit_NZ_CP050316.1_prot_WP_035689241.1_2245 [locus_tag=HL79_R011500]
29.89418	85.31746	86.740232	84.17989	85.17346	85.17346	85.17346	Crp/OxyR family	QueryHit_NZ_CP050316.1_prot_WP_035686583.1_1550 [locus_tag=HL79_R002700]
30.564777	52.9125	65.974949	65.974949	65.974949	65.974949	65.974949	grrA	QueryHit_NZ_CP050316.1_prot_WP_035686020.1_291 [genetansA]
30.866167	39.228296	52.411576	91.832797	62.733119	62.733119	62.733119	mraZ	QueryHit_NZ_CP050316.1_prot_WP_037805964.1_1278 [genetansA]
31.034843	61.674877	87.20664	87.684229	88.029557	88.029557	88.029557	trpR	QueryHit_NZ_CP050316.1_prot_WP_035685175.1_765 [genetansR]
31.04	86.992	83.2	85.868	84.464	84.464	84.464	Lys family	QueryHit_NZ_CP050316.1_prot_WP_035687666.1_1262 [locus_tag=HL79_R006400]
32.92498	61.674877	87.20664	87.684229	88.029557	88.029557	88.029557	trpR	QueryHit_NZ_CP050316.1_prot_WP_035687712.1_1402 [genetansR]
33.527132	49.96124	79.651163	84.031008	78.875869	78.875869	78.875869	DeoR/OlpR	QueryHit_NZ_CP050316.1_prot_WP_035687474.1_2124 [locus_tag=HL79_R019125]
33.627671	90.228873	89.489592	93.292826	63.14901	63.14901	63.14901	MarR/NprM	QueryHit_NZ_CP050316.1_prot_WP_035686106.1_1811 [locus_tag=HL79_R009125]
33.808881	51.016948	65.974949	65.974949	65.974949	65.974949	65.974949	Rha family	QueryHit_NZ_CP050316.1_prot_WP_037806093.1_1174 [locus_tag=HL79_R002600]
34.362394	37.644788	49.227799	94.60251	49.26959	49.26959	49.26959	DeoR/OlpR	QueryHit_NZ_CP050316.1_prot_WP_035687993.1_50 [locus_tag=HL79_R000245]
35.176651	52.22743	55.760369	71.112352	61.20022	61.20022	61.20022	alpha 34-kDa	QueryHit_NZ_CP050316.1_prot_WP_037806974.1_338 [locus_tag=HL79_R001685]
35.249042	62.145948	89.89896	89.89896	89.89896	89.89896	89.89896	Rha family	QueryHit_NZ_CP050316.1_prot_WP_037807127.1_2158 [locus_tag=HL79_R011645]
35.327103	38.87805	82.074766	83.79493	80.560748	80.560748	80.560748	ldgR	QueryHit_NZ_CP050316.1_prot_WP_035686871.1_1239 [genetansR]
36.945813	49.9888	71.756979	74.05823	76.02723	76.02723	76.02723	transcription	QueryHit_NZ_CP050316.1_prot_WP_037806293.1_840 [locus_tag=HL79_R004275]
38.190955	49.58124	88.49951	84.187605	84.241808	84.241808	84.241808	invY	QueryHit_NZ_CP050316.1_prot_WP_035686708.1_1248 [genetansR]
38.936015	84.447478	50.44478	58.40708	62.876166	62.876166	62.876166	response	QueryHit_NZ_CP050316.1_prot_WP_035687902.1_1703 [locus_tag=HL79_R003600]
39.463602	72.58705	67.241379	69.880507	72.720307	72.720307	72.720307	MarR family	QueryHit_NZ_CP050316.1_prot_WP_035688916.1_2141 [locus_tag=HL79_R011000]
40.725066	68.145161	75.072043	85.897932	78.05445	78.05445	78.05445	helicium-N	QueryHit_NZ_CP050316.1_prot_WP_035686775.1_582 [locus_tag=HL79_R002200]
40.961338	55.384615	63.932828	62.07623	63.841615	63.841615	63.841615	GenR family	QueryHit_NZ_CP050316.1_prot_WP_035686746.1_1187 [locus_tag=HL79_R000700]
41.952804	80.470163	84.954792	87.366022	84.195280	84.195280	84.195280	helicium-N	QueryHit_NZ_CP050316.1_prot_WP_035686491.1_1886 [locus_tag=HL79_R012000]
42.317073	42.073717	94.310793	93.804878	94.384346	94.384346	94.384346	DNA-binding	QueryHit_NZ_CP050316.1_prot_WP_035686488.1_2103 [locus_tag=HL79_R012005]
42.500649	74.774194	77.74194	83.20797	78.782096	78.782096	78.782096	grrA	QueryHit_NZ_CP050316.1_prot_WP_035686702.1_2126 [genetansB]
43.143313	47.197191	50.501672	58.19390	51.505017	51.505017	51.505017	nutB	QueryHit_NZ_CP050316.1_prot_WP_035686404.1_943 [genetansB]
43.444444	79.431072	82.997812	87.789934	80.891756	80.891756	80.891756	XRE family	QueryHit_NZ_CP050316.1_prot_WP_046907331.1_424 [locus_tag=HL79_R002125]
43.623237	62.022222	63.97506	62.464747	62.464747	62.464747	62.464747	rspD	QueryHit_NZ_CP050316.1_prot_WP_046907318.1_2303 [genetansA]
43.886443	79.232344	82.947598	87.379913	91.004369	91.004369	91.004369	XRE family	QueryHit_NZ_CP050316.1_prot_WP_046907547.1_1500 [locus_tag=HL79_R007685]
44.433996	63.677932	89.44334	90.815109	89.284294	89.284294	89.284294	fadR	QueryHit_NZ_CP050316.1_prot_WP_167316917.1_1106 [genetansD]
46.577947	47.414449	64.28887	83.073136	81.11871	81.11871	81.11871	hda	QueryHit_NZ_CP050316.1_prot_WP_037805811.1_1439 [genetansD]
48.444016	44.27521	71.618262	69.02866	67.638521	67.638521	67.638521	grrA	QueryHit_NZ_CP050316.1_prot_WP_037805964.1_1278 [genetansA]
48.545451	82.2							

6.9. Determinantes genéticos relacionados con Resistencia antimicrobiana

Al hacer el ensamble de los genomas de *A. paragallinarum* no se encontraron plásmidos, esto se comprobó haciendo geles *in situ* con la técnica de Eckhardt (Figura 29). Sin embargo, sí se encontraron genes de resistencia a antimicrobianos en el genoma. Además, un alineamiento del gen 16S rRNA de *A. paragallinarum* y el gen 16S de *E. coli*, mostró que a lo largo de la secuencia de *A. paragallinarum* hay varios cambios puntuales y deleciones en algunas bases (Figura 30). Estos cambios y deleciones podrían estar relacionados directamente a resistencia a antimicrobianos, especialmente a estreptomicina.



Figura 14. Gel *in situ* de Eckhardt. Agarosa al 0.8% en que se muestra DNA genómico, pero ausencia de DNA plasmídico de diferentes cepas de *A. paragallinarum*. MW: 1kb Plus DNA Ladder

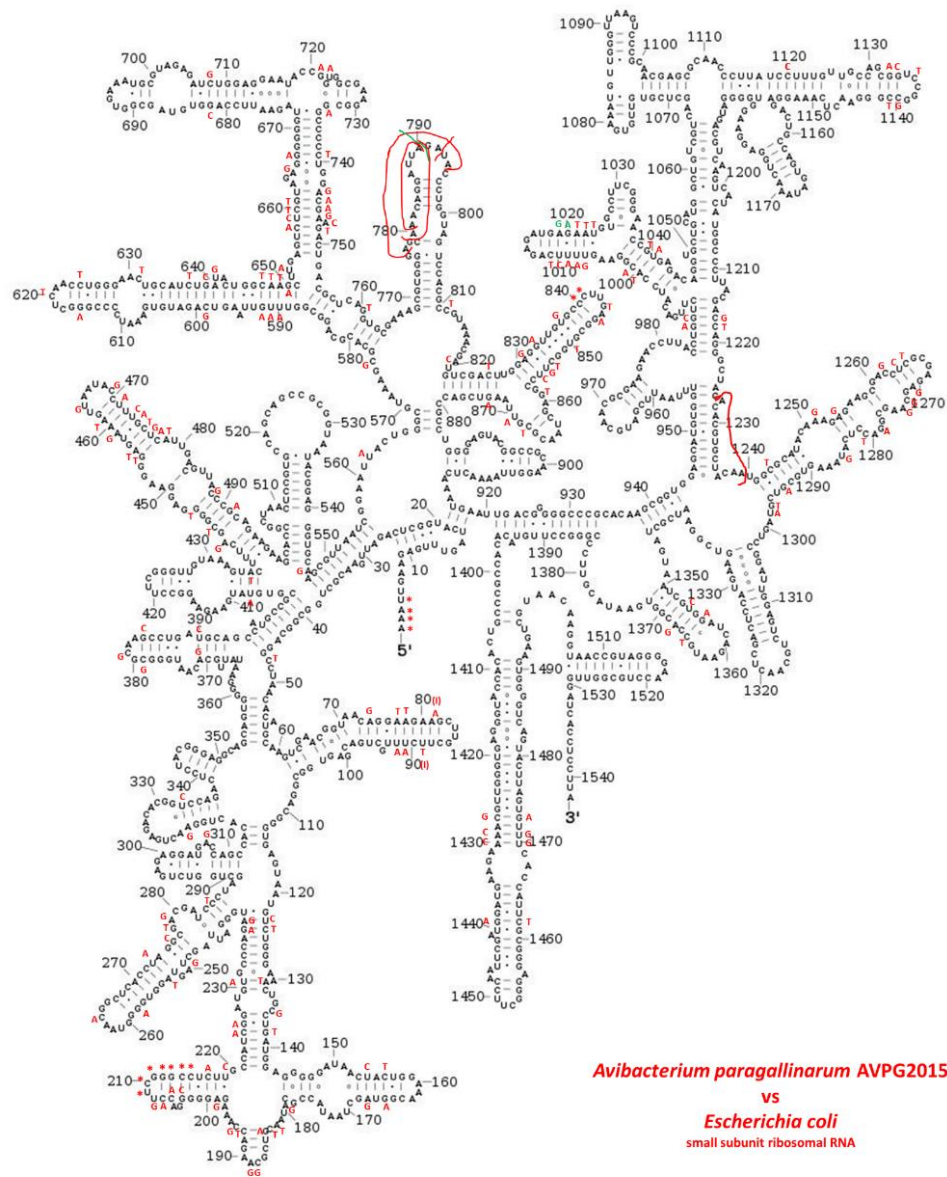


Figura 15. Estructura secundaria del 16S rRNA. En negro se muestra la secuencia del gen16S de *E. coli*, las letras en rojo representan las bases que están presente en la cepa de 2015 y que son diferentes en *E. coli*. Las líneas en rojo representan la región donde el gen 16S de *A. paragallinarum* sufre una deleción de bases.

6.10. Antibiogramas por difusión en disco

Por encontrar genes de resistencia a los antibióticos en el genoma de *A. paragallinarum* 2015, se hizo una prueba de susceptibilidad a antimicrobianos con discos impregnados con antibiótico (BioRad®). Estas sustancias se usan en el diagnóstico clínico, como suplementos alimenticios o promotores de crecimiento de aves y son comunes en el tratamiento de la coriza infecciosa. El análisis de la prueba de susceptibilidad a antimicrobianos de *A. paragallinarum* muestra que el 100% de las cepas evaluadas fueron resistentes a ampicilina, amikacina, estreptomycin y gentamicina, 88.8% fueron resistentes a tetraciclina y 38.8% fueron resistentes a cloranfenicol, de acuerdo a la interpretación sugerida por el fabricante de los sensibilizadores. Estos resultados muestran la funcionalidad de los genes identificados y son congruentes con lo observado en países como Ecuador, Holanda y Taiwán, donde se ha reportado un alto índice de resistencia a antimicrobianos en *A. paragallinarum*.

Tabla 12. Resistencia antimicrobiana de cepas de *A. paragallinarum*, R=Resistente, I=Intermedio, S=sensible

Cepa	Antibiótico											
	Ampicilina 10µg/ml		Tetraciclina 30µg/ml		Amikacina 30µg/ml		Cloranfenicol 30µg/ml		Estreptomycin 30µg/ml		Gentamicina 10µg/ml	
	mm		mm	mm		mm		mm		mm		mm
Avgp1	9	R	0.5	R	8	R	5	R	0	R	6	R
Avgp2	0	R	10	R	9	R	16	I	0	R	5	R
Avgp3	9	R	2	R	5	R	3	R	0	R	8	R
Avgp4	6	R	12	R	5	R	16	I	0	R	8	R
Avgp5	10	R	11	R	10	R	14	I	0	R	9	R
Avgp6	9	R	11	R	7	R	13	I	0	R	6	R
Avgp7	0	R	0	R	10	R	5	R	0	R	0	R
Avgp8	8	R	12	R	7	R	16	I	0	R	10	R
Avgp9	6	R	11	R	8	R	15	I	0	R	5	R
Avgp10	5	R	11	R	10	R	14	I	0	R	7	R
Esv135	5	R	1	R	8	R	4	R	0	R	8	R
2015	10	R	3	R	10	R	4	R	0	R	10	R
CL	10	R	18	I	12	R	16	I	0	R	10	R

CT	9	R	15	I	11	R	15	I	0	R	11	R
5lab	4	R	9	R	8	R	5	R	0	R	8	R
6lab	8	R	11	R	6	R	15	I	0	R	9	R
7.5	5	R	5	R	6	R	-	-	0	R	5	R
4.1	0	R	4	R	4	R	11	R	4	R	3	R

VII. DISCUSIÓN

Avibacterium paragallinarum es un microorganismo difícil de crecer por lo que manipularlo en el laboratorio puede llegar a ser complicado, a pesar de esto, existen varios trabajos que ponen en evidencia algunos de los factores de virulencia responsables de que la bacteria pueda causar enfermedad a su hospedero. La secuenciación del genoma de la bacteria ha sido de suma importancia, y países como Suiza, China, Perú y México entre otros, han depositado sus genomas en bases de datos. (Aguilar *et al.*, 2013; Xu *et al.*, 2013; Requena *et al.*, 2013; Horta *et al.*, 2017); es importante la disposición de genomas de *A. paragallinarum* de diferentes ubicaciones geográficas pues así podremos conocer similitudes y diferencias que permitirían el desarrollo de herramientas más eficientes para combatir a la bacteria sea cual sea su serotipo.

Secuenciación

Pese a que *A. paragallinarum* es de lento crecimiento, se optimizó el cultivo bacteriano y el procedimiento para extraer y purificar su material genético cromosomal en cantidad suficiente de las cepas investigadas, para proceder con la secuenciación del genoma y obtener secuencias de alta calidad (Nietsch *et al.*, 2016). El control de calidad de las lecturas sin procesar fue el primer paso para la reconstrucción del genoma y aunque hay disponibles varias herramientas para evaluar la calidad y mejorarla, en este proyecto se utilizó FastQC, pues tiene la ventaja mostrar varios parámetros de calidad en comparación con otras herramientas, por ejemplo, la calidad promedio de las lecturas, la valoración del contenido de GC e identificación de secuencias duplicadas (Andrews, 2010).

Analizada y mejorada la calidad del genoma hicimos el ensamble, para esto usamos tres herramientas, MIRA, SPAdes y ABySS sin embargo, con ninguna de las tres herramientas pudimos obtener genomas cerrados, esto probablemente se debe a que la secuenciación con Illumina nos generó lecturas cortas, de 20-301 pares de bases aproximadamente y no hubo solapamiento de algunas lecturas (Pollard *et al.*, 2018).

Algunas cepas representativas fueron seleccionadas y sus genomas se secuenciaron con la tecnología PacBio, las secuencias largas ayudaron a que algunos genomas pudieran cerrarse y el resto quedaran con un bajo número de contigs, esto nos permitió conocer el orden de todos los genes en el genoma y contar con cepas que fueran un referente mexicano. Los genomas de las cepas CL y CT no pudieron reconstruirse, probablemente por la presencia de secuencias repetidas en el genoma que confunde a los algoritmos y provoca errores, impidiendo un ensamblaje correcto (Schatz *et al.*, 2010).

La anotación de los genomas reconstruidos se llevó a cabo inicialmente en el servidor RAST, que organiza la información por subsistemas metabólicos, proporciona nombre y función de genes. Además, es un servicio gratuito, muy flexible, proporciona anotaciones bastante completas, consistentes, precisas que permiten revalorar el tipo de secuencias identificadas a diferencia de los registros finales de GenBank (Aziz *et al.*, 2008). El resultado de RAST acerca de los genes relacionados con virulencia fue adicionalmente comparado y valorado en VFDB (Virulence Factor Database). Este análisis comparativo mostró genes de virulencia conservados y comprobados funcionalmente en *A. paragallinarum*, como *avxA*, *cdtABC*, genes que codifican fimbrias, hemaglutininas etc. Sin embargo, se amplió el enfoque a aquellos que todavía no han sido descritos o anotados.

Adherencia

Un grupo de genes que identificamos en todos los genomas son los que codifican para proteínas Pil. Los pili de tipo IV son fibras flexibles que se encuentran en la superficie de las bacterias, incluidos miembros de la familia *Pasteurellaceae* como *Haemophilus influenzae*, *Pasteurella multocida* y *Actinobacillus pleuropneumoniae*; la identificación de estos genes de virulencia en *A. paragallinarum* nos da una pauta para poder comprender cómo la bacteria forma biopelícula, se adhiere a células epiteliales del huésped, interactúa con fagos, etc. Además, debido a que los pilis median la adherencia de las bacterias con las células del huésped, se pueden proponer y desarrollar estrategias que interrumpan esta interacción y que

sean una opción de tratamiento o vacuna (Bujold & MacInnes 2015; Jacobsen *et al.*, 2020).

Resistencia a antimicrobianos

Por otra parte, la presencia de genes que codifican para sistemas de eflujo, podría explicar por qué la bacteria tiene resistencia a una amplia gama de fármacos y otros compuestos nocivos para ella (Nishino *et al.*, 2006; Du *et al.*, 2015), MacAB es un transportador de eflujo de macrólidos (Nishino *et al.*, 2006), que ayuda a la expulsión de antibióticos fuera de la bacteria, así que la delección de los genes *macAB* podrían atenuar la virulencia pues induce la susceptibilidad a muchos antimicrobianos (Golparian *et al.*, 2014).

Algunos genes de resistencia a fluoroquinolonas también fueron identificados, una de las posibles causas de su presencia es que estos y otros antimicrobianos son utilizados en aves de corral para promoción del crecimiento, prevención y tratamiento de enfermedades (Gyles, 2008). Ensayos de multiresistencia antimicrobiana en *A. paragallinarum* hechos en países de Latinoamérica, mostraron alto nivel de resistencia a eritromicina, estreptomicina, gentamicina, kanamicina y tetraciclina entre otros (Luna *et al.*, 2016), incluso en Holanda la resistencia a tetraciclina es alta, y en algunos países las cepas son resistentes a al menos tres antibióticos diferentes (Hsu *et al.*, 2007). La resistencia a estreptomicina observada por parte de *A. paragallinarum*, podría deberse a la serie de cambios en la secuencia de bases del gen 16S rRNA, cuando se compara con la secuencia del gen 16S rRNA de *E. coli*, se ha demostrado que mutaciones en algunas copias del 16S rRNA son responsables de que la estreptomicina no pueda reconocer su sitio blanco y por lo tanto el antimicrobiano no tenga efecto sobre la bacteria (Burkhard *et al.*, 2021).

También, la presencia de genes que codifican para colicina V sugiere que la bacteria podría competir con otras bacterias estrechamente relacionadas que no son resistentes a la bacteriocina, y de esta forma asegurando el éxito en la

colonización (Gilson *et al.*, 1990). El nivel de bacteriocina producida por *A. paragallinarum* dependerá de la escala de competencia (Gardner *et al.*, 2004).

Proteasas

Respecto a los genes que codifican para proteasas, no es de extrañar se hayan encontrado presentes en el genoma de *A. paragallinarum*, pues tienen un papel en el desarrollo de la CI. Anteriormente se reportó que la bacteria secretaba metaloproteasas ante situaciones de estrés nutricional, y que tenían la capacidad de degradar inmunoglobulina G (IgG) de pollo (Ramón *et al.*, 2006), y ser estables hasta los 50°C (García *et al.*, 2005), suficiente como para tolerar la temperatura corporal del pollo. También se ha reportado que algunas proteasas ayudan a la formación de biopelículas (Alighami *et al.*, 2022), además son responsables del control de calidad de las proteínas y llevan a cabo la proteólisis ante estímulos ambientales, alterando así procesos biológicos (Ingmer *et al.*, 2009).

Sistema Toxina-Antitoxina

En la mayoría de los genomas de procariontes se ha reportado el hallazgo de un sistema Toxina-Antitoxina (TA) y, esto incluye el genoma de *A. paragallinarum*. En estos sistemas, las toxinas generalmente son proteínas que pueden reducir el metabolismo de la bacteria en situaciones de estrés como déficit de nutrientes, altas temperaturas, infección por bacteriófagos y exposición a antibióticos, mientras que las antitoxinas son proteínas o RNA que contrarrestan a la toxina (Singh *et al.*, 2021). Otras funciones de los sistemas TA son regular la patogenicidad y proteger contra elementos genéticos móviles. Los sistemas TA encontrados en *A. paragallinarum* son YafQ/DinJ, RelB/RelE, HigB/HigA y VapB/VapC, aunque de este último sólo identificamos a la toxina (VapC), todos los sistemas TA son de tipo II, es decir, la antitoxina inhibe la actividad de la toxina formando un complejo proteína-proteína (Goeders & Van, 2013).

En el sistema TA RelB/RelE, la toxina RelE responde ante situaciones de estrés nutricional, reduciendo el nivel de traducción que a su vez provoca inhibición de crecimiento y formación de colonias sin llegar a la muerte celular (Christensen *et*

al., 2001). Algo similar sucede con el sistema TA YafQ/DinJ que reduce el crecimiento de la bacteria y forma células persistentes ante la exposición a antibióticos (Hu *et al.*, 2015). En el sistema TA HigB/HigA, HigA es la antitoxina que inhibe la actividad RNasa de la toxina HigB. Este sistema está involucrado en la virulencia de la bacteria, pues reduce la formación de biopelícula y afectan la patogenicidad de la bacteria (Wood & Wood, 2016). Por último, en el sistema TA VapB/VapC asociado a la virulencia, la toxina VapC inhibe la traducción mediante la escisión del sitio tRNA iniciador o la escisión del sitio del bucle Sarcin-Ricin del 23S rRNA, permitiendo la reducción del nivel de traducción, sin embargo, en *A. paragallinarum* solo se encontró la toxina VapC (Winther *et al.*, 2016).

Elementos genéticos móviles

Por otra parte, en este trabajo se identificaron secuencias con homología a transposasas, integrasas y fagos. Esta información es importante, pues nos dan un panorama de las herramientas que tiene la bacteria para resistir el estrés ambiental, información relacionada con la capacidad de transferencia lateral de genes en una bacteria recalcitrante a la transformación (Wagner & Waldor, 2002), además, la presencia de transposasas podría indicar plasticidad y adaptación de la bacteria a situaciones de estrés, como baja disponibilidad de nutrientes y altas temperaturas (Vigil *et al.*, 2017); estos elementos permiten a la bacteria adaptarse al medio ambiente para poder colonizar (Vandecraen *et al.*, 2017).

Reguladores transcripcionales

Las bacterias requieren una respuesta rápida ante diferentes estímulos ambientales, por ello, los reguladores transcripcionales regulan la transcripción de genes en beneficio de la bacteria (Pis *et al.*, 2022). En el genoma de *A. paragallinarum* identificamos los genes que codifican para reguladores transcripcionales, los factores sigma son algunos de los más importantes, entre ellos están σ_{70} , σ_{32} y σ_{24} , sin embargo, σ_{54} no se encontró (Moreno Castillo 2023, datos no publicados), esto mismo ocurre con otros Pasteurellaceos como *A. actinomycetemcomitans*, *P. multocida* y *H. influenzae* (Pis *et al.*, 2022;

Studholme & Buck, 2000). σ_{54} o sigma N es codificado por *rpoN*, este regulador está relacionado con la asimilación de nitrógeno, pero no es indispensable para el crecimiento, la ausencia de este regulador podría indicar que su función no es esencial para la bacteria (Studholme & Buck, 2000). En el caso de *A. paragallinarum* se identificaron los reguladores transcripcionales GlnB y proteínas PII, que podrían cumplir la función de σ_{54} , pues también responden ante estrés por nitrógeno e interactúan con proteínas que controlan otros reguladores de la transcripción (Yurgel *et al.*, 2010; Huergo *et al.*, 2013).

VIII. CONCLUSIONES

- Se secuenciaron, ensamblaron y anotaron genomas de 18 cepas de *A. paragallinarum* con la tecnología de Illumina y PacBio, de estos genomas 5 fueron circularizados y dos ya se encuentran registrados en GenBank con los números de acceso CP058307.1 (AVPG2015) y CP050316.1(Esv135)
- Se identificaron genes que codifican para factores de virulencia y que tienen un papel importante en la adherencia (*pilACFNMPRT*), además de los ya reportados previamente en otros trabajos de investigación.
- Se determinó el perfil de resistencia en 16 cepas de *A. paragallinarum*, y se identificaron los genes que podrían estar involucrados (*macAB*, *tetA/tetR*, *parC/parE*, *cat* y *gyrA/gyrB*), los resultados sugieren que puede haber uso indiscriminado de antimicrobianos usados en la industria avícola con fines de tratamiento y promoción del crecimiento de las aves.
- Se identificaron sistemas toxina antitoxina (*higAB*, *relBE*, *yafQ/dinJ* y *vapC*), estos sistemas podrían ayudar a entender como *A. paragallinarum* regula su metabolismo en situaciones de estrés para evitar la muerte de la bacteria y que pueda persistir en el ambiente donde se encuentra.

IX. PERSPECTIVAS

- Circularizar los 11 genomas secuenciados con PacBio que quedaron con bajo número de contigs
- Elaborar proteínas recombinantes en un modelo heterólogo para evaluar su potencial inmunogénico y su rol preciso en la virulencia bacteriana.
- Llevar a cabo análisis transcriptómico de cultivos de *A. paragallinarum* para conocer el comportamiento de expresión temporal de los genes de virulencia en diferentes condiciones de cultivo.

X. REFERENCIAS

Abdel Naime WA, Fahim JR, Fouad MA, Kamel MS. (2019). Antibacterial, antifungal, and GC–MS studies of *Melissa officinalis*. *South African J Bot.* 124:228–34. doi: 10.1016/j.sajb.2019.05.011

Aguilar BL, Calderón Copete SP, Frey J, Falquet L. (2013). Draft Genome Sequence of the Virulent *Avibacterium paragallinarum* Serotype A Strain JF4211 and Identification of Two Toxins. *Genome Announc.* 15:e00592-13. doi: 10.1128/genomeA.00592-13.

Andrews S. FastQC: A quality control tool for high throughput sequence data. (2010). Available: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

Antenucci F, Arak H, Gao J, Allahgadry T, Thøfner I, Bojesen AM. (2020). Hydrostatic filtration enables large-scale production of outer membrane vesicles that effectively protect chickens against *Gallibacterium anatis*. *Vaccines.* 8(1). doi:10.3390/vaccines8010040

Antipov D, Korobeynikov A, McLean JS, Pevzner PA. (2016). HybridSPAdes: An algorithm for hybrid assembly of short and long reads. *Bioinformatics.* 32(7):1009–15. doi: 10.1093/bioinformatics/btv688

Alföldi J, Lindblad-Toh K. (2013). Comparative genomics as a tool to understand evolution and disease. *Genome Res.* 23(7):1063–8. doi: 10.1101/gr.157503.113

Aljghami ME, Barghash MM, Majaesic E, Bhandari V, Houry WA. (2022). Cellular functions of the ClpP protease impacting bacterial virulence. *Front Mol Biosci.* 9:1–19. doi: 10.3389/fmolb.2022.1054408

Atapattu DN, and Czuprynski CJ. (2005) *Mannheimia haemolytica* leukotoxin induces apoptosis of bovine lym-phoblastoid cells (BL-3) via a caspase-9-dependent mito-chondrial pathway. *Infection and immunity,* 73(9), 5504–5513. doi: 10.1128/IAI.73.9.5504-5513.2005

Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, et al. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics.* 9:1–15. doi: 10.1186/1471-2164-9-75

Bager RJ, Persson G, Nesta B, Soriani M, Serino L, Jeppsson M, et al. (2013). Outer membrane vesicles reflect environmental cues in *Gallibacterium anatis*. *Vet Microbiol.* 167(3–4):565–72. doi: 10.1016/j.vetmic.2013.09.005

Barnett DW, Garrison EK, Quinlan AR, Stürmberg MP, Marth GT. (2011). Bamtools: A C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics.* 27(12):1691–2. doi:10.1093/bioinformatics/btr174

Benz R. (2016). Channel formation by RTX-toxins of pathogenic bacteria: Basis of their biological activity. *Biochim Biophys Acta – Biomembr.* 1858(3):526–37. doi: 10.1016/j.bbamem.2015.10.025

Blackall PJ, Eaves LE, Rogers DG. (1990). Proposal of a new serovar and altered nomenclature for *Haemophilus paragallinarum* in the Kume hemagglutinin scheme. *J Clin Microbiol.* 28(6):1185–7. doi: 10.1128/jcm.28.6.1185-1187.1990

- Bolger AM, Lohse M, Usadel B.** (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–20. doi:10.1093/bioinformatics/btu170
- Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al.** (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep*. 5. doi: 10.1038/srep08365
- Bujold AR, MacInnes JI.** (2015). Identification of putative adhesins of *Actinobacillus suis* and their homologues in other members of the family *Pasteurellaceae*. *BMC Res Notes*. 8(1):1–13. doi: 10.1186/s13104-015-1659-x
- Burkhard S, Yishak G, Kidan, Therdsak P, Kerstin E, Erik C.B, and Peter S.** (2001). Mechanisms of Streptomycin Resistance: Selection of Mutations in the 16S rRNA Gene Conferring Resistance. *Antimicrob Agents Chemother*. 45(10):2877–2884. doi: 10.1128/AAC.45.10.2877-2884.2001
- Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al.** (2005). VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res*. 33:325–8. doi: 10.1093/nar/gki008
- Chen YC, Tan DH, Shien JH, Hsieh MK, Yen TY, Chang PC.** (2014). Identification and functional analysis of the cytolethal distending toxin gene from *Avibacterium paragallinarum*. *Avian Pathol*. 43(1):43–50. doi: 10.1080/03079457.2013.861895
- Chevreux B, Wetter T, Suhai S.** (1999). Genome sequence assembly using trace signals and additional sequence information. *Comput Sci Biol: Proceedings of the German Conference on Bioinformatics (GCB)*. 99: 45–56
- Christensen SK, Mikkelsen M, Pedersen K, Gerdes K.** (2001). RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci U S A*. 98(25):14328–33. doi: 10.1073/pnas.251327898
- Christensen H, Kuhnert P, Norskov-Lauritsen N, Planet P.J, Bisgaard M.** (2014). The Family *Pasteurellaceae*. *The Prokaryotes: Gammaproteobacteria*. 535-564. doi: 10.1007/978-3-642-38922-1_224
- Comas I, Cancino Muñoz I, Mariner Llicer C, Goig G, Ruiz Hueso P, Francés Cuesta C, et al.** (2020). Uso de las tecnologías de secuenciación masiva para el diagnóstico y epidemiología de enfermedades infecciosas. *Enfermedades Infecciosas y Microbiología Clínica*. 38:32-38. doi: 10.1016/j.eimc.2020.02.006
- Comstock LE, Kasper DL.** (2006). Bacterial Glycans: Key Mediators of Diverse Host Immune Responses. *Cell*. 126(5):847–50. doi: 10.1016/j.cell.2006.08.021
- Cress BF, Englaender JA, He W, Kasper D, Linhardt RJ, Koffas MAG.** (2014). Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiol Rev*. 38(4):660–97. doi: 10.1111/1574-6976.12056
- Crispo M, Blackall P, Khan A, Shivaprasad HL, Clothier K, Senties-Cué CG, et al.** (2019). Characterization of an Outbreak of Infectious Coryza (*Avibacterium paragallinarum*) in Commercial Chickens in Central California. *Avian Dis*. 63(3):486–94. doi: 10.1637/19-00081.1

Dereja IA, Hailemichael D. (2017). Infectious Coryza in Jimma Backyard Chicken Farms: Clinical and Bacteriological Investigation. *J Vet Sci Technol.* 08(01):1–6. doi: 10.4172/2157-7579.1000412

DiRienzo JM. (2014). Breaking the gingival epithelial barrier: Role of the *Aggregatibacter actinomycetemcomitans* cytolethal distending toxin in oral infectious disease. *Cells.* 3(2):476–99. doi:10.3390/cells3020476

Eshraghi A, Maldonado Arocho FJ, Gargi A, Cardwell MM, Prouty MG, Blanke SR, et al. (2010). Cytolethal distending toxin family members are differentially affected by alterations in host glycans and membrane cholesterol. *J Biol Chem.* 285(24):18199–207. doi: 10.1074/jbc.M110.112912

Eckhardt, T. A (1978). Rapid method for the identification of plasmid desoxyribonucleic acid in bacteria. *Plasmid.* 1(4):584–588. doi:10.1016/0147-619x(78)90016-1

Falconi Agapito AF, Saravia LE, Flores Pérez A, Fernández Díaz M. (2015). Naturally Occurring β -Nicotinamide Adenine Dinucleotide Independent *Avibacterium paragallinarum* Isolate in Perú. 341–3. doi: 10.1637/10969-110314-CaseR

Figaj D, Ambroziak P, Przepiora T, Skorko-Glonek J. (2019). The role of proteases in the virulence of plant pathogenic bacteria. *Int J Mol Sci.* 20(3). doi:10.3390/ijms20030672

Frey J, Kuhnert P. (2002). Mini-Review RTX toxins in Pasteurellaceae. *Int J Med Microbiol.* 292:149–58. doi: 10.1078/1438-4221-00200

Frey J. (2011). The role of RTX toxins in host specificity of animal pathogenic *Pasteurellaceae*. *Vet Microbiol.* 153(1–2):51–8. doi: 10.1016/j.vetmic.2011.05.018

Frey J. (2019). RTX toxins of animal pathogens and their role as antigens in vaccines and diagnostics. *Toxins (Basel).* 11(12). doi: 10.3390/toxins11120719

Fulsundar S, Harms K, Flaten GE, Johnsen PJ, Chopade BA, Nielsen KM. (2014). Gene transfer potential of outer membrane vesicles of *Acinetobacter baylyi* and effects of stress on vesiculation. *Appl Environ Microbiol.* 80(11):3469–83. doi: 10.1128/AEM.04248-13

García Gómez E, Vaca S, Pérez Méndez A, Ibarra Caballero J, Pérez Márquez V, Tenorio VR, et al. (2005). *Gallibacterium anatis*-secreted metalloproteases degrade chicken IgG. *Avian Pathol.* 34(5):426–9. doi: 10.1080/03079450500267866

Gargi A, Tamilselvam B, Powers B, Prouty MG, Lincecum T, Eshraghi A, et al. (2013). Cellular interactions of the cytolethal distending toxins from *Escherichia coli* and *Haemophilus ducreyi*. *J Biol Chem.* 288(11):7492–505. doi: 10.1074/jbc.M112.448118

Gardner A, West SA, Buckling A. (2004). Bacteriocins, spite and virulence. *Proc Biol Sci.* 271(1547):1529–35. doi: 10.1098/rspb.2004.2756.

Giani AM, Gallo GR, Gianfranceschi L, Formenti G. (2020). Long walk to genomics: History and current approaches to genome sequencing and assembly. *Comput Struct Biotechnol J.* 18:9–19. doi: 10.1016/j.csbj.2019.11.002

Gilson L, Mahanty HK. (1990). Genetic analysis of an MDR-like export system: the secretion of colicin V. *J Biol Chem.* 265(12):3875–84. doi: 10.1002/j.1460-2075.1990.tb07606.x

- Goeders N, Van Melderen L.** (2013). Toxin-antitoxin systems as multilevel interaction systems. *Toxins (Basel)*. 6(1):304–24. doi:10.3390/toxins6010304
- Goldman AD, Landweber LF.** (2016). What Is a Genome? *PLoS Genet*. 12(7):1–7. doi:10.1371/journal.pgen.1006181
- Golparian D, Shafer WM, Ohnishi M, Unemo M.** (2014). Importance of multidrug efflux pumps in the antimicrobial resistance property of clinical multidrug-resistant isolates of *Neisseria gonorrhoeae*. *Antimicrob Agents Chemother*. 58(6):3556–9. doi: 10.1128/AAC.00038-14
- Goodwin S, McPherson JD, McCombie WR.** (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nat Rev Genet*. 17(6):333–51. doi: 10.1038/nrg.2016.49
- Gordon A, Hannon GJ.** FastX Toolkit. Available: http://hannonlab.cshl.edu/fastx_toolkit/
- Gurevich A, Saveliev V, Vyahhi N, Tesler G.** QUAST: (2013). Quality assessment tool for genome assemblies. *Bioinformatics*. 29(8):1072–5. doi: 10.1093/bioinformatics/btt086
- Gyles CL.** (2008). Antimicrobial resistance in selected bacteria from poultry. *Anim Health Res Rev*. 9(2):149–58. doi: 10.1017/S1466252308001552
- Harper M, Boyce J.D. and Adler B.** (2012). The Key Surface Components of *Pasteurella multocida*: Capsule and Lipopolysaccharide. *Current Topics in Microbiology and Immunology*. 361: 39–51. doi: 10.1007/82_2012_202
- Heather JM, Chain B.** (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*. 107(1):1–8. doi: 10.1016/j.ygeno.2015.11.003; 10.1016/j.ygeno.2015.11.003
- Hernández MJ, Linhardt RJ.** (2001). Degradation of chondroitin sulfate and dermatan sulfate with chondroitin lyases. *Methods Mol Biol*. 171:363–71. doi: 10.1385/1-59259-209-0:363
- Hobb RI, Tseng HJ, Downes JE, Terry TD, Blackall PJ, Takagi M, et al.** (2002). Molecular analysis of a haemagglutinin of *Haemophilus paragallinarum*. *Microbiology*. 148(7):2171–9. doi: 10.1099/00221287-148-7-2171
- Hong SW, Kim BT, Shin HY, Kim WS, Lee KS, Kim YS, et al.** (2002). Purification and characterization of novel chondroitin ABC and AC lyases from *Bacteroides stercoris* HJ-15, a human intestinal anaerobic bacterium. *Eur J Biochem*. 269(12):2934–40. doi: 10.1046/j.1432-1033.2002.02967.x
- Horta Valerdi G, Sanchez Alonso MP, Perez Marquez VM, Negrete Abascal E, Vaca Pacheco S, Hernandez Gonzalez I, et al.** (2017). The Genome Sequence of *Avibacterium paragallinarum* Strain CL Has a Large Repertoire of Insertion Sequence Elements. *Genome Announc*. 5(15):1–2. doi: 10.1128/genomea.00152-17
- Hsu YM, Shieh HK, Chen WH, Sun TY, Shiang JH.** (2007). Antimicrobial susceptibility, plasmid profiles and haemocin activities of *Avibacterium paragallinarum* strains. *Vet Microbiol*. 124(3–4):209–18. doi: 10.1016/j.vetmic.2007.04.024
- Hu T, Chitnis N, Monos D, Dinh A.** (2021). Next-generation sequencing technologies: An overview. *Hum Immunol*. 82(11):801–11. doi: 10.1016/j.humimm.2021.02.012

Hu Y, Kwan BW, Osbourne DO, Benedik MJ, Wood TK. (2015). Toxin YafQ increases persister cell formation by reducing indole signalling. *Environ Microbiol.* 17(4):1275–85. doi: 10.1111/1462-2920.12567

Huergo LF, Chandra G, Merrick M. (2013). PII signal transduction proteins: Nitrogen regulation and beyond. *FEMS Microbiol Rev.* 37(2):251–83. doi: 10.1111/j.1574-6976.2012.00351.x

Huston WM. (2010). Bacterial proteases from the intracellular vacuole niche; Protease conservation and adaptation for pathogenic advantage. *FEMS Immunol Med Microbiol.* 59(1):1–10. doi:10.1111/j.1574-695X.2010.00672.x

Ingmer H, Brøndsted L. (2009). Proteases in bacterial pathogenesis. *Res Microbiol.* 160(9):704–10. doi:10.1016/j.resmic.2009.08.017

Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, et al. (2017). ABySS 2.0: Resource-efficient assembly of large genomes using a Bloom filter. *Genome Res.* 27(5):768–77. doi: 10.1101/gr.214346.116

Jacobsen T, Bardiaux B, Francetic O, Izadi-Pruneyre N, Nilges M. (2020). Structure and function of minor pilins of type IV pili. *Med Microbiol Immunol.* 209(3):301–8. doi: 10.1007/s00430-019-00642-5

Jalal Z, El Atki Y, Lyoussi B, Abdellaoui A. (2015). Phytochemistry of the essential oil of *Melissa officinalis* L. growing wild in Morocco: Preventive approach against nosocomial infections. *Asian Pac J Trop Biomed.* 5(6):458–61. doi: 10.1016/j.apjtb.2015.03.003

Jinadasa RN, Bloom SE, Weiss RS, Duhamel GE. (2011). Cytolethal distending toxin: A conserved bacterial genotoxin that blocks cell cycle progression, leading to apoptosis of a broad range of mammalian cell lineages. *Microbiology.* 157(7):1521–875. doi: 10.1099/mic.0.049536-0

Kisiela DI, Aulik NA, Atapattu DN and Czuprynski CJ. (2010), N-terminal region of *Mannheimia haemolytica* leukotoxin serves as a mitochondrial targeting signal in mammalian cells. *Cellular Microbiology,* 12: 976-987. doi: 10.1111/j.1462-5822.2010.01445.x

Koeppen K, Hampton TH, Jarek M, Scharfe M, Gerber SA, Mielcarz DW, et al. (2016). A Novel Mechanism of Host-Pathogen Interaction through sRNA in Bacterial Outer Membrane Vesicles. *PLoS Pathog.* 12(6):1–22. doi: 10.1371/journal.ppat.1005672

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. (2017). Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27(5):722–36. doi: 10.1101/gr.215087.116

Koronakis V, Koronakis E, & Hughes C. (1989). Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes. *The EMBO journal,* 8(2), 595–605. Doi: doi.org/10.1002/j.1460-2075.1989.tb03414.x

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 35(6):1547–9. doi: 10.1093/molbev/msy096

Kume K, Sawata A, Nakai T, Matsumoto M. (1983). Serological classification of *Haemophilus paragallinarum* with a hemagglutinin system. *J Clin Microbiol.* 17(6):958–64. doi: 10.1128/jcm.17.6.958-964.1983

- Küng E, Frey J.** (2013). AvxA, a composite serine-protease-RTX toxin of *Avibacterium paragallinarum*. *Vet Microbiol.* 163(3–4):290–8. doi: 10.1016/j.vetmic.2012.12.029
- Kuntu HMT, Jokinen EI, Valtonen ET, Sundberg LR.** (2011). Virulent and nonvirulent *Flavobacterium columnare* colony morphologies: Characterization of chondroitin AC lyase activity and adhesion to polystyrene. *J Appl Microbiol.* 111(6):1319–26. doi: 10.1111/j.1365-2672.2011.05149.x
- Leitão JH.** (2020). Microbial virulence factors. *Int J Mol Sci.* 21(15):1–6. doi: 10.3390/ijms21155320
- Leung KY, Wang Q, Yang Z, Siame BA.** (2019). *Edwardsiella piscicida*: A versatile emerging pathogen of fish. *Virulence.* 10(1):555–67. doi: 10.1080/21505594.2019.1621648
- Li N, Qin T, Zhang XL, Huang B, Liu ZX, Xie HX, et al.** (2015). Gene deletion strategy to examine the involvement of the two chondroitin lyases in *Flavobacterium columnare* virulence. *Appl Environ Microbiol.* 81(21):7394–402. doi:10.1128/AEM.01586-15
- Linhartová I, Bumba L, Mašn J, Basler M, Osička R, Kamanová J, et al.** (2010). RTX proteins: A highly diverse family secreted by a common mechanism. *FEMS Microbiol Rev.* 34(6):1076–112. doi: 10.1111/j.1574-6976.2010.00231.x
- Loman NJ, Pallen MJ.** (2015). Twenty years of bacterial genome sequencing. *Nat Rev Microbiol.* 13(12):787–94. doi: 10.1038/nrmicro3565
- Luna Galaz GA, Morales Erasto V, Peñuelas Rivas CG, Blackall PJ, Soriano Vargas E.** (2016). Antimicrobial Sensitivity of *Avibacterium paragallinarum* Isolates from Four Latin American Countries. *Avian Dis.* 60(3):673–6. doi: 10.1637/11398-022616-ResNote.1
- Mardis ER.** (2013). Next-generation sequencing platforms. *Annu Rev Anal Chem.* 287–303. doi: 10.1146/annurev-anchem-062012-092628
- McSweeney LA, Dreyfus LA.** (2005). Carbohydrate-binding specificity of the *Escherichia coli* cytolethal distending toxin CdtA-II and CdtC-II subunits. *Infect Immun.* 73(4):2051–60. doi:10.1128/IAI.73.4.2051–2060.2005
- Montes García JF, Vaca S, Vazquez Cruz C, Soriano Vargas E, Aguilar Romero F, Blackall PJ, et al.** (2016). Identification of a Hemagglutinin from *Gallibacterium anatis*. *Curr Microbiol.* 72(4):450–6. doi: 10.1007/s00284-015-0969-5
- Morales Erasto V, Falconi Agapito F, Luna Galaz GA, Saravia LE, Montalvan Avalos A, Soriano Vargas E, et al.** (2016). Coinfection of *Avibacterium paragallinarum* and *Ornithobacterium rhinotracheale* in Chickens from Peru. *Avian Dis.* 60(1):75–8. doi: 10.1637/11265-082015-ResNote.1
- Nietsch R, Haas J, Lai A, Oehler D, Mester S, Frese KS, et al.** (2016). The Role of Quality Control in Targeted Next-generation Sequencing Library Preparation. *Genomics, Proteomics Bioinforma.* 14(4):200–6. doi: 10.1016/j.gpb.2016.04.007
- Nishino K, Latifi T, Groisman EA.** (2006). Virulence and drug resistance roles of multidrug efflux systems of *Salmonella enterica* serovar *Typhimurium*. *Mol Microbiol.* 59(1):126–41. doi: 10.1111/j.1365-2958.2005.04940.x

Nuccio S-P, Bäumler AJ. (2007). Evolution of the Chaperone/Usher Assembly Pathway: Fimbrial Classification Goes Greek. *Microbiol Mol Biol Rev.* 71(4):551–75. doi: 10.1128/MMBR.00014-07

Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, et al. (2013). Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. 158–70. doi: 10.1007/978-3-642-37195-0_13

Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42(D1):206–14. doi: 10.1093/nar/gkt1226

Page LA. (1962). *Haemophilus* infections in chickens. I. Characteristics of 12 *Haemophilus* isolates re-covered from diseased chickens. *Am. J. Vet. Res.* 23: 85-95. 1962

Pan YC, Tan DH, Shien JH, Liu CC, He YS, Shen PC, et al. (2012). Identification and Characterization of an RTX Toxin-Like Gene and Its Operon from *Avibacterium paragallinarum*. *Avian Dis Dig.* 7(3):e47–8. doi: 10.1637/10047-122211-Reg.1

Persson G, Pors SE, Thøfner ICN, Bojesen AM. (2018). Vaccination with outer membrane vesicles and the fimbrial protein FlfA offers improved protection against lesions following challenge with *Gallibacterium anatis*. *Vet Microbiol.* 217:104–11. doi: 10.1016/j.vetmic.2018.03.010

Pis Diez CM, Juncos MJ, Villarruel Dujovne M, Capdevila DA. (2022). Bacterial Transcriptional Regulators: A Road Map for Functional, Structural, and Biophysical Characterization. *Int J Mol Sci.* 23(4). doi: 10.3390/ijms23042179

Pollard MO, Gurdasani D, Mentzer AJ, Porter T, Sandhu MS. (2018). Long reads: Their purpose and place. *Hum Mol Genet.* 27(R2):234–41. doi: 10.1093/hmg/ddy177

Pors SE, Pedersen IJ, Skjærning RB, Thøfner ICN, Persson G, Bojesen AM. (2016). Outer membrane vesicles of *Gallibacterium anatis* induce protective immunity in egg-laying hens. *Vet Microbiol.* 195:123–7. doi: 10.1016/j.vetmic.2016.09.021

Ramón Rocha MO, García González O, Pérez Méndez A, Ibarra Caballero J, Pérez Márquez VM, Vaca S, et al. (2006). Membrane vesicles released by *Avibacterium paragallinarum* contain putative virulence factors. *FEMS Microbiol Lett.* 257(1):63–8. doi: 10.1111/j.1574-6968.2006.00154.x

Ramos Salas DF. (2023). Estudio de expresión de genes relacionados con el metabolismo de carbohidratos (Condroitina) y virulencia de *Avibacterium paragallinarum*. Tesis para obtener el grado de maestra en ciencias (Microbiología). Benemérita Universidad Autónoma de Puebla.

Requena D, Chumbe A, Torres M, Alzamora O, Ramirez M, Valdivia Olarte H, et al. (2013). Genome sequence and comparative analysis of *Avibacterium paragallinarum*. *Bioinformation.* 9(10):528–36. doi: 10.6026/97320630009528

Rivero García PC, Cruz CV, Alonso PS, Vaca S, Negrete Abascal E. (2005). *Haemophilus paragallinarum* secretes metalloproteases. *Can J Microbiol.* 51(10):893–6. doi: 10.1139/w05-067

Rhoads A, Fai AU K. (2015). PacBio Sequencing and Its Applications. Elsevier Enhanced Reader.pdf. Vol. 13, Genomics Proteomics Bioinformatics. 278–89. doi: 10.1016/j.gpb.2015.08.002

- Sakamoto R, Baba S, Ushijima T, Kino Y, Honda T, Mizokami H, et al.** (2013). Development of a recombinant vaccine against infectious coryza in chickens. *Res Vet Sci.* 94(3):504–9. doi: 10.1016/j.rvsc.2012.10.027
- Schatz MC, Delcher AL, Salzberg SL.** (2010). Assembly of large genomes using second-generation sequencing. *Genome Res.* 20(9):1165–73. doi: 10.1101/gr.101360.109
- Sharma AK, Dhasmana N, Dubey N, Kumar N, Gangwal A, Gupta M, et al.** (2017). Bacterial Virulence Factors: Secreted for Survival. *Indian J Microbiol.* 57(1):1–10. doi: 10.1007/s12088-016-0625-1.
- Shenker BJ, McKay T, Pankoski L, Zekavat A, Donald R.** (2004). *Actinobacillus actinomycetemcomitans* Cytolethal Distending Toxin (Cdt): Evidence That the Holotoxin Is Composed of Three Subunits: CdtA, CdtB, and CdtC. *The Journal of Immunology.* 172(1):410-7. doi: 10.4049/jimmunol.172.1.410
- Shenker BJ, Ojcius DM, Walker LP, Zekavat A, Scuron MD, Boesze-Battaglia K.** (2015). *Aggregatibacter actinomycetemcomitans* cytolethal distending toxin activates the NLRP3 inflammasome in human macrophages, leading to the release of proinflammatory cytokines. *Infect Immun.* 83(4):1487–96. doi: 10.1128/IAI.03132-14
- Shima A, Hinenoya A, Asakura M, Sugimoto N, Tsukamoto T, Ito H, et al.** (2012). Molecular characterizations of cytolethal distending toxin produced by *Providencia alcalifaciens* strains isolated from patients with diarrhea. *Infect Immun.* 80(4):1323–32. doi:10.1128/IAI.05831-11
- Singh G, Yadav M, Ghosh C, Rathore JS.** (2021). Bacterial toxin-antitoxin modules: classification, functions, and association with persistence. *Curr Res Microb Sci.* 2:100047. doi: 10.1016/j.crmicr.2021.100047
- Sivashankari S, Shanmughavel P.** (2007). Comparative genomics - A perspective. *Bioinformation.* 1(9):376–8. doi: 10.6026/97320630001376
- Slatko BE, Gardner AF, Ausubel FM.** (2018). Overview of Next Generation Sequencing technologies (and bioinformatics) in cancer. *Mol Biol.* 122(1):1–15. doi:10.1002/cpmb.59.
- Stahl M, Friis LM, Nothaft H, Liu X, Li J, Szymanski CM, et al.** (2011). L-fucose utilization provides *Campylobacter jejuni* with a competitive advantage. *Proc Natl Acad Sci U S A.* 108(17):7194–9. doi: 10.1073/pnas.1014125108
- Staub L, Fuchs TM.** (2015). Regulation of fucose and 1,2-propanediol utilization by *Salmonella enterica* serovar Typhimurium. *Front Microbiol.* 1–11. doi: 10.3389/fmicb.2015.01116
- Studholme DJ, Buck M.** (2000). The biology of enhancer-dependent transcriptional regulation in bacteria: Insights from genome sequences. *FEMS Microbiol Lett.* 186(1):1–9. doi: 10.1111/j.1574-6968.2000.tb09074.x
- Swamy MK, Akhtar MS, Sinniah UR. Swamy MK, Akhtar MS, Sinniah UR.** (2016). Antimicrobial Properties of Plant Essential Oils against Human Pathogens and Their Mode of Action: An Updated Review. *Evid Based Complement Alternat Med.* 2016:3012462. doi: 10.1155/2016/3012462.

Swayne DE, Glisson JR, McDougald LR, Nolan Lisa, Suarez DL, et al. (2008) Diseases of Poultry. (12th edn), Blackwell Publishing Ltd., Iowa, USA.

Tapprich WE, Hill WE. (1986). Involvement of bases 787-795 of *Escherichia coli* 16S ribosomal RNA in ribosomal subunit association. Proc Natl Acad Sci. 83(3):556–60. doi: 10.1073/pnas.83.3.556

Taylor CM, Roberts IS. (2005). Capsular polysaccharides and their role in virulence. Contrib Microbiol. 12:55–66. doi: 10.1159/000081689

Terzolo H, Moreira A. (2005). *AVIBACTERIUM, HAEMOPHILUS E HISTOPHILUS*. Chapter 36, 2005:337–54.

Touchman, J. (2010) Comparative Genomics. *Nature Education Knowledge* 3(10):13

Tu T-Y, Hsieh M-K, Tan D-H, Ou S-C, Shien J-H, Yen T-Y, et al. (2015). Loss of the Capsule Increases the Adherence Activity but Decreases the Virulence of *Avibacterium paragallinarum*. Avian Dis. 59(1):87–93. doi: 10.1637/10937-091414-Reg

Van NTB, Yen NTP, Nhung NT, Cuong N Van, Kiet BT, Hoang N Van, et al. (2019). Characterization of viral, bacterial, and parasitic causes of disease in small-scale chicken flocks in the Mekong Delta of Vietnam. 99(2):783–90. doi: 10.1016/j.psj.2019.10.033

Vandecraen J, Chandler M, Aertsen A, Van Houdt R. (2017). The impact of insertion sequences on bacterial genome plasticity and adaptability. Crit Rev Microbiol. 43(6):709–30. doi: 10.1080/1040841X.2017.1303661

Velázquez Valtierra JP. (2017). Estudio genético e inmunogénico comparativo de fimbrias F17 en *Gallibacterium anatis* y *Avibacterium paragallinarum*. Tesis para obtener el grado de maestro en ciencias (Microbiología). Benemérita Universidad Autónoma de Puebla.

Vigil-Stenman T, Ininbergs K, Bergman B, Ekman M. (2017). High abundance and expression of transposases in bacteria from the Baltic Sea. ISME J. 11(11):2611–23. doi: 10.1038/ismej.2017.114

Volpi N. (2006). Therapeutic Applications of Glycosaminoglycans. Curr Med Chem. 13(15):1799–810. doi: 10.2174/092986706777452470

Volpi N. (2019). Chondroitin sulfate safety and quality. Molecules. 24(8). doi:10.3390/molecules24081447

Wagner PL, Waldor MK. (2002). Bacteriophage Control of Bacterial Virulence. Society. 70(8):3985–93. doi: 10.1128/IAI.70.8.3985-3993.2002

Wang YP, Hsieh MK, Tan DH, Shien JH, Ou SC, Chen CF, et al. (2014). The haemagglutinin of *Avibacterium paragallinarum* is a trimeric autotransporter adhesin that confers haemagglutination, cell adherence and biofilm formation activities. Vet Microbiol. 174(3–4):474–82. doi: 10.1016/j.vetmic.2014.10.013

Werneburg GT, Thanassi DG. (2018). Pili Assembled by the Chaperone/Usher Pathway in *Escherichia coli* and *Salmonella*. EcoSal Plus. 8(1). doi: 10.1128/ecosalplus.ESP-0007-2017.

- Winther K, Tree JJ, Tollervey D, Gerdes K.** (2016). VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation. *Nucleic Acids Res.* 44(20):9860–71. doi: 10.1093/nar/gkw781
- Wood TL, Wood TK.** (2016). The HigB/HigA toxin/antitoxin system of *Pseudomonas aeruginosa* influences the virulence factors pyochelin, pyocyanin, and biofilm formation. *Microbiologyopen.* 5(3):499–511. doi: 10.1002/mbo3.346
- Wu JH, Wu AM, Tsai CG, Chang XY, Tsai SF, Wu TS.** (2008). Contribution of fucose-containing capsules in *Klebsiella pneumoniae* to bacterial virulence in mice. *Exp Biol Med (Maywood).* 233(1):64-70. doi: 10.3181/0706-RM-170.
- Wu JR, Wu YR, Shien JH, Hsu YM, Chen CF, Shieh HK, et al.** (2011). Recombinant proteins containing the hypervariable region of the haemagglutinin protect chickens against challenge with *Avibacterium paragallinarum*. *Vaccine.* 29(4):660–7. doi: 10.1016/j.vaccine.2010.11.040
- Xie HX, Nie P, Chang MX, Liu Y, Yao WJ.** (2005). Gene cloning and functional analysis of glycosaminoglycan-degrading enzyme chondroitin AC lyase from *Flavobacterium columnare* G4. *Arch Microbiol.* 184(1):49–55. doi: 10.1007/s00203-005-0009-0
- Xu F, Miao D, Du Y, Chen X, Zhang P, Sun H.** (2013). Draft Genome Sequence of *Avibacterium paragallinarum* Strain 221. *Genome Announc.* 1(3):28–9. doi: 10.1128/genomeA.00290-13
- Yurgel SN, Rice J, Mulder M, Kahn ML.** (2010). GlnB/GlnK PII proteins and regulation of the *Sinorhizobium meliloti* Rm1021 nitrogen stress response and symbiotic function. *J Bacteriol.* 192(10):2473–81. doi: 10.1128/JB.01657-09
- Zepeda AA, Ramírez S, Vega V, Morales V, Talavera M, Bojesen AM, et al.** (2009). Hemagglutinating Activity of *Gallibacterium* Strains. *American Association of Avian Pathologists.* 53(1):115–8. doi: 10.1637/8375-060908-ResNote.1

Anexo de soluciones

Ampicilina 100 mg/ml (Stock)

- ❖ Ampicilina 1 g
- ❖ Agua inyectable estéril 10 ml

Amikacina 250 mg/ml (Stock)

- ❖ Amikacina 250 mg/ml
- ❖ Agua inyectable estéril 1 ml

Agar sangre 100 ml

- ❖ Base de agar sangre 4 g
- ❖ Sangre de carnero 5%
- ❖ Resuspender en agua tridestilada y aforar a 100 ml
- ❖ Esterilizar 15 lb por 15 min

Acetato de sodio (CH₃COONa) 3 M, pH=7, 50 ml

- ❖ Acetato de sodio 12.30 g
- ❖ Resuspender en agua tridestilada y ajustar pH con ácido clorhídrico
- ❖ Aforar a 50 ml,
- ❖ Esterilizar 15 lb por 15 min

Acetato de amonio (CH₃COONH₄) 7.5 M, 20 ml

- ❖ Acetato de amonio 11.562 g
- ❖ Agua tridestilada 20 ml
- ❖ Esterilizar 15 lb por 15 min

Acetato de potasio 5 M, 100 ml

- ❖ Acetato de potasio 49.07 g
- ❖ Aforar a 100 ml
- ❖ Esterilizar 15 lb por 15 min.

Acrilamida 30%

- ❖ Acrilamida 29 g
- ❖ Bisacrilamida 1 g
- ❖ Aforar a 100 ml con agua tridestilada estéril

Buffer de lisis pH=8, 100 ml

- ❖ Tris 50 mM (20 ml de Tris 250 mM, pH=8)
- ❖ EDTA 10 mM (2 ml de EDTA 0.5 M, pH=8)
- ❖ NaCl 10 mM (0.06 gr)
- ❖ Aforar a 100 ml con agua tridestilada
- ❖ Ajustar pH=8

- ❖ Esterilizar 15 lb por 15 min.

Buffer de carga (electroforesis de DNA)

- ❖ Glicerol 50%
- ❖ EDTA 1 mM, pH=8
- ❖ Azul de bromofenol 1.25%
- ❖ Xylencianol FF 0.25%

Buffer de carga (SDS-PAGE)

- ❖ Tris-HCl pH=6.8, 100 mM
- ❖ β - Mercaptoetanol 5%
- ❖ SDS 4%
- ❖ Bromofenol Blue 0.2%
- ❖ Glicerol 20%

Cloranfenicol 5 mg/ml (Stock)

- ❖ Cloranfenicol 5 mg
- ❖ Agua inyectable estéril 1 ml

Estreptomicina 85 mg/ml (Stock)

- ❖ Estreptomicina 85 mg
- ❖ Agua inyectable estéril 1 ml

Fenol-Tris pH=8, 400 ml

- ❖ Descongelar el fenol puro a baño María (de 60°C a 70°C) y tomar la cantidad a preparar.
- ❖ Añadir 1 volumen de Tris-HCL pH=9, 1M y con ayuda de una bala magnética agitar por 3 h, una vez transcurrido este tiempo dejar separar toda la noche en refrigeración.
- ❖ Quitar el sobrenadante y agregar 1 volumen de Tris HCL pH=8, 0.5 M, con ayuda de una bala magnética agitar por 3 h, una vez transcurrido este tiempo dejar separar toda la noche en refrigeración.
- ❖ Quitar el sobrenadante y agregar 1 volumen de Tris HCL pH=8, 0.1M, con ayuda de una bala magnética agitar por 3 h, medir pH con tira reactiva (pH=8) y dejar reposar.

Fenol-Cloroformo-Alcohol isoamílico 50 ml

- ❖ 25 ml de fenol-Tris
- ❖ 24 ml de cloroformo
- ❖ 1 ml de alcohol isoamílico

Gentamicina 80 mg/ml (Stock)

- ❖ Gentamicina 80 mg
- ❖ Agua inyectable estéril 1 ml

Gel de acrilamida 10%

Gel separador 15 ml

- ❖ Agua destilada estéril 5.9 ml
- ❖ Acrilamida 30%, 5 ml
- ❖ Tris 1.5 M, pH=8.8, 3.8 ml
- ❖ SDS 10%, 0.15 ml
- ❖ PSA 10%, 0.15 ml
- ❖ TEMED 0.006 ml

Gel concentrador 6 ml

- ❖ Agua destilada estéril 4.1 ml
- ❖ Acrilamida 30%, 1 ml
- ❖ Tris 1.5 M, pH=6.8, 0.75 ml
- ❖ SDS 10%, 0.06 ml
- ❖ PSA 10%, 0.06 ml
- ❖ TEMED 0.006 ml

Hidróxido de sodio (NaOH)10N 20 ml

- ❖ NaOH 8 g
- ❖ Aforar a 20 ml con agua tridestilada estéril

Medio Infusión Cerebro Corazón 100 ml

- ❖ Infusión Cerebro Corazón 5.2 g
- ❖ Aforar con agua tridestilada a 100 ml
- ❖ Esterilizar 15 lb por 15 min

Medio Infusión cerebro corazón con yema de huevo

- ❖ 5 ml de medio BHI
- ❖ NAD 10 µg/ml de medio
- ❖ Suero de caballo 5%
- ❖ Yema de huevo 500 µl

Medio Infusión cerebro corazón con sulfato de condroitina 250 ml

- ❖ 13 g de agar BHI
- ❖ Agregar 4 mg/ml de sulfato de condroitina
- ❖ Esterilizar 15 lb 15 min.

Medio Luria Bertani (LB) 250 ml

- ❖ 2.5 g de peptona de caseína
- ❖ 1.25 g extracto de levadura
- ❖ 2.5 g Cloruro de sodio
- ❖ Aforar a 250 ml
- ❖ Esterilizar 15 lb por 15 min

NADH 10 mg/ml, 10 ml

- ❖ NADH 0.100 g

- ❖ Agua inyectable 10 ml
- ❖ Esterilizar por filtración con filtro 0.022 μm

Persulfato de amonio (PSA) 10%

- ❖ PSA 0.1 g
- ❖ Agua destilada estéril 1 ml

Proteasa 250 mg/ml, 1 ml (solución stock)

- ❖ Proteasa 0.250 g
- ❖ Tris EDTA 1 ml

RNasa 10 mg/ml, 1 ml (solución stock)

- ❖ RNasa 10 mg
- ❖ Solución (Tris-HCl 10 mM, pH=7.5 y NaCl 15 MM) 1 ml
- ❖ Calentar a 100°C durante 15 min.
- ❖ Enfriar a temperatura ambiente

Sarcosil 20%

- ❖ Sarcosil 5 g
- ❖ Aforar a 25 ml con agua tridestilada
- ❖ Esterilizar 15 lb por 15 min

SDS 10%, 200 ml

- ❖ SDS 40 g
- ❖ Aforar a 200 ml con agua tridestilada
- ❖ Esterilizar 15 lb por 15 min

Solución de lisis

- ❖ Lisozima 0.1 mg/ml
- ❖ RNasa 10 $\mu\text{g/ml}$

Solución desteñidora 100 ml

- ❖ Ácido acético glacial 10 ml
- ❖ Agua destilada estéril 90 ml

Solución fijadora 100 ml

- ❖ Metanol 20 ml
- ❖ Ácido acético glacial 10 ml
- ❖ Agua destilada estéril 70 ml

Tampón I

- ❖ Tris 0.025 M pH=8
- ❖ Sarcosil 25%
- ❖ EDTA 0.250M

- ❖ Ficoll 400 7%
- ❖ Se prepara hasta que va a ser usado

Tampón II

- ❖ Tris 0.025 M pH=8
- ❖ Sarcosil 10%
- ❖ Sacarosa 25%
- ❖ Azul de bromofenol 0.07%
- ❖ Glicerol 7%
- ❖ Se prepara hasta que va a ser usado

TBE 10X, 1L

- ❖ Tris-base 108 g
- ❖ Ácido bórico 55 g
- ❖ EDTA 0.5 M 40 ml
- ❖ Se esteriliza 15bl por 15 min

Tetraciclina 25 mg/ml (Stock)

- ❖ Tetraciclina 250 mg
- ❖ Etanol absoluto 7 ml
- ❖ Agua inyectable estéril 3 ml

Tris 1.5 M, pH=8.8, 100 ml

- ❖ Trizma base 18.165 g
- ❖ Ajustar pH con ácido clorhídrico
- ❖ Aforar a 100 ml con agua tridestilada
- ❖ Se esteriliza a 15 lb por 15 min

Tris 1.5 M, pH=6.8, 100 ml

- ❖ Trizma base 18.165 g
- ❖ Ajustar pH con ácido clorhídrico
- ❖ Aforar a 100 ml con agua tridestilada
- ❖ Esterilizar a 15 lb por 15 min

Tris-base pH=9, 1M, 400 ml

- ❖ Trizma base 48.45 g
- ❖ Disolver en agua tridestilada y ajustar pH con ácido clorhídrico
- ❖ Aforar a 400 ml con agua tridestilada
- ❖ Esterilizar a 15 lb por 15 min

Tris-EDTA 10 ml

- ❖ Tris HCl 250, mM pH=8, 400 ml
- ❖ EDTA 0.5 M, pH=8, 20 µl
- ❖ H₂O estéril 9580 µl

Tris glicina 5X, 1L

- ❖ Trizma base 15.1 g
- ❖ Glicina 94 g
- ❖ SDS 10% 50 ml
- ❖ Esterilizar 15 lb, 15 min

Tris-HCL 250 mM, pH=8, 50 ml

- ❖ Trizma base 1.514 g
- ❖ Disolver en agua tridestilada y ajustar pH con ácido clorhídrico
- ❖ Aforar a 50 ml.
- ❖ Esterilizar 15 lb por 15 min.

Tris HCL pH=8, 0.5 M, 400 ml

- ❖ Trizma base 24.22 g
- ❖ Disolver en agua tridestilada y ajustar pH con ácido clorhídrico
- ❖ Aforar a 400 ml

Tris HCL pH=8, 0.1M, 400 ml

- ❖ Trizma base 4.84 g
- ❖ Disolver en agua tridestilada y ajustar pH con ácido clorhídrico
- ❖ Aforar a 400 ml

Tris HCL 50 mM, pH=7.5 con lisozima

- ❖ Tris HCL 50 mM pH=7.5 1 ml
- ❖ Lisozima 0.010 g

XII. MANUSCRITO DEL ARTICULO PUBLICADO

A Maverick-like cluster in the genome of a pathogenic, moderately virulent strain of *Gallibacterium anatis*, ESV200, a transient biofilm producer

Patricia Sanchez-Alonso^{1,§, †}, Elena Cobos-Justo^{1, †}, Miguel Angel Avalos-Rangel¹, Lucía López-Reyes¹, Gloria Paniagua-Contreras², Felipe Vaca-Paniagua², Estela Anastacio-Marcelino¹, Ana Jaqueline López-Ochoa¹, Victor M Pérez Marquez³, Erasmo Negrete-Abascal², Candelario Vázquez-Cruz^{1, §}.

1Centro de Investigaciones en Ciencias Microbiológicas, Instituto de Ciencias, Benemérita Universidad Autónoma de Puebla, Puebla México. 2Facultad de Estudios Superiores de Iztacala, UNAM, Los Reyes Iztacala, Estado de México, 3BIOVETSA Tehuacán Puebla, México.

§ Corresponding Author

† These authors contributed equally to this work and share first authorship

§Corresponding author: Centro de Investigaciones en Ciencias Microbiológicas,

Instituto de Ciencias Universidad Autónoma de Puebla, PO 72570, Puebla, Pue. Mexico.

Telephone: +52 222-229-5500 ext. 2536 Email: ecobacilos@yahoo.com

Running Title: Pathogenicity and biofilm of *G. anatis*_ESV200

KEYWORDS: *Gallibacterium anatis*; genome; Maverick, fimbriae; biofilm; pathogenesis; avian diseases

Abstract

Gallibacterium anatis causes gallibacteriosis in birds. These bacteria produce biofilms and secrete fimbrial appendages as tools to cause disease in animals. *G. anatis* strains contain up to three types of fimbriae. The fimbrial coding genes are part of modular operons classified by the encoded Usher chaperone. *Complete genome sequencing* is the strategy currently used to determine variations in the contents of fimbrial genes present in the *G. anatis* strains with certainty. Although only the genome of the UMN179 strain is complete, others are available in the database as a draft. The ESV200 strain reported here differs from the 12656-12 strain because it produces a biofilm at 20%. In liquid brain heart infusion medium, strain ESV200 appeared as a homogeneous suspension, while the 12656-12 culture became clumpy. Protein analysis of cultures of strain ESV200 using denaturing gel electrophoresis revealed several differences between protein bands that might correspond to fimbriae because fimbrial genes are transcribed as mRNAs. The amount of biofilm and the electrophoretic-protein profile were affected by DNase, proteinase K, and pronase E treatments, suggesting that the ESV200 strain produces a typical biofilm but with a different structure than that observed for strain 12656-12. ESV200 and 12656-12 strains are pathogenic for chickens, although the latter is the most virulent. Here, we show that the complete genome of the ESV200 strain is similar to that of the UMN179 strain. However, these strains have evolved with many structural rearrangements; the most striking chromosomal arrangement is a Maverick-like element present in the ESV200 strain.

1. Introduction

Gallibacterium anatis is a gram-negative, pleomorphic and capsulated gammaproteobacterium that lacks flagellar motility. The genus *Gallibacterium* was taxonomically classified as a new member of the *Pasteurellaceae* Pohl family in 2003 by Christensen et al. by reclassification of some biovars from *Pasteurella haemolytica*, “*Actinobacillus salpingitidis*” or *Pasteurella anatis* using pulsed-field gel electrophoresis (PFGE), amplified fragment length

polymorphism (AFLP), and the amplification patterns of the 16S rRNA, 23S rRNA, and ITS sequences (Christensen et al., 2003; Bojesen et al., 2007b). *G. anatis* has been isolated from the reproductive and upper respiratory organs of healthy birds as part of the microbiota (Bisgaard 1977). It has also been isolated in pure culture form and found to be associated with pathogens such as avian pathogenic *Escherichia coli* (APEC) in clinical cases of salpingitis, peritonitis, hepatitis, and septicemia in birds (Neubauer et al., 2009). These findings have led to controversy regarding its pathogenic character and its potential to cause pathogenic lesions that affect reproductive organs and egg laying in farm chickens (Bojesen et al., 2003, Bojesen et al., 2004).

However, to date, *G. anatis* has been studied as an opportunistic pathogen that can colonize chickens at four days of age and persist until they grow into laying hens (Huangfu et al., 2012). The detrimental effects of this genus on chicken breeding have enhanced its importance as an economic risk, and hence, specific tests to unequivocally identify *Gallibacterium* species have been sought and reported (Bojesen et al., 2007a; Alispahic et al., 2011).

The bacterium *G. anatis* possesses several virulence factors, such as the toxin GtxA, capsule, metalloproteases (Garcia-Gomez et al., 2005), hemagglutinins, outer membrane proteins and vesicles (Chantes-Guerra et al., 2022), biofilm production genes, and gene clusters that encode F17-like fimbriae (Kudirkiene et al., 2014; Pors et al., 2016), similar to those frequently found in several pathotypes of *E. coli*, including APEC (Bager et al., 2013a; Bertelli et al., 2017; Persson and Bojesen 2015). Among these, the toxin GtxA, which is similar to HlyA from *E. coli*, is usually identified in wild isolates, as it enables the pathogen to cause cell damage and induce apoptosis, and the presence of this toxin can be used to classify strains as hemolytic and nonhemolytic in blood agar tests (Kristensen et al., 2011).

Biofilms represent consortium-style microbial growth in which individuals of the same or different species grow embedded in a matrix composed of exopolysaccharides, proteins and deoxyribonucleic acid (DNA) (Donland 2002; Di martino 2016). Data from De Oliveira-García et al. (2003) suggest that F17-like fimbriae could be important for bacterial adhesion to mammalian cells, hemagglutination, and biofilm formation. Similar data have been reported by Wurpel et al. (2016), who found that fimbriae are involved in biofilm formation and adherence to human cells. Interesting data are obtained when comparing natural bacterial isolates; a majority of *G. anatis* strains produce filaments arranged as latices and form biofilms, exhibiting adhesion to cells and inert surfaces (Vaca et al., 2011, Salgado-Lucio et al. 2012), and it has been proposed that some *G. anatis* lineages carry different F17-like fimbriae and are more virulent (Bager et al., 2013b; Pors et al., 2016). The molecular classification of the F17-like genes into operons was performed by Kudirkiene et al. (2014). If biofilms are important for attachment to and colonization of animal

tissues by pathogenic bacteria, strains deficient in biofilm formation could exhibit altered virulence. The goal of this research was to investigate the participation of F17-like fimbriae in agglutination, biofilm formation, and pathogenesis and to determine the full genome sequence of the *G. anatis* strain ESV200 showing a dispersible phenotype.

2. Material and Methods

2.1 Strains and media

The *G. anatis* strains used in this work were 12656-12 (β -hemolytic biovar, isolated from chicken liver by Bojesen et al. in 2003) and ESV200 (β -hemolytic, isolated in Tepatitlan, Jalisco, Mexico (geographical coordinates: 20° 49' 0" north, 102° 44' 0" west), in 2009 from hens with respiratory disease by Dr. Edgardo Soriano). Strain purity was checked by streaking on blood agar medium with incubation at 37°C for 24 hours. For growth in liquid medium, strains were grown at 37°C for 24 hours in brain-heart infusion (BHI) medium (Merck[®]) with aeration, unless specified otherwise. Precultures were incubated until the optical density at 595 nm (OD₅₉₅) was 1.0 before inoculation in BHI broth in either flasks or microplates (Chantes-Guerra et al., 2022).

2.2 Biofilm formation

Biofilm formation assays were performed as described by Johnson et al. (2013), with some changes as described here, using 24-well cell-culture microplates (TrueLine, LabSource Northlake, IL, USA). For the first inoculum, *G. anatis* strains were grown in tubes containing 5 ml of liquid BHI medium at 37°C for 18 hours with aeration. Bacterial cells were concentrated by centrifugation and adjusted to an OD₆₀₀ of 2.0, and 1:100 dilutions of each strain in liquid BHI medium were prepared. For each dilution, five wells were filled with 1.5 ml of medium, and 4 of the 24 wells on the plate were reserved for the control, which consisted of BHI inoculated with phosphate-buffered saline (PBS) as described by Crampton et al. (1999). The subsequent assays were performed in triplicate, and the plates were incubated at 37°C under static conditions for 24, 48, 72, and 96 hours. After each incubation period, the culture medium and planktonic cells were removed carefully from the wells, and each well was washed twice with triple-distilled water to avoid shearing the biofilm mechanically via water pressure or contact with the pipette or removal

of the biofilm from the surface. The plates with biofilms were completely dried by placing them upside down on blotting paper, and each well was filled with 1.5 ml of 0.1% violet crystal, after which the plates were allowed to stand for 30 min at room temperature. The wells were washed four times as described above and dried at room temperature for approximately 45 to 60 min. Two milliliters of acetic acid (33% v/v) were added to each well to solubilize the crystal violet, which was quantified by measuring the OD₆₃₀ in disposable polystyrene cuvettes. To avoid variations in the analysis due to the volatility of acetic acid, the solubilization procedure was performed in five sample blocks in one row at a time to prevent acetic acid evaporation. Normalization of the biofilm data was performed using duplicates of each treatment and each replicate. For each cell culture in the plate used for biofilm quantification, another plate was used for cell quantification by resuspending the biofilms, homogenizing and measuring the OD₆₀₀.

For protein analysis, biofilms were recovered from large-volume cultures of *G. anatis* strains. For this purpose, 150 ml of BHI medium was used to dilute the standardized inoculum at 1:100 as described above. The inoculated medium was placed in 1 L glass bottles placed laying on their sides, and the bottles were incubated statically at 37°C for 24, 48, 72 and 96 hours. The medium and planktonic culture were removed carefully by decantation. To recover the biofilms, 2 ml of 50 mM Tris buffer (pH 7.5) was added, and glass beads were used to release the biofilm layer, which was recovered in 15 ml Falcon tubes. The remaining residues were recovered with 500 µl of 50 mM Tris buffer (pH 7.5). The samples were stored at -80°C. The recovered biofilms were subjected to a complex analysis of the changes in their absorbance and protein profiles after treatment with 4 M urea for 30 min, DNase (100 µg/mL), lysozyme (50 µg/mL), proteinase K (50 µg/mL) and pronase E (100 µg/mL). Denaturing 10% polyacrylamide gel electrophoresis (SDS-PAGE) of the biofilm proteins was performed to explore and distinguish planktonic cell growth from adherent cell growth, as reported previously (Lopez Ochoa et al., 2017).

2.3 Nucleic acid manipulation

Total DNA from *G. anatis* was obtained from broth cultures grown at 37°C for 24 hours. A total of 1.5 ml of the culture was centrifuged at 13,000 rpm for 5 min; the cell pellet was recovered, 0.5 ml of lysis buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 10 mM NaCl) was added, and the sample was homogenized. Then, 20 mg of lysozyme were added and the sample was incubated at 37°C for

20 min; afterward, 40 µl of 1% sodium dodecyl sulfate (SDS) were added, and the sample was homogenized. The lysate was treated with 200 µl of phenol saturated with a solution of Tris and EDTA (or TE) (Sambrook et al 2001), stirred and centrifuged at 13,000 rpm for 10 min. One volume of phenol-TE/chloroform/isoamyl alcohol (25:24:1, v/v) was added to the aqueous phase, and the sample was homogenized and centrifuged at 13,000 rpm for 10 min. The supernatant was recovered, washed twice with chloroform, and centrifuged at 13,000 rpm for 5 min; the upper phase was recovered in a new tube, either 0.1 volume of 3 M sodium acetate (pH 7) or 0.5 volumes of 7.5 M ammonium acetate (pH 7) was added, and the sample was homogenized. Total DNA was precipitated with cold (-20°C) 70% ethanol in water. The mixture was centrifuged at 13,000 rpm for 10 min, the supernatant was discarded, and the DNA pellet was washed with 1 ml of cold 70% ethanol. The DNA was dissolved in 100 µl of sterile distilled water and stored at -20°C until use.

Ribonucleic acid (RNA) was extracted from *G. anatis* under two different growth conditions, namely, agitation or static conditions, in which the OD₆₀₀ was 1 and 0.5, respectively. To obtain *G. anatis* ESV200 RNA, a QIAcube Connect (Qiagen® Cat. No./ID: 9002864) was used. In each case, 20 ml of the bacterial culture was centrifuged at 3,500 rpm for 10 min. The supernatant was discarded, and the excess broth was removed with a pipette tip. Then, 500 µl of lysis buffer (10 mM Tris-HCl (pH 8.5), 10 mM EDTA, 10 mM NaCl) was added, and the mixture was homogenized. Then, 500 µl of acid-chloroform-isoamyl phenol was added, and the sample was shaken manually for 2 min and centrifuged at 3,500 rpm for 10 min in a table top centrifuge. The supernatant was transferred to another tube, and 200 µl of 7.5 M ammonium acetate and 500 µl of acid-chloroform-isoamyl phenol were added. The sample was then mixed vigorously for 2 min before centrifugation at 3,500 rpm for 10 min. The aqueous phase was transferred to another tube and washed twice with chloroform. After centrifugation, the aqueous phase was precipitated with 2 volumes of cold absolute ethanol and allowed to precipitate overnight. To recover, the RNA mixture was centrifuged at 3,500 rpm for 10 min, and the pellet was washed twice with cold 70% ethanol. Once dry, the pellet was dissolved in 200 µl of diethyl pyrocarbonate (DEPC)-treated water. The nucleic acid purity was estimated by spectrophotometry based on its absorbance at 260, 230 and 280 nm; RNA integrity was checked by analyzing the electrophoretic profile in a 1% agarose gel, and the samples were stored at -80°C until use.

2.4 Pathogenicity tests

Bacteria were cultivated on 10% sheep blood agar at 37°C and incubated overnight in a candle jar. The bacterial inoculum was prepared by cultivation of each strain for 14 hours in BHI broth.

Specific pathogen-free chickens (Alpes I, Puebla, México) of mixed sexes were divided into two groups of 5 chickens each. The chickens were individually labeled and injected intravenously with 0.2 ml of *G. anatis* inoculum (approximately 1×10^8 viable cells). The negative control group was inoculated by intravenously injecting 0.2 ml of BHI medium.

Clinical examination, morbidity and mortality data were recorded for each bird group at seven days post-infection (dpi). Subsequently, the animals were euthanized, and the skin and mucosa (crest, chin, nostrils, eyes, skin, eyelids), respiratory system (infraorbital sinuses, trachea, syrinx, lungs), digestive system (oral cavity, esophagus, crop, proventriculus, gizzard, liver, pancreas, duodenum, jejunum and ileum, blind, rectum, tonsils, C, palatal fissure), circulatory and hemolymphatic apparatus (heart, spleen, pouch of Fabrizio, thymus), genitourinary apparatus (testis, ovary, oviduct, follicle, kidneys, ureters), and skeletal muscle system (bones, joints, plantar tissue, bone marrow) from all birds were examined for histopathological signs, followed by bacterial isolation. Samples of tissues with lesions were collected with sterilized cotton swabs, streaked on blood agar plates, and incubated for 24 hours at 37.5°C in a microaerophilic environment.

For histopathological analysis, the samples were placed in 10% neutral buffered formalin and processed, sectioned, and stained as described elsewhere. Microscopic lesions were scored on a scale of 0–3: 0= normal, 1= minimal changes, 2= mild changes, and 3= severe changes. When the observation fields of a single section had lesions with different severities, the lesion degree predominant in 50% or more of the observation fields was used for the aforementioned scoring. The average score for each group was determined by adding individual scores and dividing the total score by the number of slides examined.

The Kruskal–Wallis test was used to assess the differences in the median scores of microscopic lesions between the groups, and differences were considered significant at $p = 0.05$, as reported elsewhere. Chi-square tests were used to examine the differences between the groups in isolation (presence of the inoculated strains in either sinus) and morbidity rates (percentage of birds that showed any clinical signs during the 7-day observation period), with differences being considered significant at $p = 0.05$ (Zepeda, *et al.*, 2010).

2.5 Reverse transcription and end-point polymerase chain reaction (RT–PCR)

To analyze the changes in the transcriptional expression of the fimbrial and adhesin genes of both the 12656-12 and ESV200 strains by end-point RT–PCR, we used the UMN179 genome sequences (http://www.ncbi.nlm.nih.gov/genome/3653?project_id=66567 from Johnson et al., 2011) of the gene *Flf1_a* to design the primer pair *Fifa7upp* and *Fifa7rev* that amplified a 236-bp fragment (Table 1). The genome sequences of the 12656-12 strain (https://www.ebi.ac.uk/ena/data/view/GCA_000464615.2) were used for subsequent primer design: for the *Flf3_b*, *Flf_b*, *Flf3*, *Flf1*, and, *Flf*-related genes, pairs of primers were designed, whose names, primer sequences and sizes of the PCR amplicons are shown in Table 1. It was not necessary to make more primer designs because the size of the synthesized amplicons was appropriate to carry out the end-point PCR and qRT–PCR (see below). As a positive control, total DNA from *G. anatis* 12656-12 was used for PCR, and as a negative control, a reaction without reverse transcriptase was performed.

The reaction mixture for endpoint PCR contained 10 ng of total DNA from *Gallibacterium* strains as template, 0.3 µM each DNA oligonucleotide primer, 10 mM Tris-HCl (pH 8.4), 1.5 mM MgCl₂, 0.24 mM each dNTP and 0.3 U of Platinum *Taq*DNA polymerase (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). The mixture was homogenized and treated as follows: denaturation for 5 min at 95°C, followed by 30 cycles of the appropriate annealing temperature and 72°C for elongation in a T100 thermal cycler (Bio-Rad, Hercules, CA, USA). The amplified products were electrophoresed in a 0.8% agarose gel, stained with ethidium bromide solution (2 µg/ml) and observed under ultraviolet light.

Table 1. Oligonucleotide primers used for end point RT–PCR and qRT–PCR

Gene	Primer name	F/R*	Sequence (5' to 3')	Amplicon size (bp)
<i>flf1 a</i>	<i>Fifa7upp</i>	F	AATCAACTTTGTCTGGGGTAGGA	236
	<i>Fifa7rev</i>	R	TGCTGCAGATAATTCAAAAAGACG	
<i>flf3 b</i>	<i>FifAOPdir</i>	F	ATGGTGCTTTTGCTAGTGGTGAA	281
	<i>FifAOPrev</i>	R	AGGCTCCTTGTCTGCATAATAACG	
<i>flf b</i>	<i>FifARTup</i>	F	CTGGTGATGCAGCTCAAAATG	141
	<i>FifARTrev</i>	R	CAATGTGGCTGTTTTCTTGC	
<i>flf3 D</i>	<i>Adhlike31TRDir</i>	F	TCGATGATGGCTTTGGTGA	223
	<i>Adhlike31TRRev</i>	R	CCGTTGGATTTCGTTACA	

<i>flf D</i>	FifG7Dir	F	CAGGTCTACAAATTACGCCAAG	147
	Fifg7Rev	R	CGTGTCTGCATCAGAAAAGG	
<i>flf D</i> -related	22AdhlikeTREDir	F	GGTAAATCTCGCCCCCTGTT	189
	22AdhlikeTRERev	R	TTTCTTTTAGTACTACCGTTCTGG	
<i>rpoA</i>	rnapolMdir	F	TACCAACAAACGACCGATAGG	244
	rnapolMrev	R	TAAAGTGCAGAATAAAGATGACG	
<i>fusA</i>	EfactorGFMDir	F	TGAATGCCAATGCTGAGAATG	248
	EfactorGFMRv	R	TGGCGGTGAAGAATAACTG	

*F/R indicates the forward or reverse primer for each gene target

2.6 Quantitative and end-point PCR

To analyze the changes in transcriptional expression by RT-PCR, the Ct cycle method was used. Using the constitutively expressed genes as a reference, we chose the hypothetical 901-bp *rpoA* gene, which encodes the RNA-polymerase alpha-subunit RpoA, as well as the 2,106-bp putative *fusA* gene from *G. anatis* UMN179, which encodes a hypothetical elongation factor G (EFG). The primers for these two genes and the primers used for analysis of the expression of the fimbrial genes are listed in table 1. Due to the small size of the amplicons and the fact that the primers were well evaluated for their amplification efficiency, the same primers were used in the end-point PCR and quantitative RT-PCR. We expected that the transcripts of the housekeeping genes would show stable levels of expression under different conditions for normalization, and we chose to apply the formula described by Pfaffl (2001).

Real-time PCR was performed in an IQTM5 Multicolor Real Time PCR Detection System (Bio-Rad). For each qPCR sample, three technical replicates and two biological replicates were examined. For each primer pair, a no-template control (NTC) was included to detect the formation of primer dimers. The reaction mixture was prepared at room temperature as follows: 1 µl of cDNA, 1 µl of 10 µM forward primer, 1 µl of 10 µM reverse primer, 10 µl of SYBR Green PCR master mix (Thermo Scientific) and 7 µl of water. The SYBR Green mix already contains DNA polymerase, SYBR Green and a passive reference dye (ROX), along with the required reaction components. Individual concave-cap tubes were used. The reaction was carried out according to the supplier's instructions: 1 cycle of 95°C/10 min; 40 cycles of 95°C/15 s, 60°C/15 s; 72°C/30 s. For the denaturation curve, 60 cycles at 65°C were performed, with the time increasing by 5 s in each cycle. Once the experimental Ct values were obtained, the dynamic range was calculated using log₁₀ dilutions for the test, covering a range of 10⁰ to 10⁸ molecules. The efficiency of the qRT-PCRs was

determined, and those with efficiencies greater than 96% and an R^2 value greater than 0.86 were considered valid. The supplementary material includes graphs showing the expression of some genes.

2.7 Genomic analysis of *G. anatis* ESV200

The total DNA of *G. anatis* ESV200 was obtained to carry out nucleotide sequencing of the genome. Two sequencing technologies, namely, the NextSeq500 platform from Illumina[®] and the Sequel System from PacBio[®], were used to obtain a unique chromosomal molecule. The software packages used to join reads into a single chromosomal sequence were FASTQC 11.9 (Wingett and Andrews 2018), Velvet 1.0.19 (Zerbino and Birney 2008), Spades 3.15.5 (Bankevich et al., 2012) and Canu 2.3 (Koren et al., 2017). The gene content in the genome of the ESV200 strain was explored by comparing the information deposited in GenBank under accession number CP114281 (WGS- Project-CP114281.1) and three other records for strains UMN179 (CP002667), F149 (JPHN01000000) and 12656-12 (AVOX00000000). The genomes included correspond to the most well-studied strains of *G. anatis*, and because the pathogenicity was compared with reference to the virulent strain 12656-12, UMN179 is the only completely circularized genome. Complete chromosomes were compared (Dhillon et al., 2013) with Rapid Annotation using Subsystem Technology (RAST) (Aziz et al., 2008; Overbeek et al., 2014) and displayed in a graphical representation prepared with the Easyfig program (Sullivan et al., 2011). Additionally, a review of the gene annotation obtained by the GenBank pipeline was performed. The contents of various RNA types were verified using Basic Local Alignment Search Tool - BLASTn (www.ncbi.nlm.nih.gov; Madden et al., 1996) and tRNAscan-SE 2.0 (Chan et al., 2021). The phylogenetic analysis of the four fimbrial operon genes was performed as described Kudirkiene et al. (2014), although the sequences of the genes encoding the Usher chaperone, the major protein, and the adhesin were separated for a more appropriate comparison. Clusters were constructed by comparison of *G. anatis* gene sequences from operons with Clustal Omega (Madeira et al., 2022): the sequences included here were from the *G. anatis* ESV200, UMN179, and 12656-12 strains. The phylogenetic analysis of *G. anatis* ESV200 was performed by comparing eight concatenated housekeeping genes (*adk*, *atpD*, *fumD*, *gyrB*, *inf2*, *mdh*, *recN*, and *tdh*) previously used as multilocus sequencing typing (MLST) markers in *G. anatis* (Johnson et al., 2013). The neighbor-joining method was used to draw the dendrogram using the online programs Clustal Omega

(www.ebi.ac.uk/Tools/msa/clustalo/), Kalign (Lassmann 2019) and iTol (Letunic and Bork 2021). A phylogram was constructed with or without a root of two MLST sequences from *Avibacterium endocarditis* 20186H4H1 and *A. gallinarum* NCTC11188 strains (Accession numbers: NZ_PQVI01000000 and NZ_PQVJ00000000.1, respectively). MEGA v.11 software was used to assess the confidence of the phylogram of *G. anatis* MLST sequences through a bootstrapping test with 500 resamples.

3. Results

3.1 *G. anatis* growth and biofilm production

Usually, *G. anatis* 12656-12 and ESV200 strains cultured in the same conditions show different turbidity or clumping phenotypes in liquid BHI medium, and 12656-12 strain cultures agglutinate after a few minutes if kept under static conditions; on the other hand, ESV200 strain cultures form stable dispersed suspensions, and this behavior could be a result of inoculum size, aeration and biofilm capacity. To analyze the development of turbidity and cell agglomeration, two types of cultures were prepared by incubation at 37°C, one agitated at 200 rpm to force the formation of a homogeneous suspension and the other unagitated. The shaken cultures consistently exhibited an OD₆₀₀ of 1.9 at 24 hours, while the static cultures showed a maximal OD₆₀₀ of 1.9 at 96 hours for *G. anatis* 12656-12 and a maximal OD₆₀₀ of 1.2 at 24 hours for *G. anatis* ESV200. Two conditions were established for comparative quantitation of gene expression in *G. anatis* at different culture times, as shown below. The results for the quantitation of biofilm formation are depicted in Figure 1, showing an increase in biofilm quantity for both strains at 48 vs. 24, 72, and 96 hours. As depicted in Figure 1A, strain ESV200 showed maximal production of biofilm at 48 h, which is 25% more than the amount of biofilm produced by the 12656-12 strain. However, the ESV200 biofilm was unstable, showing a drastic decrease of approximately 80% between 48 and 72 hours. The homogeneous suspension and the minor quantity of biofilm formed confirmed the predominant growth of ESV200 in liquid culture medium as a major planktonic suspension, with a temporary exception at 48 hours. In contrast, the 12656-12 strain was clumpy in liquid BHI medium and maintained the biofilm structure for a long time until the cell density was double that of ESV200 at 72 hours. These differences observed for the ESV200 strain cultures indicated that biofilm formation was turned on in a transient manner after disintegration of the biofilm structure but did not indicate an inability to produce a biofilm. The 12656-12 strain showed a robust biofilm.

Figure 1. Biofilm production and protein expression pattern from *G. anatis* 12656-12 and ESV200 strains cultured under different conditions and subjected to treatment.

A.- Biofilm production analysis as a function of culture incubation time (normalized); the optical density (OD_{630}) corresponds to the absorbance of the crystal violet dye retained by the biofilm. B.- Protein profiles from biofilms produced by *G. anatis* cultured at different temperatures. The arrows indicate the disappearance of protein bands at high temperature. WCE-whole cell extract. C.- Alteration of the enzyme-treated *G. anatis* biofilm produced in wells of polystyrene plates (as described in the Materials and methods); the optical density (OD_{630}) corresponds to the absorbance of the crystal violet dye retained by the biofilm after treatment. D.- Protein profiles of the *G. anatis* biofilms harvested and previously treated with DNase, lysozyme, or urea before separation by denaturing electrophoresis (SDS-PAGE) (as described in the Materials and methods). LWT-Left without treatment.

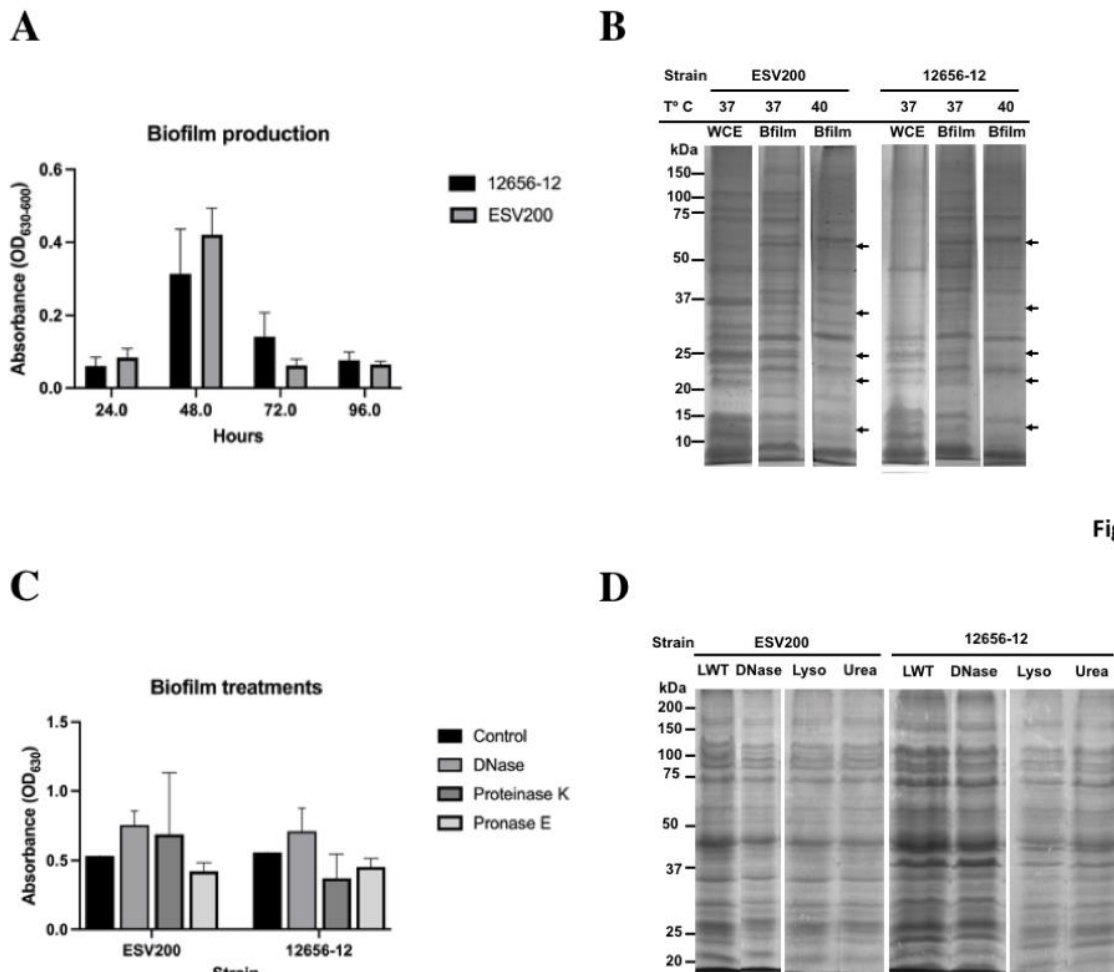


Figure 1

3.2 Protein production analysis of the *G. anatis* ESV200 and 12656-12 strains

The protein profiles of the planktonic and biofilm cultures were studied using SDS-PAGE to determine the importance of this type of cellular component in the formation of aggregates or in maintaining cellular dispersion. The protein patterns of bacterial growth of both strains under agitation showed higher levels of recoverable and visible proteins in the SDS-PAGE with differences in the levels of some proteins, which suggests that the proteins could be relevant in the adherence or dispersibility of both strains, phenotypes that differed between these two strains (Figure 1B). Therefore, it can be inferred that the ESV200 strain, because of its proteins, was more compatible with the aqueous medium because it remained dispersed for a longer time. Cultures of

the same strain with planktonic cells and biofilms also showed differences in the expression patterns of different proteins. The ESV200 strain showed five protein bands with increasing concentrations in the biofilm structure at 37°C. The concentrations of the visible proteins decreased at 42°C, indicating regulation mediated by temperature.

In *G. anatis* 12656-12, an increase in the levels of more than 6 accumulated protein bands in the adherent state was observed (the arrows show the 5 most evident bands of less than 100 kDa). This profile was maintained at 37 or 42°C of incubation; therefore, this variation in the culture conditions led to little alteration in 12656-12 strain metabolism.

As it has been reported that bacterial biofilms represent a state of cell agglomeration with a complex matrix, the biofilm was treated for 48 h with 4 different enzymes (two proteases, DNase and lysozyme) and urea (Figure 1C and 1D) to explore the changes in protein profile in the biofilm matrix. These treatments would increase the turbidity of the biofilm suspension if they separated the cell clumps. Figure 1C shows the change in cell agglomeration caused by DNase, proteinase K and pronase E. In samples from the ESV200 biofilm, DNase and proteinase K treatments increased the turbidity measured at 630 nm, while pronase E reduced the turbidity. For the 12656-12 biofilm, only treatment with DNase disintegrated the agglomerates and induced an increase in the OD; the proteases used did not have this effect. The protein profile was not altered by DNase, lysozyme or urea, although the separation of the proteins was slightly improved with respect to the profile of the proteins obtained by simple resuspension in the 2X loading buffer used for SDS-PAGE.

3.3 Pathogenicity test of *G. anatis* ESV200

Because the two strains of *G. anatis* investigated had different agglomeration behaviors, their pathogenic capacity was analyzed in groups of 5 birds at 8 weeks of age, and relatively similar results were obtained when the birds were inoculated intravenously with 0.2 ml of fresh culture. The symptoms appeared on the 4th day postinoculation, although there was greater decay in the group inoculated with ESV200 (Table 2). There was no evident damage to the respiratory system of birds inoculated with any of the strains during necropsy. However, a higher rate of damage was observed in the organs (liver, spleen or kidney) of the birds inoculated. Damage to the heart was observed in birds inoculated with the *G. anatis* 12656-12 strain but not in birds inoculated with the ESV200 strain. The injected bacteria were unable to be recovered from the most frequently damaged organs.

Table 2. *Gallibacterium anatis* pathogenesis test

Experimental infectious disease assay		Data assay		
<i>G. anatis</i> strain		Control (-)	12656-12*	ESV200**
Infection route		intravenous	intravenous	intravenous
Number of chickens by group		5	5	5
Chicken age (weeks old)		8	8	8
Symptoms appearance time (dpi)		none	3	3
External signs	Decayed (3dpi)	none	5/5	5/5
	Limp combs	none	2/5	5/5
Euthanasia and necropsy (dpi)		7	7	7
Pathological data observed				
	Respiratory tract damage	absent	absent	absent
	Liver congestion	none	3/5	3/5
	Liver atrophy	none	1/5	none
	Heart hydropericardium	none	1/5	none
	Hydropericardium onset	none	3/5	none
	Splenomegaly***	none	3/5	5/5
	Spleen atrophy	none	1/5	none
	Spleen congestion	none	3/5	3/5
	Kidneys	none	5/5	3/5
Bacterial recovery from organs				
	Spleen	Not tested	Negative	Negative
	Liver	Not tested	Negative	Negative

The potential pathogenicity of the *G. anatis* 12656-12 and ESV200 strains was separately tested in two groups of chickens plus a negative control group (only 0.2 mL of BHI injected/chicken). *Bojesen et al., 2003. Liver isolate from Denmark, **Edgardo Soriano, Tepatitlán Jal. México, ***Moderate splenomegaly with petechial hemorrhages. dpi.- days post-infection. Rational numbers represent the proportion of affected between unaffected animals of each group.

3.4 Fimbrial mRNA quantitation in *G. anatis* ESV200

As changes in bacterial aggregation have been observed through treatment with proteases and fibrillar structures by electron microscopy of *G. anatis* biofilms (Salgado-Lucio et al., 2012; Vaca et al., 2011), we asked how some extracellular proteins of *G. anatis* are expressed, as they could form cellular junctions within the biofilms or bridges of cells with solid surfaces, so the abundances of the mRNA of the fimbrial genes *flf1_a*, *flf3_b*, *flf_b*, *flf1_D*, *flf3_D* and *flf_D* (Figure 2A) were determined at the transcriptional level (Figure 2B). Absolute quantitation of the transcripts by qRT-PCR was performed by using standard curves prepared for each of the genes; the expression of each transcript was different, and the *flf3* gene primers did not amplify any fragment from ESV200 DNA, either due to variations in the target sequence of the primers or due to the absence of the gene in the genome of the ESV200 strain. Another gene that is poorly expressed in *G. anatis* ESV200 is *flf_D*. However, the structural and adhesin genes of the *flf_b* operon of ESV200 correspond to two different operons in *G. anatis* 12656-12; the homologous but more divergent adhesin (89% identity) is part of the *flf3_b* operon, and the remaining three genes, including the gene encoding the major fimbrial protein (98% identical), are present in *flf_b*. According to the expression analysis of the fimbrial genes in strain 12656-12, higher expression was induced under static conditions (Figure 3).

Figure 2. Gene organization in the F17-like operons from *G. anatis* and gene expression tests.

A.- Gene arrangement in the three F17-like operons harbored in *G. anatis* UMN179 and 12656-12 strains, and from the two operons found on the genome of *G. anatis* ESV200 strain; The molecular classification of the genes into operons was made according to Kudirkiene et al. (2014), and the name operons are shown for each one; the coordinates for operons within each genome is also shown. Capital letters A, B, C or D were added to distinguish between the genes encoding the major protein, the first chaperone, the Usher chaperone and the adhesin. In subscript the name of the strain was added. B.- Amplified DNA fragments obtained by the qRT-PCR test of the expressed fimbrial and adhesin genes from *G. anatis* 12656-12 and ESV200 strains; because the Flf3_b fimbrial gene (star labeled) homologous to the present in *G. anatis* 12656-12 strain does not exist in *G. anatis* ESV200 strain, as well as the Flf_D gene, spurious cDNA amplification in low quantity, was observed. +/- Positive or negative reaction controls. C.- Diagrams of the indicated genes in A are shown to scale, and the oligonucleotide DNA primers used as well as the expected size of the amplified DNA product are shown.

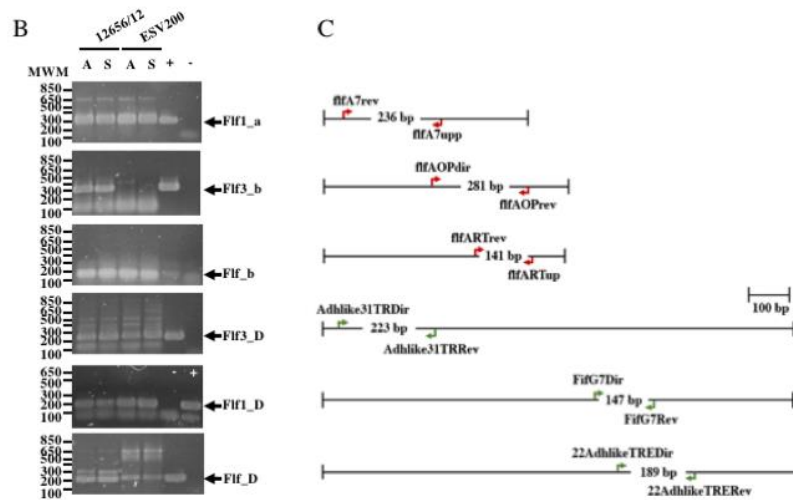
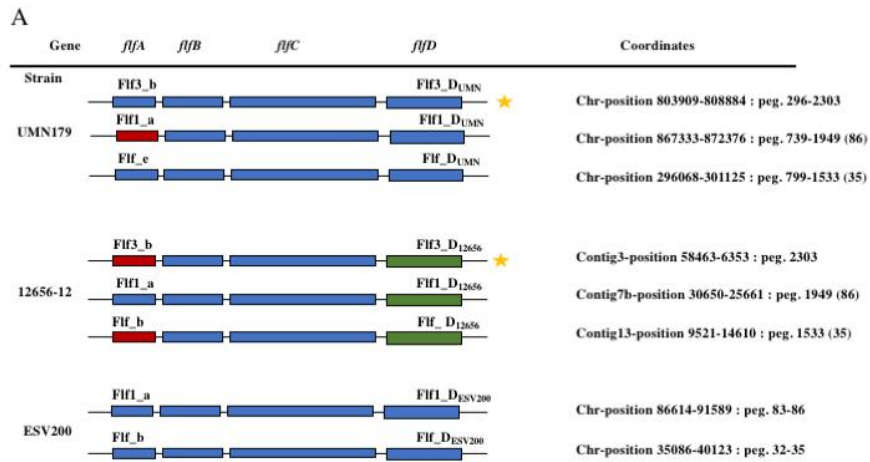
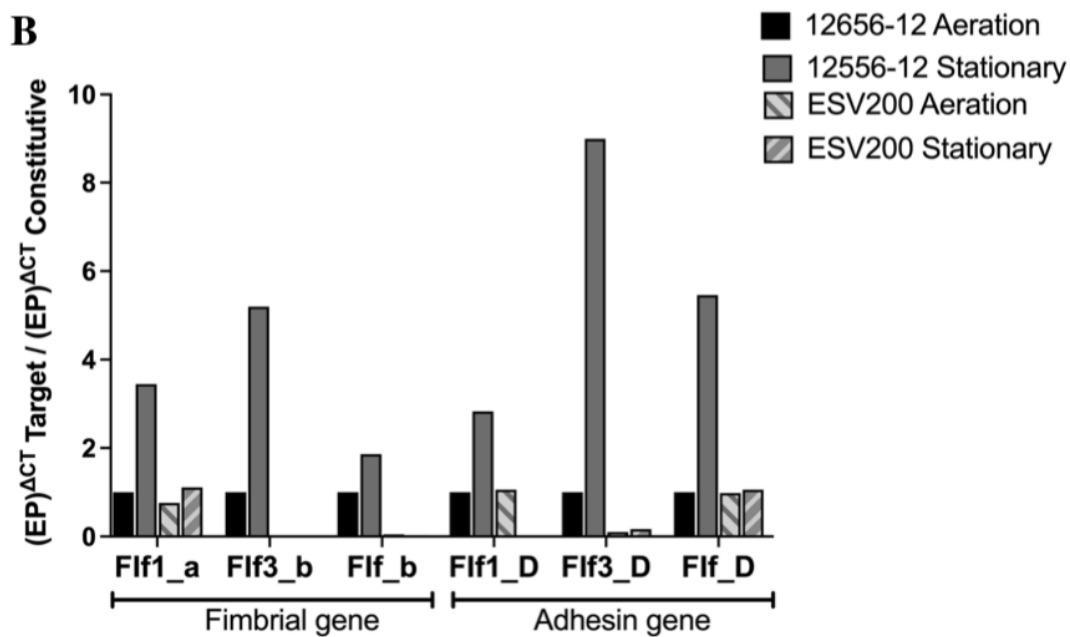
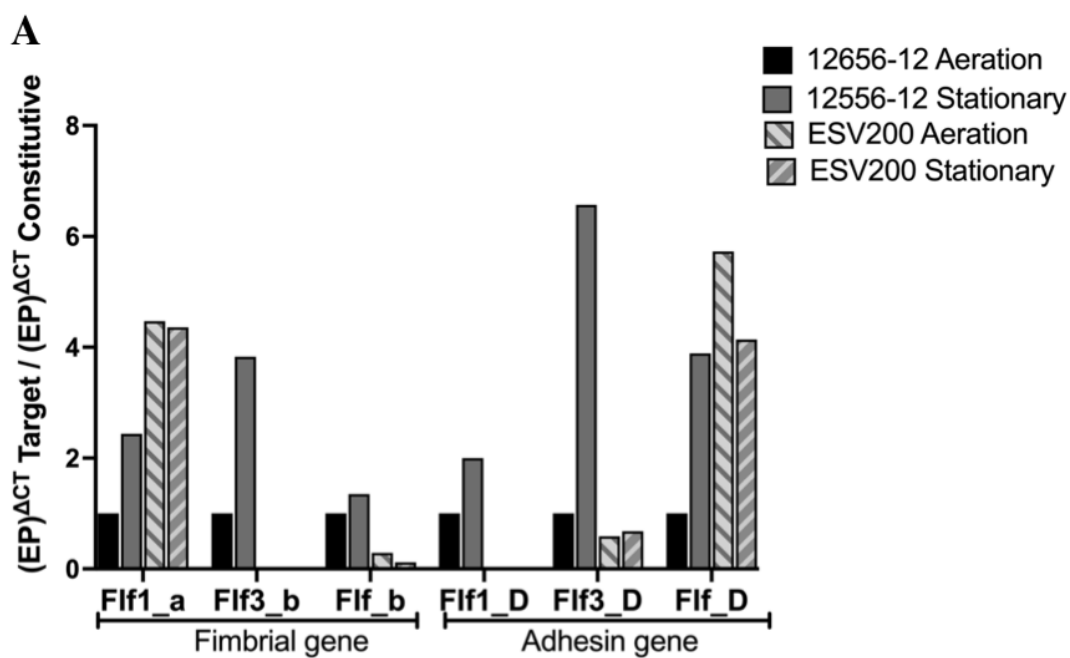


Figure 2



C

```

12656 Op1 MALIFHKILAIALRMAILPIRHYLVHVVNNSVIPILSLSVSNKVMVQLLLVTPILVPLPVLVHGVVVVVKVINKCLLGLLELL* 85
F149 Op2 MALIFHKILAIALRMAILPIRHYLVHVVNNSVIPILSLSVSNKVMVQLLLVTPILVLLPVLVHGVVVVVKVINKCLLVLELL* 85
UMN179Op1 MALIFHKILAIALRMAILPIRHYLVHVVNNSVIPILSLSVSNKVMVQLLLVTPILVLLPVLVHGVVVVVKVINKCLLVLELL* 85
*****

```

Figure 3

Figure 3. Inducible expression of F17-like genes in *G. anatis* 12656-12 strain and noninducible expression in *G. anatis* ESV200 strain.

A.- The gene expression quantitation of mRNA from F17-like operon genes was achieved by qRT-PCR. In this assay, the comparison with the constitutive *rpoA*-gene encoding the RNA-polymerase alpha subunit was included. B.- The gene expression quantitation of F17-like genes was performed by comparing the expression of the gene encoding elongation factor G. The efficiency of the qRT-PCRs was greater than 96%, with an $R^2 \geq 0.86$. C.- Amino acid sequence deduced from the open reading frames identified in the genome of several *G. anatis* strains but absent in the genome from *G. anatis* ESV200 strain; this protein was not identified in any database consulted, such as GenBank, ExpASy, PDB, or by searching with BLASTp.

The relative expression of fimbrial mRNAs with respect to the expression of mRNA of the RNA polymerase gene (Figure 3A) showed a constant expression level in the 12656-12 strain under static conditions, and the expression levels of the three mRNAs of the three fimbrial genes (0.5 at 4X) and the three mRNAs encoding adhesins (1 to 7X) changed under agitation conditions.

For the ESV200 strain, 4X greater expression of the mRNA encoding the Flf1_a protein was observed, regardless of the culture conditions, without the expression of Flf1_D. For the major protein Flf3_b, there was no expression in the ESV200 strain, and for adhesin, low expression was observed, perhaps as an important housekeeping gene. For the major protein Flf_b in ESV200, vestigial expression was observed, while increased expression (4X) of Flf was observed. In a comparative analysis of expression of the target genes with that of the housekeeping gene encoding EFG (Figure 3B), basal expression was also observed under aeration conditions in the 12656-12 strain, while the expression of the fimbrial proteins increased from 2X to 5X under static conditions; the expression of adhesins increased from 3X to 9X. In strain ESV200, there were no important changes in the expression of the majority fimbrial protein or adhesin, although under aeration, low expression of the adhesin Flf_D was observed (Figure 3).

3.5 The ESV200 strain genome contains fimbrial genes

To further understand the form of expression of the fimbriae in *G. anatis* ESV200, the sequence of the ESV200 genome was obtained, and information on the fimbrial operons was extracted. This

information was compared with that of the genomes of strains UMN179, 12656-12 and F149^T. The ESV200 strain genome was found to contain two fimbrial operons, while the other three strains possessed three operons (Figure 2). These operons are based on a secretion system that involves the Usher chaperone located on the outer membrane of the bacterium. This chaperone is the largest protein of the chaperone/Usher system that, together with another smaller periplasmic chaperone encoded by the fimbrial operon, is involved in the secretion of the fimbrial structure. The sequence of the usher chaperone gene of the ESV200 strain was compared with others (Figure 4), and it was found that, in addition to the open reading frame (ORF) for the usher gene, there was an internal reading frame in one of the usher genes in strains 12656-12, F149^T and UMN179 (Figure 3C), but this reading frame was absent in the ESV200 strain. The absence of this extra ORF may be related to the lower biofilm production by this strain.

3.5.1 ESV200 genome reconstruction and comparative genomic analysis

Initially, the genome of the ESV200 strain was assembled from 55 contigs containing 2.54 Mbp, which were obtained from reads from Illumina and 454 technologies after filtering for sequence quality with FastQC and applying Velvet software to reconstitute the genome into contigs. A circular genome was obtained with long reads from Sequel PacBio technology and assembled with Spades and Canu software. The ESV200 genome was reconstructed de novo with 32X coverage and N50 = 2538577. A full-length revision was made with RAST and deposited in the GenBank database with accession number CP114281.1 for final annotation with the NCBI Prokaryotic Genome Annotation Pipeline. The genome has a 40% GC content and contains 7, 6 and 6 copies of the 5S, 16S and 23S rRNA genes, respectively, and 61 tRNA copies. The Gtx-toxin operon was found to share 94% identity with the nucleotide sequence (HQ316167), and GtxA shared 100% identity with WP_094873868.1 and 99% identity with toxin proteins from the UMN179 and 12656-12 strains. Additionally, two chaperone/usher fimbrial operons were found (see above, Figures 2 and 4). There were 3 CRISPR arrays and 71 mobile element proteins present. To date, 39 genomes of this species have been reported, and only UMN179 has 1 circular genome with 39.9% GC content, 2.68 Mbp, and 7, 6, and 6 copies of the 5S, 16S and 23S rRNA genes, respectively, making this strain a high-confidence reference for exploring the core genes of the ESV200 strain. The genetic core of the 4 genomes consisted of 2,061 correctly annotated genes with known functions, and hypothetical genes were discarded. Table 3 shows the set of genes shared between genome-

strain pairs selected from GenBank: AVOX00000000.1, UGGZ00000000.1, CP114281.1 and CP002667.1. The comparison provides useful information about common genes and other strain-specific genes, for example, the genes for fimbrial attachment and colonization of host tissues mentioned above (Figure 4).

Table 3. *Gallibacterium anatis* ESV200 genome comparison

Strains*	12656-12	NCTC11413	ESV200	UMN179
12656-12	<u>2337</u> **	-	-	-
NCTC11413	2343	<u>2256</u>	-	-
ESV200	2341	2297	<u>2315</u>	-
UNM179	2359	2322	2326	<u>2324</u>

* Four GenBank genomes from *G. anatis* strains were compared by pairs; the selected genomes in this analysis were those containing the longest genome size or few contigs.

** Matrix values indicate the total number of homologous genes contained between strains. The values in the diagonal direction indicate the total gene number from each strain reported in the GenBank database. The table shows the number of genes shared between strains in a pairwise comparison. The hyphen symbol is included to avoid duplicating values in pairwise comparison. The number of probable genes shown is derived from the comparison between genomes using the RAST platform. Numerical differences are due to gene duplications or cross-homologous sequences containing multiple domains.

Figure 4. Molecular analysis of DNA and proteins encoded by *G. anatis* F17-like fimbrial operons.

4.1.-Phylogeny of F17-like genes. Clusters were constructed by gene comparison of *G. anatis* sequences from usher containing operons with Clustal Omega (Madeira et al., 2022): the sequences included here were from *G. anatis* ESV200, UMN179, and 12656-12 strains. A.- The phylogram of fimbrial genes from three *G. anatis* strains. B.- The cluster of usher-chaperone genes from three *G. anatis* strains. C.- The cluster of adhesin genes from *G. anatis*. The genes named here contain the GenBank database tags. Examples: fimbrial genes – Fim plus tag; usher genes – Ush plus tag; and adhesin genes -Ad plus tag. *flfA*- Reference *flfA* operon genes are indicated in the cladograms. The length of the branch is indicated in each sequence name.

4.2.- Phylogeny of F17-like fimbrial proteins from the *G. anatis* ESV200 strain. The phylogenetic trees were reduced leaving the *G. anatis* closely related sequences and eliminating duplicated records from GenBank: A.- Tree with sequences closely related to fimbrial protein ESV200-tag-32; B- Tree with sequences closely related to fimbrial protein ESV200-tag-83.

4.3.- Phylogeny of F17-like adhesin proteins from the *G. anatis* ESV200 strain. The phylogenetic trees were reduced leaving the *G. anatis* closely related sequences and eliminating duplicated records from GenBank: A.- Tree with sequences closely related to adhesin protein ESV200-tag-35; B- Tree with sequences closely related to adhesin protein ESV200-tag-86.

Dendrogram of *G. anatis* ESV200 usher related proteins

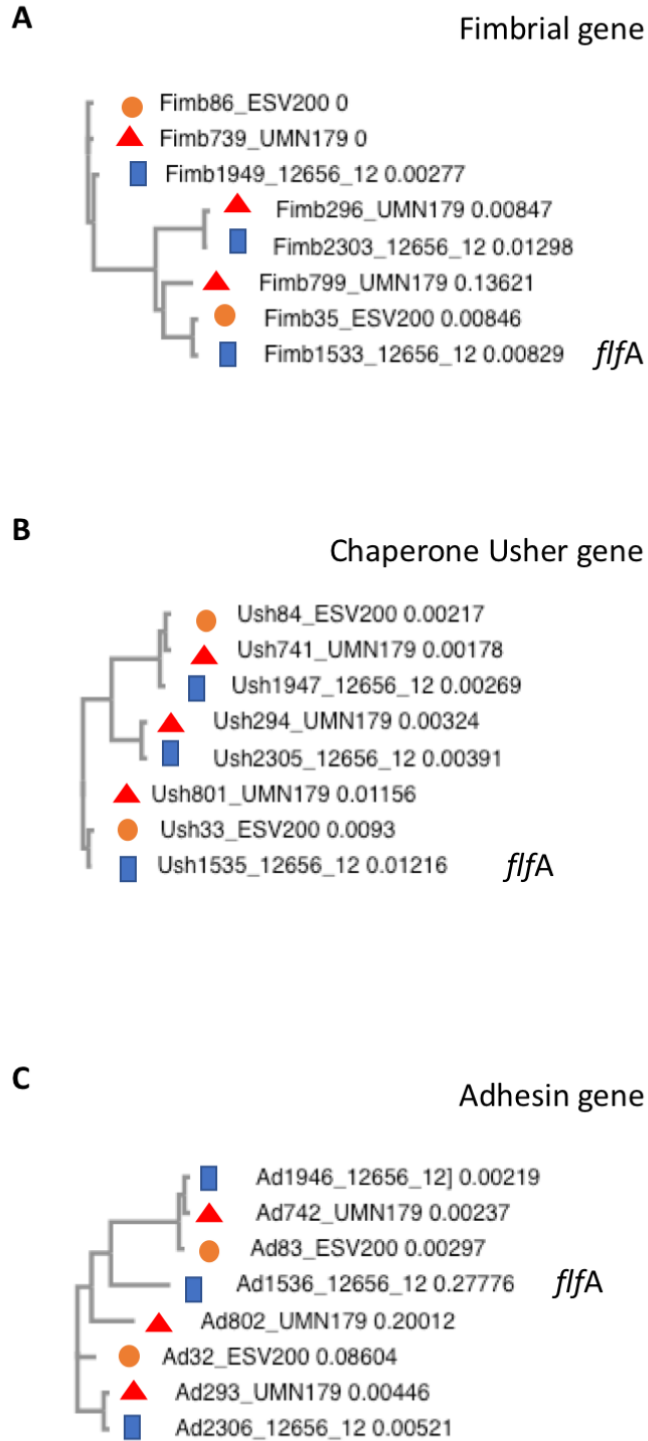


Figure 4

Phylogeny of *G. anatis* ESV200 fimbrial proteins

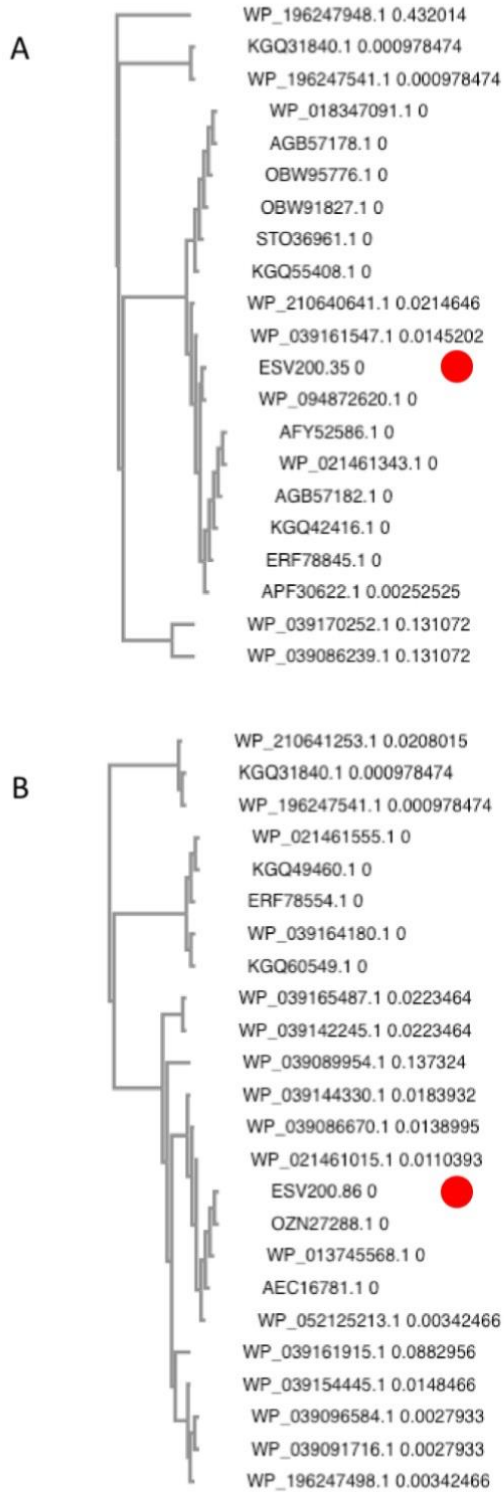


Figure 4.2

Phylogeny of *G. anatis* ESV200 adhesin proteins

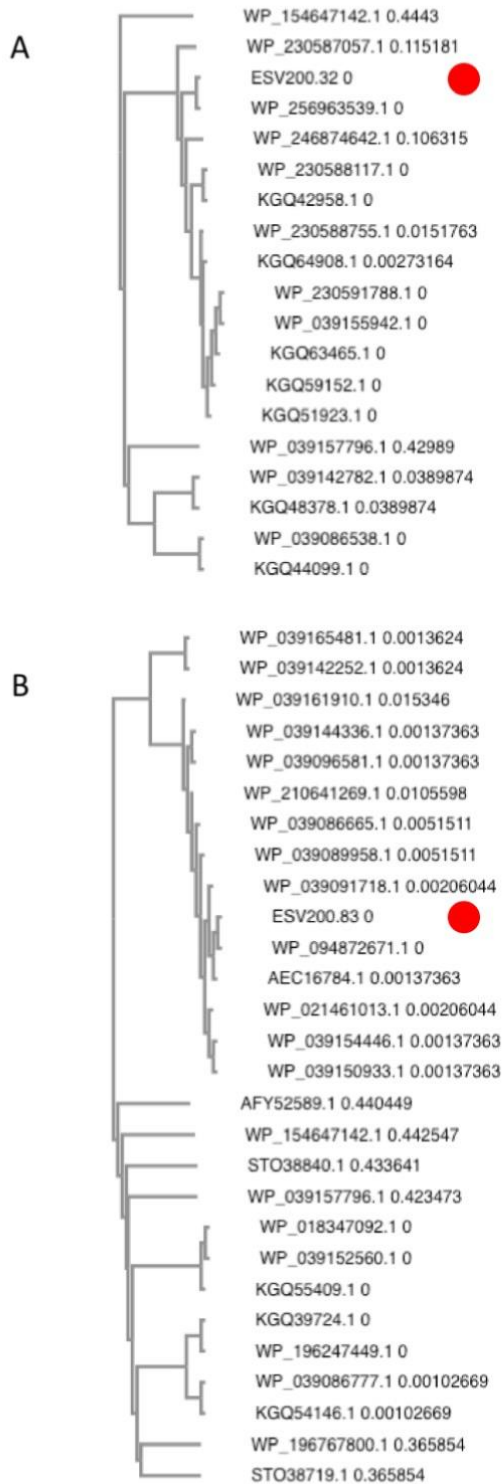


Figure 4.3

The linear comparison of the genomes of the strains ESV200 and UMN179 showed a significant number of chromosomal rearrangements but maintenance of the functionality responsible for pathogenicity in birds (Figure 5). The phylogenetic comparison of the strains that have had their genome sequenced to date, by using a group of genes suggested to be MLST markers, showed interesting groupings, although there was not much distinction among the sequenced strains, at least with this method (Figure 6). The evolutionary history was inferred using the maximum parsimony method. The most parsimonious tree with a length of 3105 is shown. The consistency index was 0.423833 (0.348744), the retention index was 0.509460 (0.509640) and the composite index was 0.215926 (0.177671) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm with search level 1 in which the initial trees were obtained by randomly adding sequences (10 replicates). This analysis involved 30 nucleotide sequences. A total of 12441 positions was included in the final dataset. The evolutionary analysis was conducted using MEGA 11 software.

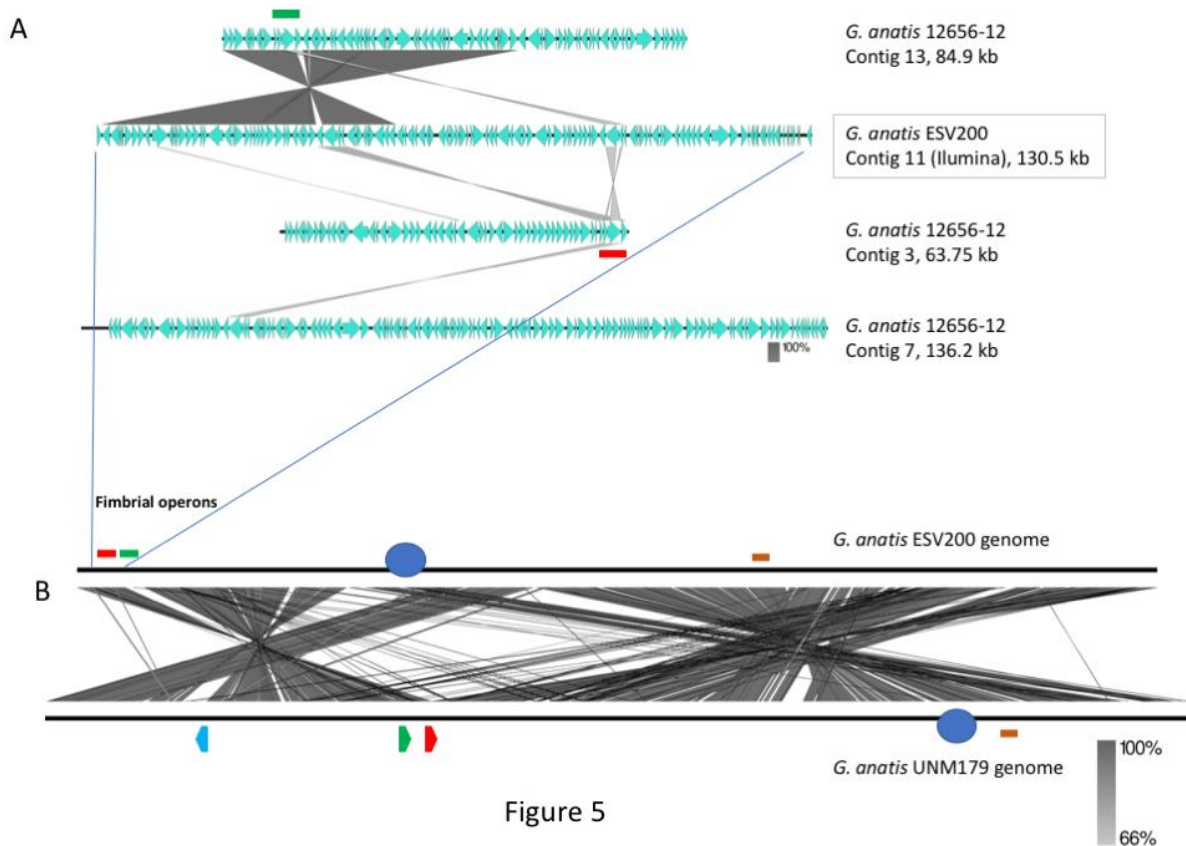


Figure 5

Figure 5. Identification and localization of F17-like fimbrial operons harbored in the *G. anatis* ESV200 genome.

A.- Identification and localization of F17-like fimbrial operons harbored in *G. anatis* ESV200 and 12656-12 strains. The *G. anatis* fimbrial operons containing the usher-chaperone gene were identified by BLASTn and BLASTx (www.ncbi.nlm.nih.gov; Madden et al., 1996) and Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/) programs. The image shows the identity regions between 130,591-base-long contig 11 from the *G. anatis* ESV200 strain and three contigs (No. 3, 7, and 13) from the *G. anatis* 12656-12 strain. Contig 11 from the *G. anatis* ESV200 strain has a large region with identity toward contig 13 from the *G. anatis* 12656-12 strain and harbors two fimbrial operons (highlighted with green and red colors). Each *G. anatis* 12656-12 contig harbor only one fimbrial operon.

B.- Full alignment of *G. anatis* ESV200 and UMN179 genomes. The ESV200 genome was reconstructed through de novo sequencing using Sequel PacBio technology with 32X coverage.

Two black horizontal lines show the comparison along the two *G. anatis* genomes, and the crossed lines between the two parallels indicate mainly stretches of homologous sequences. The crossed lines are the result of hundreds of direct and inverted matches, and here are observed two notorious regions for being inverted. The two sections are approximately equivalent to 30 and 60% of the compared genomes. Image obtained with Easyfig software.

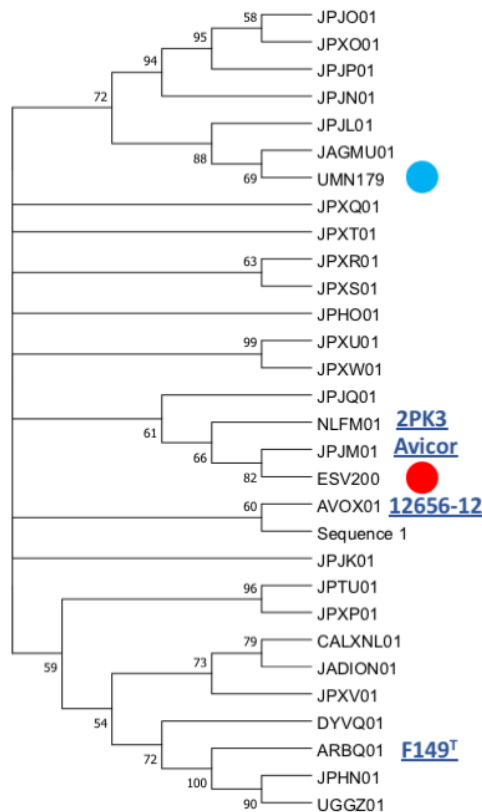


Figure 6. Phylogenetic grouping of the *G. anatis* ESV200 strain by MLST.

The analysis was performed using the next concatenated sequences harbored in thirty-one *G. anatis* genomes: *adk*, *atpD*, *fumD*, *gyrB*, *inf2*, *mdh*, *recN*, and *tdh*. This phylogram was obtained by Kalign (Lassmann 2019) and iTol (Letunic and Bork 2021) online programs. Here, *G. anatis* MLST sequences form a very compact group, although the *G. anatis* ESV200 sequences are closer to the

G. anatis 2PK3 and Avicor sequences. The names in the phylogeny tree are the accession number prefixes used in GenBank except ESV200 and UMN179 to easily identify them. Four names of the reference *G. anatis* strain are also highlighted on the outside of the tree. The evolutionary history was inferred using the maximum parsimony method. The most parsimonious tree with a length of 3105 is shown. The consistency index was 0.423833 (0.348744), the retention index was 0.509460 (0.509640) and the composite index was 0.215926 (0.177671) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm with search level 1 in which the initial trees were obtained by randomly adding sequences (10 replicates). This analysis involved 30 nucleotide sequences. A total of 12441 positions was included in the final dataset. The evolutionary analysis was conducted using MEGA v.11 software.

3.6 Integrases in the genome of *G. anatis* ESV200

Nine integrase-related amino acid sequences, ranging in size from 60 to 415 amino acids, were deduced *in silico* from the *G. anatis* ESV200 genome. A 168 amino acid sequence was annotated as an integrating regulator; therefore, the regulator size selection criterion was used to separate the integrase sequences into two categories. Thus, five large sequences remained as integrases (252 to 415 aa) and the four short ones as possible integration regulators (60 to 170 aa). The short sequences showed a conserved amino-terminal domain of 60 amino acids, although the integrase regulator had more significant differences from the other short proteins. Essential amino acids that are part of the active site for integration, including tyrosine, were identified in proteins from 252 to 415 aa. Three amino acid sequences related to integrases delimited a 539.6 kb fragment. This sequence is flanked at the ends by two segments of mobile genetic elements, at one end an episome type and at the other a prophage element, with 50.5 and 46.3 kb, respectively. The remarkable aspect of the chromosome fragment is that it contains chromosome, phage, and plasmid replication gene clusters. Mobilization genes and plasmid transfer. Genes that can provide resistance to antimicrobial substances as ejection pumps. Genes for RNA polymerase peptide biosynthesis. Genes involved in ribosome biosynthesis, such as ribosomal RNA and ribosomal proteins. Genes for the biosynthesis of aminoacyl-tRNA synthetases. Genes for the biosynthesis of chaperones for the export of extracellular proteins. Genes for fimbriae biosynthesis and control of biofilm formation. By size, this segment resembles a bacterial minichromosome. However, due to the lack of essential genes, it has an integron-type structure, similar to those called Maverick. This chromosome segment is equivalent to a domain called "D1", composed of 22% of all genes. The

"D2 and D3" domains, delimited by the mobile element regions, contain 25.2 and 52.8% of all genes, respectively. The D1 domain contains 5% hypothetical and DNA mobility genes. Regarding the exclusive DNA mobility genes, D1 contained 3.1%, D2 6.1%, and D3 0.94%. This means more significant potential for DNA mobility in the small domains than in the large domain of the *G. anatis* ESV200 chromosome.

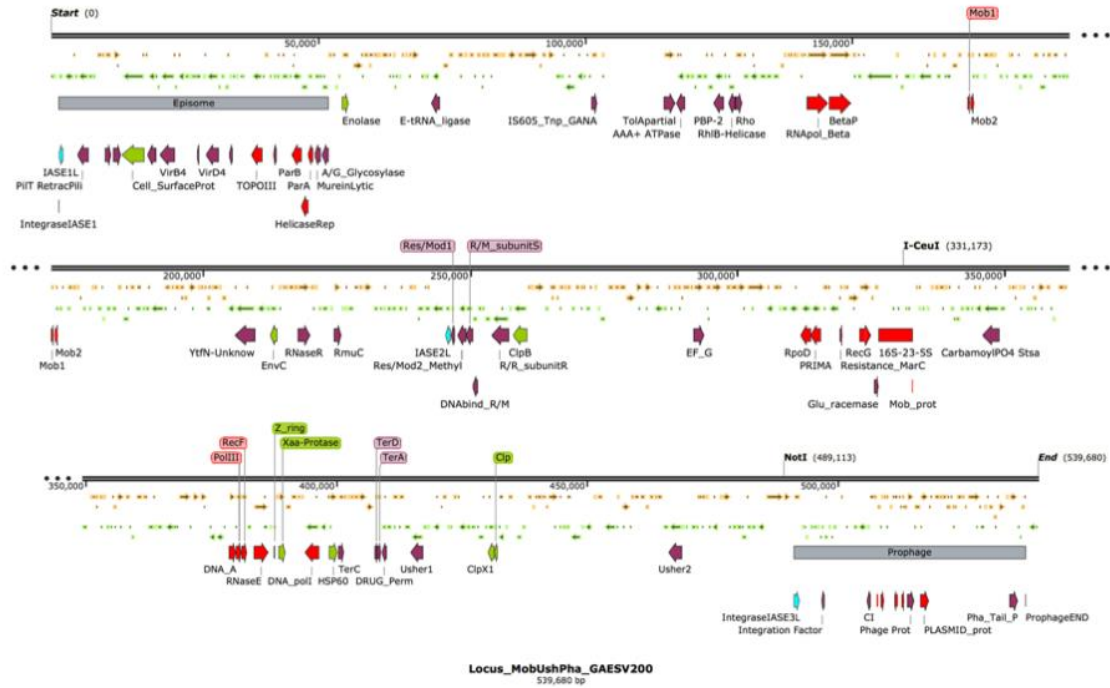


Figure 7. Organization of the chromosomal domain D1 of *G. anatis* ESV 200 displaying a complex Maverick-like structure.

Domain D1 of the three chromosomal domains of *G. anatis* ESV200 is shown here. The domains are separated based on the presence of large mobile genetic elements, the D1 domain is composed of 539.6 kb, and 520 ORFs have been deduced. The D1 domain contains an episome

and a prophage at the ends, and there are also open reading frames encoding putative integrases. Some sequences that code for proteins related to DNA mobility, integral DNA maintenance, recombination, host colonization, participation in the control of biofilm formation, drug resistance and cell division have been highlighted.

4. Discussion

Several diseases in birds are caused by *G. anatis* and are generically known as gallibacteriosis. Biofilm formation requires the establishment of a colony on superficial tissues and, if conditions allow it, on internal surfaces or internal organs; for this, the interaction of bacterial adhesins with host receptors is essential. These colonization processes can be very beneficial for the bacterium under antimicrobial treatment, as bacteria exhibit resistance in biofilms, which can be formed by several pathogens, including those of the *Gallibacterium* genus. It has also been reported that fimbriae are important for the formation of biofilms and that they are even targets of vaccination therapy via the development of fimbrial protein antigens. Here, we examined the presence of genes that encode fimbrial proteins and their transcriptional expression in the ESV200 strain, a pathogenic strain that apparently produces little biofilm in vitro.

In this work, the biofilm formation phenotype of strain ESV200 was compared with that of strain 12656-12, a bacterium included in several previous reports. The ESV200 strain forms highly homogeneous suspensions in liquid cultures, while the 12656-12 strain forms clumpy cultures that can settle rapidly when left under static conditions for a short time. However, biofilm quantitation performed on cell culture plates showed that the ESV200 strain produced as much biofilm as the 12656-12 strain at 24 and 48 hours, but this production was reduced after 72 hours. This behavior could be attributed to the change in the expression of genetic regulators that are still unknown under laboratory conditions or to loss of viability. The strain 12656-12 showed maintenance of the biofilm structure over time. Under such conditions, it would be expected that there would be differences in virulence if the aggregation conditions and biofilm formation were highly relevant in the pathogenic process due to better adhesion to host tissues (Salgado-Lucio et al., 2012).

The pathogenicity tests showed a similar pattern of damage produced by the two strains tested, with heart damage caused by strain 12656-12 but not by ESV200. We could not confirm that the lower quantity of biofilm produced at 72 hours by the ESV200 strain was the cause of the lack of heart damage, since molecular components of the genetic background that have not been evaluated may

have played a role; for example, the *in silico* analysis indicated that the biosynthesis of cardiolipins that are present in both strains might indicate the surface compatibility of the bacterium with cardiac tissue (Rossi et al., 2017). The genomes of the ESV200 and 12656-12 strains contain two genes homologous to those involved in the pathogenicity of *Salmonella enterica*, *pbgA* and *icsA* (Chiu et al., 2005; Cian et al., 2019).

The biofilm produced by *G. anatis* was altered by the biofilm-disrupting treatment conditions (Figure 1C). DNase and proteinase K altered the ESV200 strain biofilm, as indicated by the increase in turbidity, indicating the participation of DNA and proteins in biofilm structure formation and cell attachment. Disintegration was also observed for the 12656-12 strain biofilm treated with DNase. The reduction in turbidity may also have been due to extensive activity of the pronase E and proteinase K enzymes; therefore, the presence of proteins in the biofilm of *G. anatis* was explored. The proteins were revealed by electrophoretic separation of the total extracts of both strains and analyzed, and the sizes ranged from > 10 to <75 kDa; the sizes of fimbrial proteins ranged from 22 to 43 kDa. Lysozyme and urea further improved the protein resolution in the SDS-PAGE system, which may be because some proteins became free in solution and by inhibition of proteolytic activity (Tamura et al., 1998; Garcia-Gomez et al., 2005).

Regarding the possibility of the biosynthesis of fimbriae of the F17 type proposed by Bojesen, two fimbrial operons were identified in the genome of the ESV200 strain, which corresponded to the chaperone/usher secretion system, one of which showed 99% conservation with the gene present in the 12656-12 strain, while the other corresponded to a secretion system, exhibiting 98, 92 and 97% identity with the major fimbrial protein, the chaperone and Usher, respectively, whereas the putative adhesin exhibited 89% identity with the third operon of the 12656-12 strain, which suggests that an important recombination process led to the deletion of an operon in the ESV200 strain and the exchange of genes as modules within the fimbrial operons, and perhaps in this way, it gains greater genetic and functional variability, as occurs with other microorganisms (Dufresne et al., 2018). This conclusion was made by sequencing the genome of strain ESV200, which made it possible to clarify why not all the transcripts present in strain 12656-12, evaluated at the beginning of the investigation, could be quantified. Investigation of the identities of the adhesins showed that the adhesin FlfA shared 99% identity with that present in the 12656-12 strain, 30% with the *E. coli* adhesin TPA (HAN7666271.1) and 29% with the *S. enterica* protein (EDB7133990.1), while the adhesin FlfA7 shared 88% identity with the protein from 12656-12, 36% with the *E. coli* adhesin F17 (STJ01015.1) and 32% with F17G (EFA5996531.1). The adhesin absent in the ESV200 strain

but present in the 12656-12 strain genome shared 33% identity with the *E. coli* adhesin TPA (HAN7666271.1) and 31.6% identity with the *S. enterica* adhesin (EAM2645026.1).

The presence of several F17-type operons varies among *G. anatis* strains, since more than one fimbrial operon may be present and since some of them have mutations that change them into pseudogenes, so it is expected that the content and type of fimbriae will depend on the type of strain tested (Johnson et al., 2013; Pors et al., 2016). In the ESV200 strain, the two operons seem to express themselves with different transcription rates according to the estimates made with two housekeeping genes used in this work, namely, RNA polymerase and elongation factor G. Interestingly, it was observed that the ESV200 Flf_a mRNA was strongly expressed under static conditions and agitation, similar to the adhesin Flf. For the mRNAs of the fimbrial proteins and adhesins of the 12656-12 strain, notable increases in expression levels of 1x to 6x more than the transcription of the RNA polymerase subunit were observed for all the operons in the static condition, perhaps due to the distribution of molecular regulators. Therefore, the immunity strategies based on the fimbrial proteins proposed by Bojesen are functional in strains that exhibit the expression behavior seen in strain 12656-12 or in strains that have different types of fimbrial operons, which seem to be conserved in *G. anatis* for their importance in the colonization of avian tissues (Bojesen et al., 2004; Salgado-Lucio et al., 2012; Pors et al., 2016).

It is currently important to obtain a new fully sequenced and circularized *G. anatis* genome to accurately compare the possible presence or absence of genes that are frequently detected as multiple copies, whether they are involved in basic functions such as those of the biosynthetic machinery or other functions related to virulence, such as colonization.

Location of the fimbrial operons in the genome of *G. anatis* ESV200. It has been documented that bacteria can fix their genome to perform better in the environment where they live, be it natural or laboratory (Nagy-Staron et al., 2021; Saleski et al., 2021). It has been proposed that genes in the genome have different levels of expression depending on their location and their syntenic organization, which in many bacteria has allowed the correlation of gene clusters with the virulence abilities of pathogenic bacteria (Le Berre et al., 2022). One of the outstanding aspects at the moment in the genetics of bacteria is the presence of islands of pathogenicity (Desvaux et al., 2020), which are short segments of the genetic material that concentrate genes that encode virulence factors (Wu, S. 2015), antibiotic resistance functions (Debroas, et al 2019; Ellabaan, et al 2021; Maciel-Guerra, A., 2022), proteins that ensure DNA perpetuation, and enzymes involved in the internal mobility of DNA (Hallstrom et al., 2015; Smyshlyayev et al., 2021). Due to their functional

structure, they can be contained in plasmids, transposons, integrons, or phages (Dobrindt et al., 2015). In the genome of bacteria of the *Pasteurellaceae* family, the presence of a significant number of sequences related to DNA mobility, such as integrases or transposases, has been observed (Rodt et al., 2012; Horta et al., 2017). In the genome of *G. anatis* ESV200, together with the Usher operons that encode the fimbriae, an insertion sequence was found that could represent an integron, similar to pathogenicity islands (Carniel et al., 1996; Mazel D. 2006). When inspecting toward the ends of the operons, two insertion sequences were located, one next to 50.5 kb of an episome and another next to 46.3 kb of a prophage. The segment or D1-domain delimited by both mobile elements is approximately 539.6 kb in size (Figure 7). The inspection of this segment of DNA concentrates most of the vital functions for the cell, DNA Primase, DNA polymerase, components of the main RNA polymerase, various recombinases, antitoxin-toxin system, DNA translocation, transfer of genetic material, production of fimbriae, general mechanisms of resistance to antibiotics, genes responsible for the synthesis of ribosomal RNA 16-23-5S and control of biofilm formation. This large sequence resembles the mobile structures of eukaryotes called Maverick, associated with c-integrase, and in size, they almost reach the size of a *Mycoplasma* genome (Suthers et a., 2009).

5. Conclusions

The genome of *G. anatis* strain ESV200 is the second genome of this species to be completely sequenced since 2011, with an estimated genome size of 2,538,577 nucleotides. With a group of classic MLST genes, a phylogenomic association can be achieved, although with another group of genes, other close evolutionary relationships between *G. anatis* strains can be observed, which is always interesting when characterizing new isolates to differentiate them from some clonal types.

The ESV200 strain is virulent in birds and produces fimbriae related to colonization and virulence; it contains several virulence genes, one of which encodes the toxin GtxA, the main virulence factor of *G. anatis*. Shown here is the structure of a chromosomal domain, which is armed in a way recently described, named Maverick, and which can impact the evolution of pathogens, such as *G. anatis*.

Strain ESV200 can be used as a molecular model for gene knockout studies as well as for genetic complementation because it naturally contains only two Usher-type fimbrial operons and exhibits a weaker clumping phenotype in biofilms.

Acknowledgments

We are grateful to CONACYT México for the fellowship for Cobos Justo ME, to CONACYT México for project CB-2015-01-259209, BUAP-VIEP-00265, BUAP-VIEP-100103133-2022, UNAM-DGAPA PAPIIT IN204122 and to Edgardo Soriano for the kind donation of the bacterial ESV200 strain used in this study.

References

1. Alispahic, M., Christensen, H., Hess, C., Razzazi-Fazeli, E., Bisgaard, M., Hess, M., 2011. Identification of *Gallibacterium* species by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry evaluated by multilocus sequence analysis. *Int J Med Microbiol.* 301(6):513-22. doi: 10.1016/j.ijmm.2011.03.001.
2. Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., et al., 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics.* 9:75. doi: 10.1186/1471-2164-9-75.
3. Bager, R.J., Persson, G., Nesta, B., Soriani, M., Serino, L., Jeppsson, M., Nielsen, T.K., Bojesen, A.M., 2013a. Outer membrane vesicles reflect environmental cues in *Gallibacterium anatis*. *Vet Microbiol.* 167(3-4):565-72. doi: 10.1016/j.vetmic.2013.09.005.
4. Bager, R.J., Nesta, B., Pors, S.E., Soriani, M., Serino, L., Boyce, J.D., Adler, B., Bojesen, A.M., 2013b. The fimbrial protein FlfA from *Gallibacterium anatis* is a virulence factor and vaccine candidate. *Infect Immun.* 81(6):1964-73. doi: 10.1128/IAI.00059-13.

5. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455-77. doi: 10.1089/cmb.2012.0021.
6. Bertelli, C., et al., 2017. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research.* May 2. doi: 10.1093/nar/gkx343
7. Bisgaard, M., 1977. Incidence of *Pasteurella haemolytica* in the respiratory tract of apparently healthy chickens and chickens with infectious bronchitis. Characterization of 213 strains. *Avian Pathol.* 6:285–292. doi: 10.1080/03079457708418238.
8. Bojesen, A.M., Nielsen, O.L., Christensen, J.P., Bisgaard, M., 2004. In vivo studies of *Gallibacterium anatis* infection in chickens, *Avian Pathology*, 33:2, 145-152, DOI: 10.1080/03079450310001652059
9. Bojesen, A.M, Shivaprasad, H.L., 2007a. Genetic diversity of *Gallibacterium* isolates from California turkeys, *Avian Pathology*, 36:3, 227-230. DOI: 10.1080/03079450701332352
10. Bojesen, A.M., Torpdahl, M., Christensen, H., Olsen, J.E., Bisgaard, M., 2003. Genetic diversity of *Gallibacterium anatis* isolates from different chicken flocks. *J Clin Microbiol* 41: 2737-2740. doi: 10.1128/JCM.41.6.2737-2740.2003.
11. Bojesen, A.M., Vazquez, M.E., Robles, F., Gonzalez, C., Soriano, E.V., Olsen, J.E., Christensen, H., 2007b. Specific identification of *Gallibacterium* by a PCR using primers targeting the 16S rRNA and 23S rRNA genes. *Vet Microbiol.* 123(1-3):262-8. doi: 10.1016/j.vetmic.2007.02.013.

12. Carniel, E., Guilvout, I., Prentice, M. 1996. Characterization of a large chromosomal "high-pathogenicity island" in biotype 1B *Yersinia enterocolitica*. *Journal of Bacteriology*, 178(23), 6743–6751. <https://doi.org/10.1128/jb.178.23.6743-6751.1996>
13. Cian, M.B., Giordano, N.P., Masilamani, R., Minor, K.E., Dalebroux, Z.D., 2019. *Salmonella enterica* Serovar Typhimurium Uses PbgA/YejM To Regulate Lipopolysaccharide Assembly during Bacteremia. *Infect Immun*. 2019 Dec 17;88(1):e00758-19. doi: 10.1128/IAI.00758-19.
14. Chan, P.P., Lin, B.Y., Mak, A.J., Lowe, T.M. 2021. tRNAscan-SE 2.0: Improved Detection and Functional Classification of Transfer RNA Genes. *Nucl. Acids Res*. 49:9077–9096
15. Chantes-Guerra, A., Maldonado-Puga, S., Rojas-Ruiz, N., Rea-Hernandez, I., Montes-Garcia, F.J., Trujillo-Ruiz, H., Yañez-Aguilar, I.E., Vazquez-Cruz, C., Sanchez-Alonso, P., Negrete-Abascal, E., 2022. A putative siderophore receptor of *Gallibacterium anatis* 12656-12 under Fur control also binds hemoglobin. *Frontiers in microbiology*, 13, 951173. <https://doi.org/10.3389/fmicb.2022.951173>
16. Chiu, C.H., Tang, P., Chu, C., Hu, S., Bao, Q., Yu, J., Chou, Y.Y., Wang, H.S., Lee, Y.S. 2005. The genome sequence of *Salmonella enterica* serovar Choleraesuis, a highly invasive and resistant zoonotic pathogen. *Nucleic Acids Res*. 33(5):1690-8. doi: 10.1093/nar/gki297. Erratum in: *Nucleic Acids Res*. 2005;33(7):2351.
17. Christensen, H., Bisgaard, M., Bojesen, A.M., Muttters, R., Olsen, J.E., 2003. Genetic relationships among avian isolates classified as *Pasteurella haemolytica*, *Actinobacillus salpingitidis* or *Pasteurella anatis* with proposal of *Gallibacterium anatis* gen. nov., comb. Nov. and description of additional genomospecies within *Gallibacterium* gen. nov. *Int J Syst Evol Microbiol*, 53, 275-287. doi: 10.1099/ijms.0.02330-0.
18. de Oliveira-Garcia, D., Dall'Angol, M., Rosales, M., Azzuz, A.C.G.S., Alcantara, N., Martinez, M.B., Giron, J.A., 2003. Fimbriae and adherence of *Stenotrophomonas maltophilia* to epithelial cells and to abiotic surfaces. *Cell. Microbiol*, 5, 625-636. doi: 10.1046/j.1462-5822.2003.00306.x.

19. Debroas, D., Siguret, C., 2019. Viruses as key reservoirs of antibiotic resistance genes in the environment. *The ISME journal*, 13(11), 2856–2867. <https://doi.org/10.1038/s41396-019-0478-9>
20. Desvaux, M., Dalmasso, G., Beyrouthy, R., Barnich, N., Delmas, J., Bonnet, R., 2020. Pathogenicity Factors of Genomic Islands in Intestinal and Extraintestinal *Escherichia coli*. *Frontiers in microbiology*, 11, 2065. <https://doi.org/10.3389/fmicb.2020.02065>
21. Dhillon, B.K., Chiu, T.A., Laird, M.R., Langille, M.G., Brinkman, F.S., 2013. IslandViewer update: Improved genomic island discovery and visualization. *Nucleic Acids Res.* 41(Web Server issue):W129-32. doi: 10.1093/nar/gkt394.
22. Di Martino P., 2016. Bap: A New Type of Functional Amyloid. *Trends Microbiol.* 24(9):682-684. doi: 10.1016/j.tim.2016.07.004.
23. Dobrindt, U., Tjaden, S., Shah, S., Hacker., J., 2015. Mobile genetic elements and pathogenicity islands encoding bacterial toxins. Chp 2, Pag 40-76. Editor(s): Joseph Alouf, Daniel Ladant, Michel R. Popoff. *The Comprehensive Sourcebook of Bacterial Protein Toxins (Fourth Edition)*, ISBN 9780128001882. Academic Press. <https://doi.org/10.1016/B978-0-12-800188-2.00002-1>.
24. Donlan, R.M., 2002. Biofilms: microbial life on surfaces. *Emerg Infect Dis.* 8 (9), 881-890. doi: 10.3201/eid0809.020063.
25. Dufresne, K., Saulnier-Bellemare, J., Daigle, F., 2017. Functional Analysis of the Chaperone-Usher Fimbrial Gene Clusters of *Salmonella enterica* serovar Typhi. *Frontiers in Cellular and Infection Microbiology*. <https://doi.org/10.3389/fcimb.2018.00026>
26. Ellabaan, M., Munck, C., Porse, A., Imamovic, L., Sommer, M., 2021. Forecasting the dissemination of antibiotic resistance genes across bacterial genomes. *Nature communications*, 12(1), 2435. <https://doi.org/10.1038/s41467-021-22757-1>
27. Garcia-Gomez, E., Vaca, S., Perez-Mendez, A., Ibarra-Caballero, J., Perez-Merquez, V., Tenorio, V. R., Negrete-Abascal, E., 2005. *Gallibacterium anatis*-secreted metalloproteases degrade chicken IgG. *Avian pathology: journal of the W.V.P.A.*, 34(5), 426–429. <https://doi.org/10.1080/03079450500267866>

28. Hallstrom K.N., McCormick B. A., 2015. Pathogenicity Islands: Origins, Structure, and Roles in Bacterial Pathogenesis, Chp 16, Pag 303-314. Editor(s): Yi-Wei Tang, Max Sussman, Dongyou Liu, Ian Poxton, Joseph Schwartzman. Molecular Medical Microbiology (Second Edition), ISBN 9780123971692. Academic Press. <https://doi.org/10.1016/B978-0-12-397169-2.00016-0>.
29. Horta-Valerdi, G., Sanchez-Alonso, M.P., Perez-Marquez, V.M., Negrete-Abascal, E., Vaca-Pacheco, S., Hernandez-Gonzalez, I., Gomez-Lunar Z., Olmedo-Alvarez G., Vázquez-Cruz, C., 2017. The genome sequence of *Avibacterium paragallinarum* strain CL has a large repertoire of insertion sequence elements. Genome Announcements, 5(15), e00152-17. <https://doi.org/10.1128/genomeA.00152-17>
30. Huangfu, H., Zhao, J., Yang, X., Chen, L., Chang, H., Wang, X., Li, Q., Yao, H., Wang, C., 2012. Development and preliminary application of a quantitative PCR assay for detecting *gtxA*-containing *Gallibacterium* species in chickens. Avian Dis. 56(2):315-20. doi: 10.1637/9907-082511-Reg.1.
31. Johnson, T.J., Danzeisen, J.L., Trampel, D., Nolan, L.K., Seemann, T., Bager, R.J., Bojesen, A.M., 2013. Genome analysis and phylogenetic relatedness of *Gallibacterium anatis* strains from poultry. PLoS One. 8(1):e54844. doi: 10.1371/journal.pone.0054844.
32. Johnson, T.J., Fernandez-Alarcon, C., Bojesen, A.M., Nolan, L.K., Trampel, D.W., Seemann, T., 2011. Complete genome sequence of *Gallibacterium anatis* strain UMN179, isolated from a laying hen with peritonitis. Journal of bacteriology, 193(14), 3676–3677. <https://doi.org/10.1128/JB.05177-11>
33. Kristensen, B.M., Frees, D., Bojesen, A.M., 2011. Expression and secretion of the RTX-toxin GtxA among members of the genus *Gallibacterium*. Vet Microbiol. 153(1-2):116-23. doi: 10.1016/j.vetmic.2011.05.019.

34. Kudirkiene, E., Bager, R.J., Johnson, T.J. et al., 2014. Chaperone-usher fimbriae in a diverse selection of *Gallibacterium* genomes. *BMC Genomics* 15, 1093 <https://doi.org/10.1186/1471-2164-15-1093>
35. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., Phillippy, A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27(5):722-736. doi: 10.1101/gr.215087.116.
36. Lassmann, T., 2019. Kalign 3: multiple sequence alignment of large data sets. *Bioinformatics*, 36, 1928–1929, <https://doi.org/10.1093/bioinformatics/btz795>.
37. Le Berre, D., Reverchon, S., Muskhelishvili, G., Nasser, W., 2022. Relationship between the Chromosome Structural Dynamics and Gene Expression-A Chicken and Egg Dilemma?. *Microorganisms*, 10(5), 846. <https://doi.org/10.3390/microorganisms10050846>
38. Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation, *Nucleic Acids Research*, Volume 49, Issue W1, 2 July 2021, Pages W293–W296, <https://doi.org/10.1093/nar/gkab301>
39. Lopez-Ochoa, J., Montes-Garcia, J.F., Vazquez, C., Sanchez-Alonso, P., Perez-Marquez, V.M., Blackall, P.J., Vaca, S., Negrete-Abascal, E., 2017. *Gallibacterium* elongation factor-Tu possesses amyloid-like protein characteristics, participates in cell adhesion, and is present in biofilms. *Journal of microbiology (Seoul, Korea)*, 55(9), 745–752. <https://doi.org/10.1007/s12275-017-7077-0>
40. Maciel-Guerra, A., Baker, M., Hu, Y., Wang, W., Zhang, X., Rong, J., Zhang, Y., Zhang, J., Kaler, J., Renney, D., Loose, M., Emes, R.D., Liu, L., Chen, J., Peng, Z., Li, F., Dottorini, T., 2022. Dissecting microbial communities and resistomes for interconnected humans, soil, and livestock. *The ISME journal*, 10.1038/s41396-022-01315-7. Advance online publication. <https://doi.org/10.1038/s41396-022-01315-7>
41. Madden, T.L., Tatusov, R.L., Zhang, J., 1996. Applications of network BLAST server. *Methods in enzymology*, 266, 131–141. [https://doi.org/10.1016/s0076-6879\(96\)66011-x](https://doi.org/10.1016/s0076-6879(96)66011-x)

42. Madeira, F., Pearce, M., Tivey, A.R.N., et al. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*. Apr:gkac240. DOI: 10.1093/nar/gkac240.
43. Mazel D., 2006. Integrons: agents of bacterial evolution. *Nat Rev Microbiol* 4:608 – 620
44. Nagy-Staron, A., Tomasek, K., Caruso Carter, C., Sonnleitner, E., Kavčič, B., Paixão, T., Guet, C. C., 2021. Local genetic context shapes the function of a gene regulatory network. *eLife*, 10, e65993. <https://doi.org/10.7554/eLife.65993>
45. Neubauer, C., De Souza-Pil, M., Bojesen, A.M., Bisgaard, M., Hess, M., 2009. Tissue distribution of haemolytic *Gallibacterium anatis* isolates in laying birds with reproductive disorders. *Avian Pathol*, 38, 1-7. doi: 10.1080/03079450802577848.
46. Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens, R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42(Database issue):D206-14. doi: 10.1093/nar/gkt1226
47. Persson, G., Bojesen, A.M., 2015. Bacterial determinants of importance in the virulence of *Gallibacterium anatis* in poultry. *Vet Res.* 46(1):57–67. doi: 10.1186/s13567-015-0206-z
48. Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research*, 29(9), e45. <https://doi.org/10.1093/nar/29.9.e45>
49. Pors, S.E., Skjerning, R.B., Flachs, E.M. et al., 2016. Recombinant proteins from *Gallibacterium anatis* induces partial protection against heterologous challenge in egg-laying hens. *Vet Res* 47, 36 <https://doi.org/10.1186/s13567-016-0320-6>
50. Roodt, Y., Bragg, R.R., Albertyn, J., 2012. Identification of prophages and prophage remnants within the genome of *Avibacterium paragallinarum* bacterium. *Sequencing* 2012:953609. doi:10.1155/2012/953609.

51. Rossi, R.M., Yum, L., Agaisse, H., Payne, S.M., 2017. Cardiolipin Synthesis and Outer Membrane Localization Are Required for *Shigella flexneri* Virulence. *mBio*. 8(4):e01199-17. doi: 10.1128/mBio.01199-17.
52. Saleski, T.E., Chung, M.T., Carruthers, D.N., Khasbaatar, A., Kurabayashi, K., Lin, X.N., 2021. Optimized gene expression from bacterial chromosome by high-throughput integration and screening. *Science advances*, 7(7), eabe1767. <https://doi.org/10.1126/sciadv.abe1767>
53. Salgado Lucio, M.L., Vaca, S., Vázquez, C., Zenteno, E., Rea, I., Pérez-Márquez, V.M., Negrete-Abascal, E., 2012. Adhesion of *Gallibacterium anatis* to Chicken Oropharyngeal Epithelial Cells and the Identification of Putative Fimbriae. *Advances in Microbiology*, Vol. 2(4):505-510. doi: 10.4236/aim.2012.24064
54. Smyshlyaev G., Bateman A., Barabas O., 2021. Sequence analysis of tyrosine recombinases allows annotation of mobile genetic elements in prokaryotic genomes. *Molecular Systems Biology* (2021)17:e9880, <https://doi.org/10.15252/msb.20209880>
55. Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. *Bioinformatics*. 27(7):1009-10. doi: 10.1093/bioinformatics/btr039.
56. Suthers, P.F., Dasika, M.S., Kumar, V.S., Denisov, G., Glass, J.I., Maranas, C.D., 2009. A genome-scale metabolic reconstruction of *Mycoplasma genitalium*, iPS189. *PLoS computational biology*, 5(2), e1000285. <https://doi.org/10.1371/journal.pcbi.1000285>
57. Tamura, N., Lottspeich, F., Baumeister, W., Tamura, T., 1998. The role of tricorin protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell*. 95(5):637-48. doi: 10.1016/s0092-8674(00)81634-7.
58. Vaca, S., Monroy, E., Rojas, L., Vazquez, C., Sánchez, P., Soriano-Vargas, E., Bojesen, A.M. Negrete-Abascal, E., 2011. Adherence of *Gallibacterium anatis* to Inert Surfaces. *Journal of Animal Veterinary Advances*, 10:(13)1688-1693. doi:10.3923/javaa.2011.1688.1693

59. Wingett, S.W., Andrews, S., 2018. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res*. 7:1338. doi: 10.12688/f1000research.15931.2.
60. Wu, S., 2015. Invasion. Chp25, Pag 423-448. Editor(s): Tang, Y-W., Sussman, M., Liu, D., Poxton, I., Schwartzman J. *Molecular Medical Microbiology (Second Edition)*. Academic Press, ISBN 9780123971692. <https://doi.org/10.1016/B978-0-12-397169-2.00025-1>.
61. Wурpel, D.J., Totsika, M., Allsopp, L.P., Webb, R.I., Moriel, D.G., Schembri, M.A., 2016. Comparative proteomics of uropathogenic *Escherichia coli* during growth in human urine identify UCA-like (UCL) fimbriae as an adherence factor involved in biofilm formation and binding to uroepithelial cells. *J Proteomics*. 131:177-189. doi: 10.1016/j.jprot.2015.11.001.
62. Zepeda-Lopez, H.M., Perea-Araujo, L., Miliar-Garcia, A., Dominguez-Lopez, A., Xoconostle-Cazarez, B., Lara-Padilla, E., Ramirez Hernandez, J.A., Sevilla-Reyes, E., Orozco, M.E., Ahued-Ortega, A., Villaseñor-Ruiz, I., Garcia-Cavazos, R.J., Teran, L.M., 2010. Inside the outbreak of the 2009 influenza A (H1N1)v virus in Mexico. *PloS one*, 5(10), e13256. <https://doi.org/10.1371/journal.pone.0013256>
63. Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–9. doi: 10.1101/gr.074492.107