



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE
PUEBLA

FACULTAD DE CIENCIAS BIOLÓGICAS



MAESTRÍA EN CIENCIAS BIOLÓGICAS

LÍNEA DE GENERACIÓN Y APLICACIÓN DEL CONOCIMIENTO
(LGyAC)

ESTRUCTURA Y FUNCIONAMIENTO DE LOS SERES VIVOS

TESIS:

“IDENTIFICACIÓN, CARACTERIZACIÓN Y EVALUACIÓN DE LA
ACTIVIDAD DEL PROMOTOR DEL GEN *LGALS9*”

QUE PARA OBTENER EL GRADO DE MAESTRO EN CIENCIAS
BIOLÓGICAS PRESENTA:

Q.F.B. JESÚS APARICIO FLORES

DIRECTORA DE TESIS:

D. EN C. LORENA MILFLORES

CO-DIRECTORA DE TESIS:

D. EN C. VERÓNICA VALLEJO RUIZ



PUEBLA, PUE.

DICIEMBRE 2019

Declaratoria de No Plagio

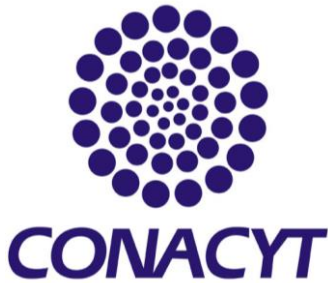
Yo, QFB Jesús Aparicio Flores, con número de matrícula 217470715 alumno de la maestría en Ciencias Biológicas inscripta a la Facultad de Ciencias Biológicas de la Benemérita Universidad Autónoma de Puebla, autor de esta tesis que lleva por título: “Identificación, caracterización y evaluación de la actividad del promotor del gen *LGALS9*”, declaro que el presente trabajo de investigación es original, siendo el resultado de mi trabajo personal, el cual no he copiado de otro trabajo de investigación y de mi autoría. También que las citas de otros autores han sido debidamente identificadas en el presente trabajo de investigación.

Atte.

QFB Jesús Aparicio Flores

Puebla, Pue. Diciembre 2019.

Agradecimientos



Al Consejo de Ciencia y Tecnología (CONACYT) por la beca otorgada para la realización de este trabajo de investigación.

No. de becario: 712280.

A la Vicerrectoría de Investigación y Estudios de Posgrado (VIEP) por la beca otorgada.



A la maestría en Ciencias Biológicas de la Facultad de Ciencias Biológicas.

Índice

Resumen.....	1
Introducción	2
Antecedentes	9
Planteamiento del problema	12
Justificación.....	12
Preguntas de investigación.....	13
Objetivo general	14
Objetivos específicos	14
Esquema de trabajo	15
Material y métodos.....	16
Resultados	27
Discusión de resultados.....	37
Conclusiones	41
Perspectivas.....	42
Referencias.....	43
Anexos.....	48

Resumen

La Galectina 9 (Gal-9) es una galectina de tipo repetición en tándem, con un peso molecular de entre 34 a 39 KDa, con un par de dominios de reconocimiento a carbohidratos (DRC N- y C-) unidos por un polipéptido corto llamado dominio de unión. En la superficie extracelular la Gal-9 se une a oligosacáridos que contienen galactosa, también puede interactuar con glicoproteínas de la matriz extracelular (laminina, fibronectina), formando interacciones bivalentes o multivalentes. La participación de Gal-9 en la agregación celular, la inhibición de la invasión y la unión al endotelio vascular fue observada en células de cáncer de mama y carcinoma hepatocelular.

Gal-9 es codificada por el gen *LGALS9*, se encuentra localizado en el cromosoma 17, en el brazo largo región 11.2 (p11.2), hasta ahora el promotor del gen *LGALS9* no ha sido descrito.

En el presente trabajo se realizó un análisis *in silico* para poder encontrar la región promotora del gen *LGALS9*, al realizarlo se obtuvo dos posibles regiones promotoras río arriba de este gen. Al analizar estas regiones se localizaron sitios de unión a factores de transcripción como TFDII y TBP los cuales reconocen cajas TATA, además de estos sitios fueron localizados sitios Inr. Junto con estos sitios fueron hallados sitios de unión a factores de transcripción como Sp1, AP-2, AP-1 y GATA-1, los cuales se han observado como activadores de la transcripción.

Se propone que el promotor del gen *LGALS9* es un promotor clásico el cual es activado por cajas TATA las cuales son reconocidas por TFDII y TBP. Al realizar el análisis de las diferentes versiones de la posible región promotora, se observó que solo una región presentó actividad con respecto a las otras analizadas.

Haciendo una comparación con el promotor del gen *LGALS3*, el cual codifica para la Galectina-3, este promotor solo presenta sitios Inr y Sp1 por los cuales se mantiene la actividad transcripcional, además de carecer de cajas TATA, para el promotor de *LGALS9* se encontró que puede ser activado por cajas TATA, sitios Inr y por otros factores de transcripción que pueden aumentar su transcripción.

Palabras clave: Galectina-9, factores de transcripción, cajas TATA, sitios Inr.

Introducción

El término galectina fue introducido en 1994 de la manera siguiente: “la membresía en la familia galectina requiere el cumplimiento de un par de criterios: afinidad por β -galactósidos y una similitud en la secuencia del sitio de unión a carbohidratos cuyos aminoácidos relevantes hayan sido identificados por cristalografía de rayos X” (Barondes *et al.*, 1994). Todas las galectinas contienen uno o dos dominios de reconocimiento de carbohidratos (DRC) de alrededor de 130 aminoácidos, el cual presenta afinidad a β -galactosidos (Rabinovich *et al.*, 2002). Las galectinas pueden actuar como monómeros, dímeros u oligómero de orden superior según el caso y las condiciones específicas (concentración, presencia de ligandos) (Leffler, Carlsson, Hedlund, Qian & Poirier, 2004). Hasta ahora 15 miembros de la familia de galectinas han sido identificados en vertebrados y son clasificadas con base en su estructura molecular en tres tipos: 1) las prototipo son las que presentan un solo DRC que es capaz de dimerizar (1, 2, 5, 7, 10, 11, 13, 14 y 15); 2) las de repetición tipo tándem compuesta por dos DRC conectados por un péptido de unión (4, 6, 8, 9 y 12); 3) el tipo quimera el cual consiste de un DRC en el extremo C terminal unido a un dominio diferente de lectina en el extremo N terminal, solo contiene un miembro, la galectina 3 (Yang, Rabinovich & Liu, 2008), como se muestra en la figura 1.

La identidad de secuencia entre galectinas de humano que se reporta en la literatura se encuentra entre el 20% al 50%, algunas excepciones notables incluyen el DRC C terminal de galectina 12 con menos del 12% con respecto a otros DRC, las galectinas 10, 13 y 14 presentan de más del 50% de identidad (Liu & Rabinovich, 2009).

Las galectinas son proteínas que se localizan intracelular y extracelularmente y ejercen diversas funciones dependiendo de su localización. En el espacio extracelular modulan la señalización celular mediante la unión a glicoconjugados y la adhesión celular a través haciendo puente en las interacciones célula-célula y célula-matriz extracelular (Hughes, 2001). Por ejemplo, pueden reconocer laminina, fibronectina y vitronectina en la matriz extracelular (Brinchmann, Patel & Iversen, 2018). En el citosol participando en vías de señalización. En el núcleo, regulando procesos como el splicing de ARNm, la progresión del

ciclo celular y proliferación (Elola, Wolfenstein-Todel, Troncoso, Vasta & Rabinovich, 2007).

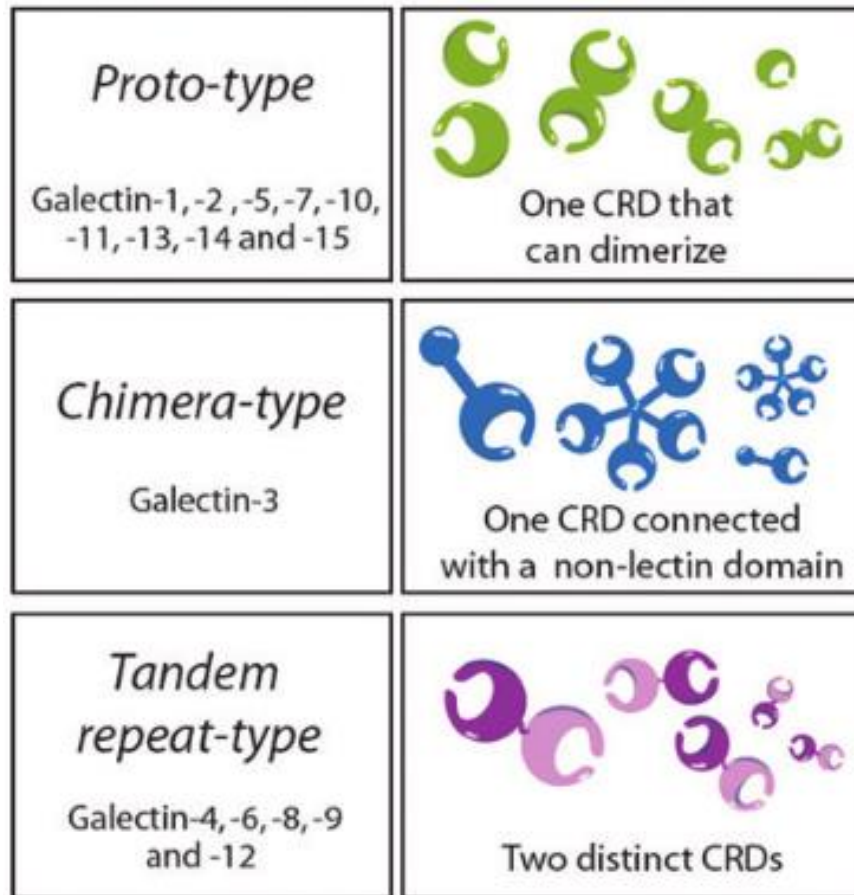


Figura 1.- Clasificación de las galectinas (tomado de Liu & Rabinovich, 2009).

Entre los ligandos extracelulares conocidos para las galectinas, a nivel extracelular, se encuentran CD45, CD7, CD43, CD2, CD3, CD4, CD107, CEA, laminina, fibronectina, glicosaminoglicanos, integrinas, gangliósido GM1, polipéptido HBGP82, antígeno de cáncer CA125, glicoproteína 90K/MAC-2BP y receptor de células pre-B. Las galectinas pueden ser moduladores positivo o negativos del mismo proceso, esto sugiere plasticidad en sus isoformas y una función diferente en relación a su localización (Brinchmann, Patel & Iversen, 2018).

En humanos, la Galectina 1 está presente en el espacio extracelular, es expresada por células estromales alrededor de las células pre-B, donde se une al receptor de células pre-B (BCR) y está involucrada en la formación sináptica entre células pre-B y células estromales, también regula su proliferación de manera negativa mediante la unión del coactivador transcripcional OCA-B. Se ha mostrado que afecta la homeostasis de las células T modulando la producción de citosinas, proliferación celular y apoptosis. La apoptosis de células T es inducida por la unión a glicoproteínas de la superficie celular y liberando vías de señalización de apoptosis. Un aumento en la expresión de la proteína de galectina 1 es asociado con mal pronóstico y la adquisición del fenotipo metastásico. Galectina 1 aumenta su expresión en tumores favoreciendo el crecimiento e influencia su progresión y metástasis modulando varios eventos biológicos como la migración celular, adhesión y angiogénesis. Esta se une a H-RAS y promueve su anclaje a la membrana plasmática, este proceso es clave en el papel de esta proteína en la transformación neoplásica, produciendo la activación de RAF-1 y de MAPK1 (Yang, Rabinovich & Liu, 2008).

La Galectina 3 es el único miembro de esta familia con una región N terminal compuesta por repeticiones en tándem de cortos segmentos de aminoácidos (120 aminoácidos) conectado al DRC C terminal. La actividad antiapoptótica de galectina 3 parece dependiente de la fosforilación en la Serina 6. En la matriz extracelular puede unirse a integrinas como $\alpha 1\beta 1$ y $\alpha M\beta 1$, de esta manera influencia la unión de la célula a la matriz celular. Es capaz de formar pentámeros al unirse a un glicano multivalente, la formación del pentámero permite actuar como un puente entre célula y célula (Yang, Rabinovich & Liu, 2008). Se ha mostrado que galectina 3 regula la expresión de ciertos genes, incluyendo los relacionados con cáncer, como el de ciclina D1, factor de transcripción de tiroides 1 y mucina 2. Galectina 3 induce apoptosis en células T y neutrófilos, activa varias células mieloides y linfoides, incluyendo mastocitos y monocitos, resultando en la liberación de superóxido y citosina (Nakamura & Raz, 2007). El mecanismo por el cual galectina 3 está involucrada en la transformación celular y el crecimiento puede estar relacionado a su habilidad de interactuar con K-RASy facilitar su transducción de señal a RAF y a fosfoinositol 3 quinasa, además puede influenciar la tumorigénesis por la regulación del ciclo celular, por la unión a β -catenina y ser involucrada en la expresión de ciclina D y c-myc (Yang, Rabinovich & Liu, 2008).

Galectina 9 es clasificada como una galectina de repetición en tándem por sus dos DRC, está formada por 355 aminoácidos (aa) y con peso aproximado de 39.5 KD. El DRC N terminal de 148 aminoácidos comparte el 39% de su secuencia con el DRC C terminal de 149 aa. La cristalografía de rayos X indicó que el N-DRC está formado por seis hebras (S1-S6) y cinco láminas β (F1-F5) además de una corta hélice- α con motivos tipo sándwich- β . La estructura del C-DRC, como el N-DRC, consiste de dos láminas β antiparalelas (S1-S6) y hebras b (F1-F5) con una hélice- α . Ambos N- y C-DRC tienen bolsillos de unión a carbohidratos con hebras S4, S5 y S6, las cuales difieren en la secuencia de aminoácidos uno del otro, resultando en una afinidad por diferentes β -galactósidos. El péptido de unión es flexible permitiendo la rotación de los DRC dándole el potencial para formar complejos más diversos glicano-galectina-9. Galectina 9 presenta afinidad por N-glicanos ramificados y repeticiones de oligolactosamina, la afinidad por los N-glicanos es compartida por los DRC N y C terminal. Aunque el N-DRC se une al pentasacárido Frossman, A-hexasacárido y algunos glicolípidos.

La expresión de galectina 9 (proteína) se distribuye en hígado, intestino delgado, timo, riñón, bazo, pulmón, músculo cardíaco y esquelético, cerebro, placenta, páncreas, próstata y colon. También se ha detectado en leucocitos responsables de la inmunidad innata y adquirida, timocitos, células endoteliales y fibroblastos activados por IFN (Heusschen, Griffioen & Thijssen, 2013).

Las células Th1 presentan un receptor para galectina 9, Tim-3, cuando las células T se diferencian totalmente de células T nativas. La interacción galectina-9-Tim-3 guía a las Th1 maduras a la apoptosis con flujo de calcio y agregación celular, disminuyendo la inmunidad por Th1. Las células T CD+8 y CD+4 son susceptibles al estímulo de galectina 9 exógena y experimentan apoptosis basada en la vía calcio-calpaina-caspasa-1. Galectina 9 también induce apoptosis de células B, las cuales no expresan Tim-3 en la superficie celular, aunque la apoptosis es ejecutada por la familia de las caspasas y tres subfamilias de Bcl-2, caspasa-1 juega un rol más prominente en la piroptosis que en apoptosis. La piroptosis evoca la inflamación mientras que la apoptosis no lo hace, esta muerte celular programada se observa en macrófagos y células dendríticas, es encontrada en situaciones como la infección e inflamación (Fujita, Iwama, Oura, 2017).

Se ha observado un efecto regulador dependiente de galectina 9, producido por la propia célula, sobre Tim-3 y las respuestas inflamatorias en miocitos y macrófagos a través de las vías TLR (receptor tipo Toll) modulando la transcripción de los genes de citosinas, lo que resulta en el aumento de la expresión de IL-12 y disminución de IL-23, este efecto puede ser amplificado por la estimulación de TLR (Toll like receptor) (Matsuura, Tsukada, Mizobe, *et al.*, 2009).

Además, galectina 9 modula la expresión de su receptor (Tim-3) y la transcripción de genes de citosinas inflamatorias a través de la señalización STAT-3. Galectina 9 ha sido reportada como un factor de respuesta al lipopolisacárido (LPS), la cual se transloca hacia el núcleo interaccionando con NF-IL6, activando algunos genes de citosinas (IL-1 α , IL-1 β , y IFN- γ) en monocitos. Galectina 9 intracelular activa la transcripción de IL-12 e inhibe a IL-23 a través de la vía de señalización STAT3, de manera interesante el bloqueo de la vía STAT3 anula la transcripción IL-12/IL-23 mediado por galectina 9, pero no afecta la expresión de Tim-3, lo que sugiere un mecanismo diferente para la regulación génica de Tim-3 (Ma, Li, Cheng, *et al.*, 2013). Galectina 9 intracelular induce la expresión de ARNm de IL-1 α y IL-1 β en monocitos, en contraste galectina 9 extracelular fallo no activa IL1A en las mismas células, además Gal-9 intracelular indujo la activación de los factores de transcripción NF-IL6 y AP-1, sin activar a NF- κ β (Matsuura, Tsukada, Mizobe, *et al.*, 2009).

En fibroblastos las vías de señalización como MAPK, PI3K y JAK/STAT son activadas por IFN- γ , las cuales regulan la expresión de genes involucrados en la iniciación, propagación y resolución de las respuestas inmune e inflamatoria. El tratamiento con inhibidores de estas cinasas reprimió la expresión de galectina 9 estimulada por IFN- γ en citosol como en membrana, demostraron que las vías de señalización JAK/STAT/, PI3K, JNK y p38 representan un rol crucial en la expresión de galectina 9 inducida por IFN- γ (Park, Jung, Park, *et al.*, 2011) y en astrocitos TCR (TNF/TNFR1/JNK/c-Jun) (Zhu *et al.*, 2005).

Las citosinas pro-inflamatorias TNF, IL-1 β y IFN- γ son capaces de inducir la expresión de galectina 9 en cultivos primarios de astrocitos. TNF media la producción de galectina 9 en astrocitos que depende de TNFR1 y de la activación de JNK y c-Jun, un resultado que es respaldado por la presencia de secuencias que reconoce c-Jun dentro del promotor de galectina 9. En células endoteliales de vena umbilical humana estimuladas con IFN- γ , la

expresión de galectina 9 es regulada por la histona deacetilasa 3 (HDAC3) y la fosforilación de IRF3 mediada por PI3K. La vía PI3K ha sido implicada en la regulación de galectina 9 en el pólipo nasal humano de fibroblastos. La presencia de una secuencia de IRF3 en el promotor de galectina 9 de ratón respalda su papel en la regulación de su transcripción. Después de la estimulación de astrocitos con TNF, galectina 9 aumentó su expresión por un mecanismo distinto a la vía IFN/PI3K/IRF3 y requiere la activación de JNK/c-Jun (Steelman, Smith III, Welsh & Li, 2013).

HDAC3 es requerida en la expresión constitutiva de galectina 9. HDAC3 es activada por IFN- γ , en células endoteliales, y forma un complejo con PI3K e IRF3, incrementando la fosforilación de IRF3 mediada por PI3K, guiando a un aumento en la regulación de la transcripción de galectina 9, donde HDAC3 puede regular la expresión en dos maneras: como una proteína de andamio para la interacción de PI3K e IRF3, la cual es dependiente de la actividad de HDAC3, el complejo HDAC3-PI3K-IRF3 puede ser el responsable de la expresión constitutiva de galectina 9, por otro lado la HDAC3 activa puede promover la fosforilación de IRF3 y su acceso al promotor del gen. La inhibición de la actividad de HDAC3 por apicina o knock-out disminuye los niveles de proteína de galectina 9. IFN- γ ejerce un efecto dual en galectina 9: inducción de su expresión y localización en la membrana (Alam, Li, Maragariti, *et al.*, 2011).

Dentro de las Treg Foxp3 juega un papel crítico en el desarrollo y regula un amplio espectro de funciones celulares de Treg, su expresión depende principalmente de la señalización de TGF- β R y de IL-2R. La pérdida de la señalización de TGF- β o IL-2 resulta en una deficiencia en la generación de células Treg. TGF- β activa a las proteínas Smad, las cuales son factores de transcripción que promueven la inducción de un número de moléculas requeridas para la generación de células Treg incluyendo a Foxp3. La activación de TGF- β R directamente desencadena la fosforilación y translocación del receptor que regula las proteínas Smad, las cuales subsecuentemente mediante la unión al locus *Foxp3*, desencadena la transactivación de la expresión de *Foxp3*. La activación de Smad3 induce modificaciones en histonas, resultando en la promoción y estabilización del polipéptido Foxp3. Ensayos de inmunoprecipitación de cromatina (ChIP) fueron encontrados múltiples sitios de Smad3 dentro del promotor *LGALS9* en células iTreg (Wu, Thalhamer, Franca, *et al.*, 2014).

Se ha observado en líneas celulares que provienen de cáncer de mama una disminución en los niveles de ARN mensajero de galectina 9, esta disminución favorece la transformación hacia la malignidad del tumor (Irie, Yamauchi, Kontani, *et al.*, 2005). Por otra parte, se ha observado un incremento en la expresión de ARN mensajero en líneas celulares de leucemia y cáncer de colon (Lahm, Andre, Hoeflich, *et al.*, 2001). Se ha encontrado una disminución en la expresión de galectina 9 en cáncer cervicouterino asociada con el potencial oncogénico (Liang, Ueno, Oomizu. *et al.*, 2008)

Antecedentes

Para la regulación transcripcional de los genes es necesario que presenten elementos como un promotor, el cual está compuesto por un promotor central y elementos reguladores cercanos, además de elementos reguladores distales, los cuales pueden ser enhancers, silenciadores o aislantes (Maston, Evans & Green, 2006). El promotor central es la región que sirve como sitio de acoplamiento para la maquinaria basal de transcripción y el ensamble de un complejo de preiniciación de la transcripción (CPT) (Smale & Kadonaga, 2003).

El primer elemento del promotor central descrito fue la caja TATA, la cual es una secuencia rica en /T localizada entre 25 a 30 nucleótidos río arriba del sitio de inicio de la transcripción (SIT), presenta una secuencia consenso, TATA (A/T) A (A/T), la cual es reconocida por la subunidad TBP del complejo TFIID (Smale & Kadonaga, 2003). Otro elemento del promotor central es el iniciador (Inr), el cual contiene secuencias ricas en pirimidinas alrededor del SIT. Estos sitios son capaces de dirigir de manera precisa el inicio de la transcripción por si solos o con otros elementos, los componentes TAF1/TAF2 de TFIID han sido implicados en el reconocimiento de Inr (Chalkley & Verrijzer, 1999). El ensamblaje de del CPT no depende de un solo punto, sino que muchos de los elementos del promotor central interactúan con TFIID estabilizando su ensamble.

Debido a la participación de las galectinas en diferentes procesos celulares, el estudio de la regulación transcripcional de los genes de galectina es interesante (Chiariotto, Salvatore, Frunzio & Bruni, 2004). Las regiones promotoras de las galectinas 1, 2, 3, 4, 6 y 10 de diferentes especies, además de galectina 9 y 12 de rata han sido clonadas (Gitt, Jordan & Leffler, 1997).

Una pequeña región genómica que abarca el sitio de inicio de la transcripción (-63/+45) es suficiente para la actividad transcripcional del gen de galectina 1 en ratón (*Igals1*). Los elementos *cis* río arriba y río abajo son necesarios para la eficiente actividad transcripcional.

Un sitio de inicio de la transcripción ha sido identificado en la posición -31, han sido encontrados también sitios Sp1 (-57), Inr (-48) que se encuentra sobrepuesta a una caja TATA no canónica (De Gregorio, Chiariotti & Di Nocera, 2001).

Pocos elementos reguladores se han reconocidos en la región promotora del gen *lgals2*, entre ellos un sitio a AML, el factor AML ha sido considerado como un maestro en la regulación de hematopoyesis, su rol en la regulación de este gen solo está restringido al tracto gastrointestinal (Gitt, Jordan & Leffler, 1997; Lutterbach & Hiebert, 2000).

En el promotor del gen *LGALS3* se han encontrado cinco sitios putativos para Sp1, así como también cinco sitios para CRE, cuatro para AP-1 y un sitio para AP-4; además de contar con dos sitios para NF- κ B así como también un elemento sis inducible (SIE). La presencia de sitios de unión Sp1 es una característica de genes que se expresan de manera constitutiva, la presencia de los sitios para CRE y NF κ B sugiere que la activación puede ser regulada a través de las vías de señalización que involucran a la proteína de unión a elementos de respuesta cAMP o NF- κ B, por ejemplo las vías CREB/ATF y NF κ B/Rel pueden estar involucradas en la regulación de la expresión de Galectina 3 por la proteína Tax durante la infección de HTLV-I en células T (Kadrofske, Openo & Wang, 1997; Ahmed & Bandyopadhyaya, 2015).

El promotor del gen *LGALS6* de ratón contiene una secuencia entre -354 a -367 pb con una alta homología en 19 pb de la secuencia con la región río arriba de la apolipoproteína B, el promotor de *lgals6* contiene algunas cajas E, en las cuales se encuentran sitios que se asemejan al sitio de unión para Myc y a MyoD (Gitt, *et al.*, 1998).

Las construcciones del promotor de galectina 10 (*LGALS10*) han sido analizadas encontrando una región entre -292 y -411 pb, la cual contiene dos sitios de unión a GATA, una región rica en purinas como un potencial sitio unión para PU1, además de sitios para Sp1, Oct y EoTF que son necesarios para la total actividad promotora. Al igual que galectina 1, la expresión de galectina 10 puede ser inducida por butirato requiriendo la integridad del sitio Sp1 (Gitt, Jordan & Leffler, 1997; Dyer & Rosenberg, 2001)

En la región promotora del galectina 11 (*LGALS11*) contiene secuencias para c-myb y δ EF1, la presencia de sitios para δ EF1 en la región promotora y en el exón 4 sugiere un rol en la regulación de la expresión de este gen (Ogden *et al.*, 1998).

En la región promotora de galectina 12 (*LGALS12*) de humano, se encuentran localizados sitios Sp1 y AP2, un potencial sitio de unión a CEBP.

Galectina 9 es codificada por el gen *LGALS9* localizado en el brazo corto del cromosoma 17 (17q11.2) el cual consiste de 11 exones (Leffler, Carlsson, Hedlund, Qian & Poirier, 2004).

Para el gen *LGALS9* (galectina 9) el análisis de la región promotora proximal de rata y ratón revela una identidad del 76% entre ellos, los promotores de estas dos especies tienen sitios de unión a factores de transcripción en común, como AP3, CF1 y caja E2, y otros en los que difieren como son Insulin II enhancer, AP1 para rata, AP2, TCF2 α , GATA 1/2/3 para ratón. Un análisis por medio de TRANSFAC (Transcription Factor Analysis Software) muestra que en el promotor de *LGALS9* de humano se encuentran sitios de unión a Kid3, WT1, p300, SREBP, TCFII, STAT, entre otros (Mishra, Grzybek, Niki, *et al.*, 2010). La estimulación de galectina 9 es específica de cada célula, activada por de IFN- γ , IFN- β , IL-1 β , IL-1 α , PMA, LPS, Poly IC y TLR-3. La expresión de galectina 9 puede ser inducida por IFN- γ en células endoteliales (vía HDAC3/P13K/IRF3) (Alam, Li, Maragariti, *et al.*, 2011), fibroblastos (vía MAPK, PI3K and JAK/STAT) (Park, Jung, Park, *et al.*, 2011) y astrocitos TCR (TNF/TNFR1/JNK/c-Jun) (Zhu *et al.*, 2005). En cultivos primarios de neuronas tratados con IFN- γ se encontró un aumento en la transcripción de Gal-9 mediada por STAT1 (Bi, Hong, Lee & Baum, 2011). Otros factores de importancia que pueden incrementar la expresión de galectina 9 son KLF2 y SOX1 (Mishra, Grzybek, Niki, *et al.*, 2010). Resultados preliminares de nuestro grupo de investigación muestran una disminución en la expresión en el ARNm del gen *LGALS9* en presencia de la oncoproteína E6 de VPH-18 en la línea celular HaCat (Caporal, 2017). Hasta ahora el promotor del gen *LGALS9* en humanos no ha sido identificado y mucho menos caracterizado, la búsqueda de la región promotora del gen *LGALS9* para poder conocer de qué manera se regula y cuáles son los factores de transcripción que pueden controlar su activación, así como también de su expresión.

Planteamiento del problema

El promotor del gen *LGALS9* no ha sido reportado a la fecha, por lo que identificar su posible región promotora y realizar su caracterización, representa el punto de inicio para conocer los posibles factores que participen en la regulación de su actividad.

Justificación

En diferentes tipos de cáncer se ha observado una disminución en la expresión de galectina 9, codificada por el gen *LGALS9*, asociándose con la progresión de la enfermedad y el pronóstico de vida. Poco se sabe sobre los mecanismos de regulación de la expresión de este gen por lo que profundizar en su estudio nos ayudara a comprender los factores que podrían estar participando en el cambio de expresión durante el cáncer.

Preguntas de investigación

1. ¿Cuál es la región promotora del gen *LGALS9*?
2. ¿Cuál es el promotor mínimo que mantiene la actividad transcripcional del gen *LGALS9*?

Hipótesis

El promotor del gen *LGALS9* comprende una región que incluye los elementos de respuesta necesarios para formar el complejo de inicio de la transcripción y que contenga elementos de respuesta a factores de transcripción que regulen su actividad.

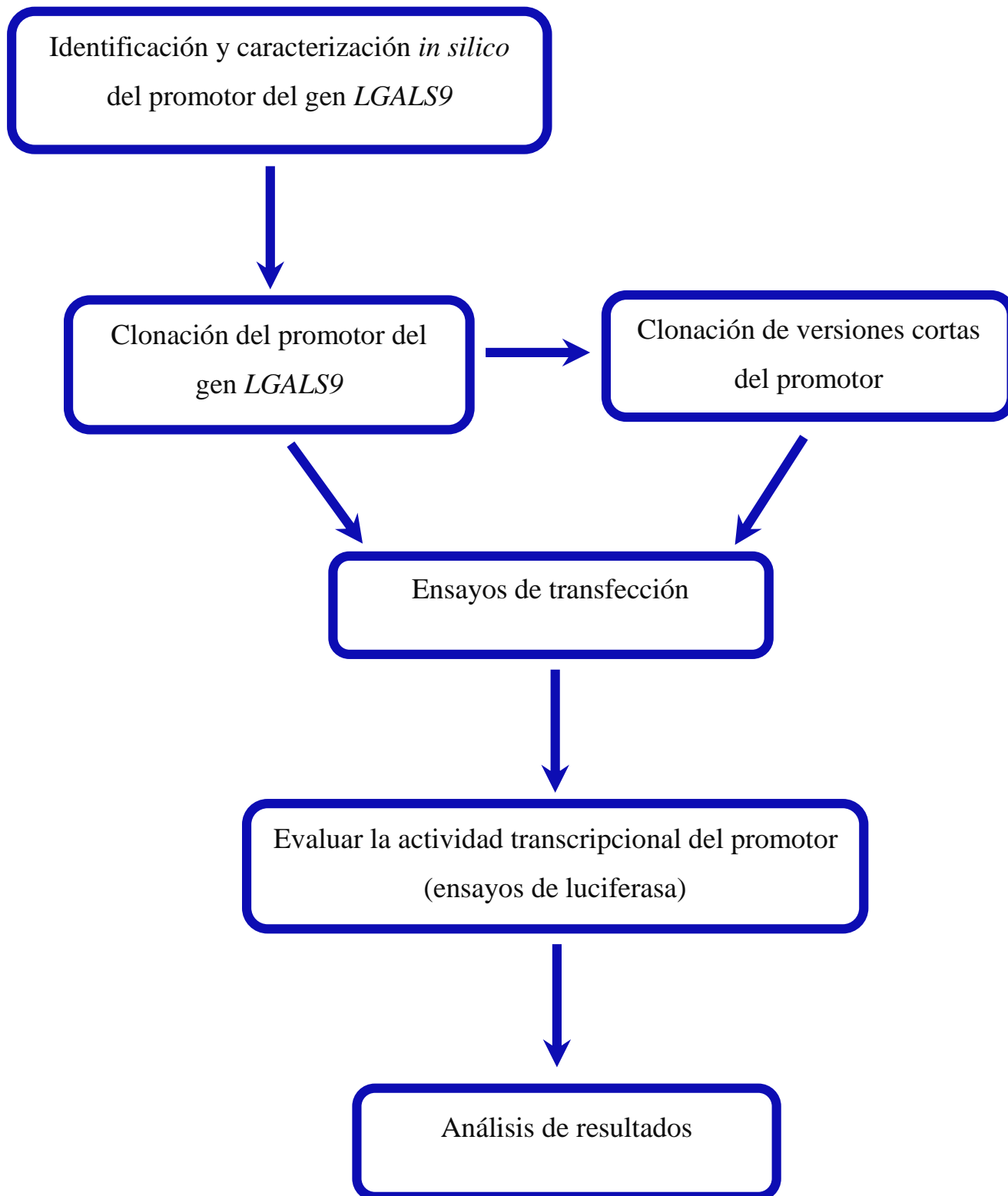
Objetivo general

- Identificar, caracterizar y evaluar la actividad del promotor del gen *LGALS9*.

Objetivos específicos

1. Identificar el posible promotor del gen *LGALS9*
2. Caracterizar el posible promotor del gen *LGALS9*
3. Evaluar la actividad del posible promotor del gen *LGALS9*
4. Identificar el promotor mínimo que mantiene la actividad transcripcional del gen *LGALS9*

Esquema de trabajo



Material y métodos

Identificación del posible promotor del gen *LGALS9*

La secuencia nucleotídica del gen se obtuvo a partir de datos en el *Gene bank* (NC_000017.11), posteriormente por medio del programa EPD (Eukaryotic Promoter Database) del sitio web <https://epd.vital-it.ch/index.php> se identificó la posible región promotora, con el número de referencia FP022909 LGALS9_1. Además por medio de las páginas <http://vega.archive.ensembl.org/index.html> y <https://www.ncbi.nlm.nih.gov/> se obtuvieron dos variantes del ARNm reportadas, las cuales se utilizaron para poder encontrar el codón de inicio de la transcripción del gen *LGALS9*.

Identificación “*in silico*” de los sitios de unión a factores de transcripción del posible promotor del gen *LGALS9*

Para identificar los sitios de unión a factores de transcripción dentro del promotor del gen *LGALS9* se hizo uso del programa Patch1.0 (Public 1.0 Pattern Search for Transcription Factor Binding Sites) del sitio web <http://gene-regulation.com/pub/programs.html>.

Una vez realizado el análisis de los sitios de unión a factores de transcripción en la región del posible promotor del gen *LGALS9* se diseñaron oligonucleótidos con la finalidad de amplificar y clonar dicha región, al mismo tiempo se diseñaron oligonucleótidos para poder identificar la región mínima, se utilizaron los siguientes oligonucleótidos mostrados en la tabla 1:

Tabla 1.-Secuencia de los oligonucleótidos utilizados para la amplificación del promotor.

Oligo	Pb	Secuencia	Tm (°C)
1Gal9	26	GAACATTTCCAGCCTCTCGAGGACTC	59
2Gal9	26	CCAACCTTTCTCGAGCTGGTCTCAAG	59
4Gal9	24	CAGAAGCTCCTGGGTACCCCATAC	59
5Gal9	25	CAGTCCACCCGGTACCATACCTTTCTCTCTC	59
bGal9	26	CTCACTGGACTCAGGTAGGAAGCTTG	60

Los oligonucleótidos fueron diseñados con sitios de corte para enzimas de restricción, el oligonucleotido reverso bGal9 contiene un sitio de corte para la enzima *HindIII*, para los oligonucleótidos delanteros 1Gal9 y 2Gal9 contienen un sitio de corte para la enzima *XhoI*, los oligonucleótidos 4Gal9 y 5Gal9 contienen un sitio de corte para la enzima *KpnI* para su posterior clonación.

Línea celular HaCat

La línea celular HaCat está formada por queratinocitos inmortalizados, los cuales fueron aislados de piel humana. Fue elegida esta línea celular por su fácil propagación “in vitro” además que estudios previos se ha identificado la presencia de la proteína galectina 9. El cultivo celular se realizó en monocapa, utilizando el medio de cultivo Dulbecco’s Modified Eagle’s Medium (DMEM) alto en glucosa 13.4 g/L (SIGMA), suplementado con 10% de suero fetal bovino (SFB) (GIBCO) y adicionado con antibiótico (10,000 U de penicilina, 10 mg de estreptomomicina) y 25 µg/ml de anfotericina B. La incubación se realizó a la temperatura de 37°C y con una atmosfera de 5% de CO₂.

Extracción de ADN y reacción en cadena de la polimerasa (PCR)

De la línea celular HaCat se realizó la extracción de ADN genómico siguiendo el inserto del kit comercial (Zymo Research) para obtener un templado con mayor pureza y concentración.

La concentración de la muestra de ADN se determinó a una longitud de onda de 260 nm, mientras que la relación de absorbancias A260/280 se utilizó para evaluar la pureza de la muestra.

Condiciones para la reacción de PCR

Las condiciones generales para la reacción en cadena de la polimerasa (PCR) fueron: 95°C 5 minutos (desnaturalización); 95°C 1 minuto, 62°C 1 minuto y 72°C 2 minutos 20 segundos (40 ciclos); 72°C 7 minutos (elongación), utilizando la enzima DreamTaq PCR Master Mix (2x) de Thermo Scientific, usando el termociclador miniPCR.

Los tamaños de los fragmentos amplificados se muestran en la tabla 2.

Tabla 2.- Oligonucleótidos usados y tamaño de los fragmentos.

Oligonucleótidos	Tamaño del fragmento (pb)
1Gal9/bGal9	2089
2Gal9/bGal9	1479
4Gal9/bGal9	727
5Gal9/bGal9	222

La electroforesis de los fragmentos amplificados se realizó en un gel de agarosa al 1% teñido con el reactivo Midori Green Advance, visualizados por medio del transiluminador UVP High Performance.

Vectores a utilizar

Las diferentes regiones amplificadas del promotor del gen *LGALS9* fueron clonadas dentro del vector pGL4.12 el cual contiene al gen reportero de luciferasa *luc2CP* de *Photinus pyralis* (luciérnaga), en la figura 1 se muestra las características del vector, el cual presenta un sitio

de multiclonación con diferentes sitios para enzimas de restricción y además un gen de resistencia a ampicilina el cual fue utilizado como marcador de selección de clonas. Las clonaciones se realizaron a través del sitio de multiclonación dentro del vector pGL4.12 y a los sitios de corte producidos por los oligonucleótidos en los fragmentos amplificados, usando las enzimas de restricción *HindIII*, *XhoI* y *KpnI*.

Además, se utilizó el vector pGL4.72 el cual contiene el gen *hRlucCP* de *Renilla reniformis* (esponja de mar) el cual no presenta promotor, sirvió como control interno evaluando la eficiencia de la transfección además de normalizar los valores de la actividad de luciferasa, en la figura 2 se muestra las características del vector pGL4.72.

Además, el vector pGL3 fue utilizado como control positivo en la actividad de luciferasa, el cual contiene el promotor SV40 rio arriba del gen *luc+*, en la figura 3 se muestra las características de este vector.

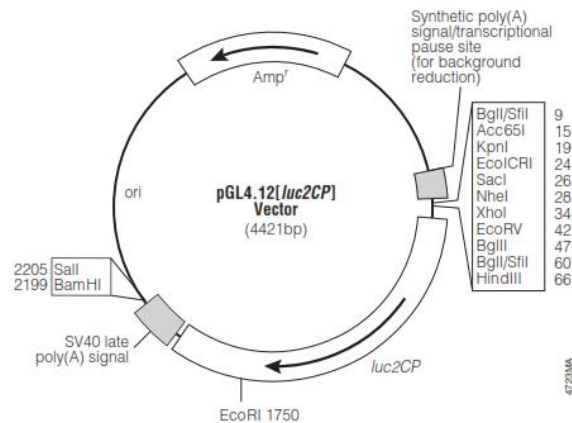


Figura 1.- Representación esquemática del vector pGL4.12

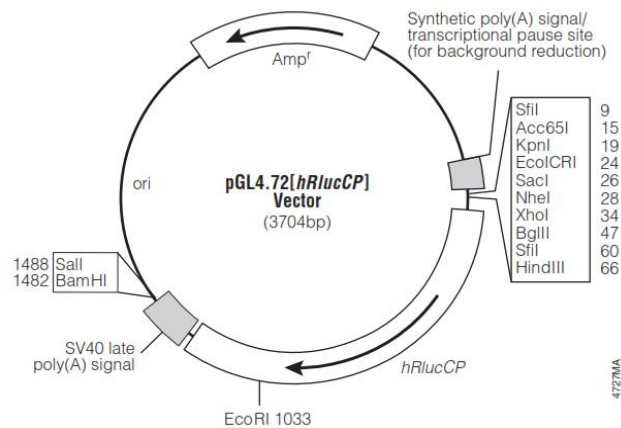


Figura 2.- Representación esquemática del vector pGL4.72

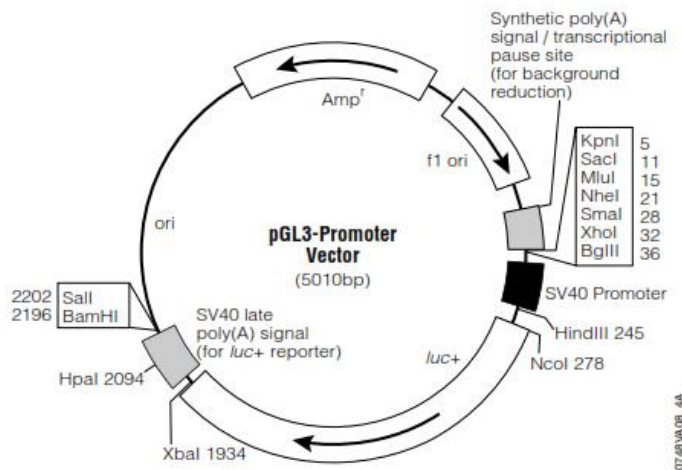


Figura 3.- Representación esquemática del vector pGL3

Extracción de banda

Se realizó una electroforesis en gel de agarosa al 1% para la extracción de banda de los fragmentos amplificados utilizando el kit Zymoclean Gel DNA Recovery, por medio del siguiente protocolo: se visualizó en gel de agarosa el fragmento cortando el fragmento de manera cuidadosa, se agregó el buffer ADB incubando a 37°C durante 10 minutos,

mezclando cada 2 minutos aproximadamente; la mezcla se transfirió a la columna zymo-spin centrifugando a 12 000 RPM 1 minuto, agregado después 200 µl de DNA Wash Buffer centrifugando a 12 000 RPM por 35 segundos, repitiendo este paso 2 veces. Se agregó 16 µl siguiente de agua destilada estéril, incubando 1 minuto a temperatura ambiente y centrifugando a 12 000 RPM durante 1 minuto.

Condiciones de restricción y ligación.

Los productos amplificados y el vector pGL4.12 fueron digeridos por las enzimas de restricción *HindIII*, *XhoI* y *KpnI* del kit FastDigest Value Pack de Thermo Scientific, siguiendo los volúmenes siguientes (Tablas 3, 4 y 5):

Tabla 3.- Cantidades utilizadas en la digestión de los fragmentos 2089 y 1479 pb.

Fragmentos 2089 y 1479 pb	
PCR (fragmento)	1 µg
<i>XhoI</i>	1 U/µl
<i>HindIII</i>	1 U/µl
Buffer	2 µl
H ₂ O	Necesario para 20 µl

Tabla 4.- Cantidades utilizadas en la digestión de los fragmentos 727 y 222 pb.

Fragmentos 727 y 222 pb	
PCR (fragmento)	1 µg
<i>KpnI</i>	1 U/µl
<i>HindIII</i>	1 U/µl
Buffer	2 µl
H ₂ O	Necesario para 20 µl

Tabla 5.- Cantidades utilizadas en la digestión del vector pGL4.12

Vector pGL4.12	
Vector	1 µg
<i>XhoI/KpnI</i>	1 U/µl
<i>HindIII</i>	1 U/µl
Buffer	2 µl
H ₂ O	Necesario para 20 µl

Tabla 6.- Cantidades utilizadas en la ligación de los fragmentos (inserto) y el vector pGL4.12.

Vector	50 ng
Fragmento de PCR (inserto)	5:1 con relación al vector
T4 DNA Ligase	1 unidad
Buffer	2 µl
H ₂ O	Necesario para 20 µl

Todas las reacciones de digestión fueron incubadas a 37°C en termoblock Dry Thermostat Scientific durante dos horas.

Para la reacción de ligación de los fragmentos y el vector, utilizando las especificaciones de la enzima T4 DNA Ligase de Thermo Scientific.

Todas las reacciones fueron incubadas a 22°C por 3 horas y después a 4°C toda la noche en termoblock Dry Thermostat Scientific.

Preparación de las células competentes

La cepa *E. coli* Top 10 One Shot (F-mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG) fue utilizada para preparar células competentes por CaCl₂ 0.1 M, por medio de la siguiente técnica: se tomó una colonia de la cepa *E. coli* Top 10 One Shot que fue inoculada en 50 ml de medio LB incubando toda la noche a 37°C, de este cultivo se utilizaron 50 µl para ser inoculados en 5 ml de medio LB, incubados durante 3 horas hasta alcanzar la densidad óptica de 0.6, el cultivo fue incubado en hielo durante 15 minutos, transcurrido el tiempo el cultivo fue centrifugado a 4000 RPM por 10 minutos, decantando el sobrenadante y agregando 1 ml de CaCl₂ 0.1 M

frio, incubando durante 10 minutos en hielo. Después se centrifugo 4000 RPM por 10 minutos, decantando el sobrenadante. Al botón celular se le agrego CaCl_2 0.1 M más de glicerol al 20%.

Transformación y selección de clonas

Para la transformación con células competentes por CaCl_2 0.1 M se utilizó la siguiente metodología: se utilizaron 200 μl de células competentes, añadiendo 10 μl de la reacción de ligación, fueron incubadas durante 10 minutos en hielo, incubándolas después a 42°C por 45 segundos. Transcurrido el tiempo se incubó durante 2 minutos en hielo, Se añadió 1 ml de medio LB y se incubo durante 1 hora a 37°C en agitación. Después de la incubación fueron sembrados 100 μl del cultivo en placas de medio LB con ampicilina (LB/Amp) a una concentración de 100 $\mu\text{g}/\mu\text{l}$. Incubando a 37°C durante toda la noche.

Las colonias con forma puntiforme y bien aisladas fueron seleccionadas, sembradas en tubos con 5 ml de medio LB/Ampicilina (LB/Amp), incubando a 37°C toda la noche. El ADN plasmídico de las clonas fue extraído de la siguiente manera: se usaron 4.5 ml del cultivo centrifugando a 12 000 RPM por 5 minutos, decantando y agregando 100 μl de solución STE (solución Tris-HCl, EDTA), mezclando la solución con vortex e incubando por 5 minutos a temperatura ambiente, después se agregaron 200 μl de la solución SDS/NaOH incubando por 5 minutos en hielo. Se agregó 150 μl de la solución de acetato de sodio 3 M incubando por 30 minutos en hielo. Transcurrido el tiempo se centrifugó a 12 000 RPM durante 10 minutos. El sobrenadante fue transferido a un tubo nuevo agregando 800 μl de isopropanol, centrifugando a 12 000 RPM por 5 minutos, se decantó y agrego 1 ml de etanol al 70%, centrifugando 12 000 RPM por 5 minutos. Se retiró el sobrenadante y dejo secar el tubo a temperatura ambiente, después se agregaron 50 μl de agua estéril.

En la tabla 7 se muestran las construcciones obtenidas por ligación del vector y los fragmentos del promotor del gen *LGALS9*:

Tabla 7.- Construcciones obtenidas.

FRAGMENTO DEL PROMOTOR DEL GEN <i>LGALS9</i>	CONSTRUCCIÓN
2089	pGL4.2089pGal9
1479	pGL4.14799pGal9
727	pGL4.727pGal9
222	pGL4.222pGal9

Las clonas seleccionadas fueron sometidas a digestión con la enzima *PstI* kit FastDigest Value Pack de Thermo Scientific, los volúmenes usados se muestran en la tabla 8:

Tabla 8.- Cantidades utilizadas en la digestión de las clonas.

ADN	800 ng
<i>PstI</i>	1 unidad/ μ l
Buffer	2 μ l
H2O	Necesario para 20 μ l

Las reacciones fueron incubadas a 37°C durante 2 horas.

En la tabla siguiente se muestran los fragmentos obtenidos por la digestión con la enzima *PstI*:

Tabla 9.-Fragmentos producidos por la digestión de las clonas.

Constructos	Fragmentos (pb)	
pGL4.2089pGal9	5090	1364
pGL4.1479pGal9	4490	1364
pGL4.727pGal9	3716	1364
pGL4.222pGal9	3214	1364
pGL 4.12	3057	1364

Transfección en la línea celular HaCat

Para la transfección de las clonaciones de los diferentes fragmentos del promotor del gen *LGALS9* se utilizó la línea celular HaCat, la cual fue crecida en medio DMEM con 10% de suero fetal bovino (SFB) y antibiótico (Ampicilina/Anfotericina) (Ab), en botellas T25. La confluencia celular usada para transfectar fue del 80% en las botellas T25, usando la cámara de Neubauer se realizó el conteo de células, fueron sembradas 80 000 células/ml por cada pozo, en placas Corning Costar de 24 pozos, con medio de cultivo DMEM al 10% SFB con Ab incubando por 24 horas a 37°C con 5% de CO₂. Una vez adheridas las células, se les retiró el medio de cultivo, lavando cada pozo con 500 µl de buffer PBS, se retiró el buffer PBS y se agregó 250 de medio de cultivo DMEM sin SFB y sin Ab, incubando las placas en las condiciones antes mencionadas.

Para la preparación de los complejos de transfección se utilizó el reactivo *Lipofectamina 2000* de Invitrogen, los fragmentos del promotor del gen *LGALS9* fueron co-transfectados con el vector pGL4.72 (control de eficiencia de la transfección), además de utilizar los vectores pGL3 (control positivo) y pGL4.12 (control negativo). Se utilizaron las concentraciones de 800 ng de ADN para cada construcción, 20 ng para pGL4.72, 200 ng para pGL3 y 800 ng para pGL4.12. Cada complejo de transfección fue diluido en medio de cultivo DMEM sin SFB y sin Ab, después de prepararlos fueron incubados durante 20 minutos a temperatura ambiente, transcurrido este tiempo se agregó a cada pozo la mezcla de transfección por goteo y se incubó durante 6 horas a 37°C con 5% CO₂. Terminada la incubación se retiró el medio, se hicieron lavados con el buffer PBS, agregando 500 µl de medio de cultivo DMEM al 10% de SFB y Ab, incubando por 24 horas.

Ensayo de luciferasa

Después de transcurridas las 24 horas de incubación se realizó el ensayo de luciferasa ocupando el kit Dual-Luciferase® Reporter Assay System de Promega. Se lavaron las células con PBS, posteriormente se agregó a cada pozo 100 µl del buffer de lisis PLB1x, incubando en agitación durante 2 horas a temperatura ambiente, se tomaron 20 µl del lisado a los cuales se le agregó el reactivo LARII para medir el gen reportero de luciferasa (*Photinus agrepyralis*) que generó una señal luminiscente estable, la cual corresponde a la actividad del promotor del gen *LGALS9*, una vez cuantificada la luminiscencia la reacción es inactivada con el reactivo Stop & Glow y de manera simultánea se agrega el sustrato para la segunda luciferasa (*Renilla reniformis*), usando el luminómetro GloMax20/20 de Promega.

Análisis de resultados

Los valores obtenidos de la actividad de luciferasa fueron analizados para determinar la actividad del posible promotor del gen *LGALS9*. Para determinar si existen diferencias estadísticamente significativas entre las diferentes condiciones experimentales, se realizó una prueba de ANOVA de una vía.

Resultados

Análisis *in silico*

Para identificar la posible región promotora del gen *LGALS9* se realizaron análisis *in silico* para ello se obtuvo la secuencia del gen en la base de datos del NCBI (https://www.ncbi.nlm.nih.gov/nucleotide/NC_000017.11?report=fasta&from=27631148&to=27649560) con número de referencia NC_00017.11.

El gen *LGALS9* se encuentra localizado en el cromosoma 17q11.2, el cual está formado por 11 exones, comprende la región 27631148 a 27649560 pb, con un tamaño de 18412 pb. La secuencia fue analizada por el programa Eukaryotic Promoter Database (EPD) para la búsqueda de posibles regiones promotoras, el programa EPD es una base de datos de promotores eucarióticos que presentan sitios de inicio de la transcripción y características clásicas como cajas TATA y sitios Inr, los cuales son reconocidos por la maquinaria de la transcripción. En un inicio el programa EPD mostró dos posibles regiones promotoras, la primera entre las posiciones 27623260 del gen *KSR-1* y 27649392 pb del gen *LGALS9*; la segunda región entre las posiciones 27630224 y 27630824 pb dentro del gen *LGALS9* (datos obtenidos en la página <http://atlasgeneticsoncology.org/index.html>). Por esta razón, se diseñaron oligonucleótidos para amplificar un fragmento de 2089 pb que incluyeran los dos posibles promotores, se realizó una búsqueda nuevamente de la región promotora resultando que solo una posible región promotora de 600 pb (en las posiciones 27630224 y 27630824 pb dentro del gen *LGALS9*, líneas amarillas) la cual se muestra en la figura 4. Con la ayuda de la página <https://www.ncbi.nlm.nih.gov/> se identificaron las diferentes variantes del ARNm del gen, a partir de las secuencias de éstas se identificó el sitio de inicio de la transcripción (indicado con una flecha roja) al cual se le dio el +1 en la secuencia del fragmento amplificado. La posible región promotora de 600 pb comprende de la posición +162 a la -426 en el fragmento de 2089 pb.

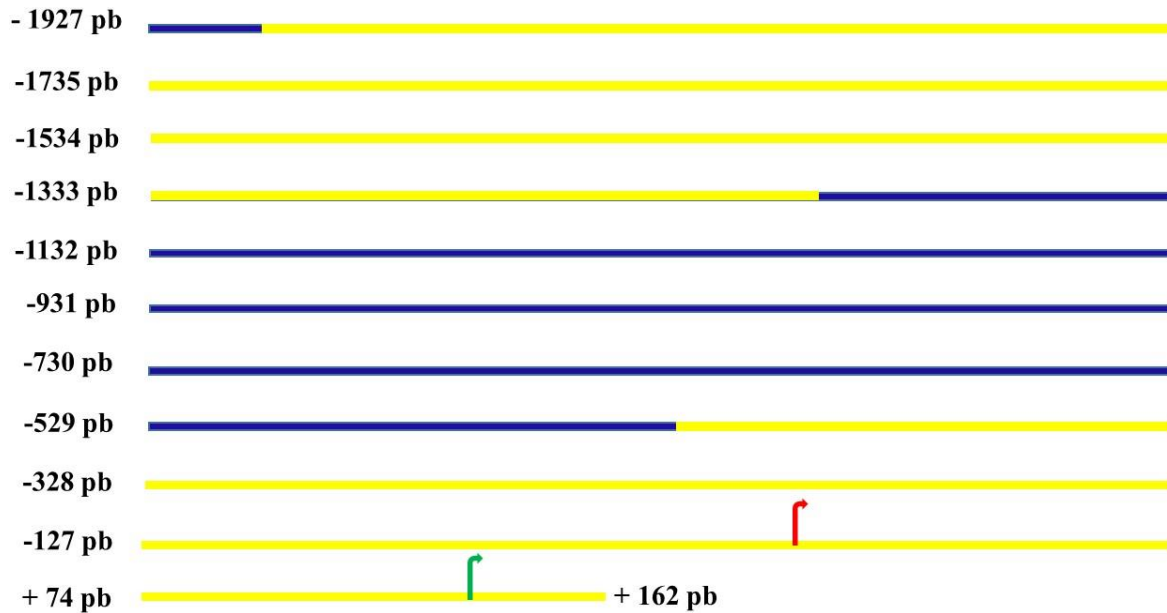


Figura 4.-Representación esquemática del fragmento amplificado de 2089 pb, en color amarillo las dos posibles regiones promotoras. La flecha en rojo (↑) indica el sitio de inicio de la transcripción, la flecha en verde (↑) indica el sitio de inicio de la traducción.

Una vez identificadas las posibles regiones promotoras se realizó el análisis para encontrar los sitios de unión a factores de transcripción del fragmento de 2089 pb, el análisis se realizó a través del programa Patch 1.0 del sitio web <http://gene-regulation.com/pub/programs.html>, es un programa que identifica sitios de unión a factores de transcripción en la secuencia de ADN, el criterio utilizado para elegir los factores de transcripción mostrados en la figura 4 fue que el score para cada factor de transcripción fuera de 100%, de esta manera se obtuvieron los siguientes sitios a factores de transcripción en la secuencia.

Dentro de la secuencia, mostrada en la figura 5, se encontraron cajas TATAA las cuales son sitios de unión a los factores TFIID y TBP, en las posiciones +5, los cuales forman parte de la maquinaria basal de la transcripción, estos factores nos hacen pensar que el promotor del gen *LGALS9* es un promotor clásico. También fueron encontrados sitios de unión a Sp1, Sp2, Sp3 y Sp4 estos factores de transcripción contienen estructuras de dedos de zinc, y se conoce que son potentes activadores transcripcionales, también se ha reportado que Sp1 y Sp3

interactúan con factores de transcripción de la maquinaria basal. Adicionalmente, fueron localizados sitios para los factores AP-1 y AP-2 los cuales también se han observado como activadores de la transcripción en genes que se encuentran implicados en el desarrollo del crecimiento, diferenciación y proliferación. La unión de AP-2 junto con Sp1 potencializa su efecto activador.

Clonación de las posibles regiones promotoras

Una vez identificados los factores de transcripción fueron diseñados oligonucleótidos en sitios específicos en los cuales fueron eliminados algunos de estos sitios de unión a factores de transcripción, por lo cual se realizó una PCR (ver material y métodos) para amplificación de los diferentes fragmentos del posible promotor.

En la figura 6 se muestra la electroforesis de los fragmentos amplificados de 2089 pb y 1479 pb para su clonación. Para su posterior clonación los fragmentos amplificados, así como el vector pGL4.12 fueron digeridos con las enzimas de restricción *HindIII* y *XhoI*.

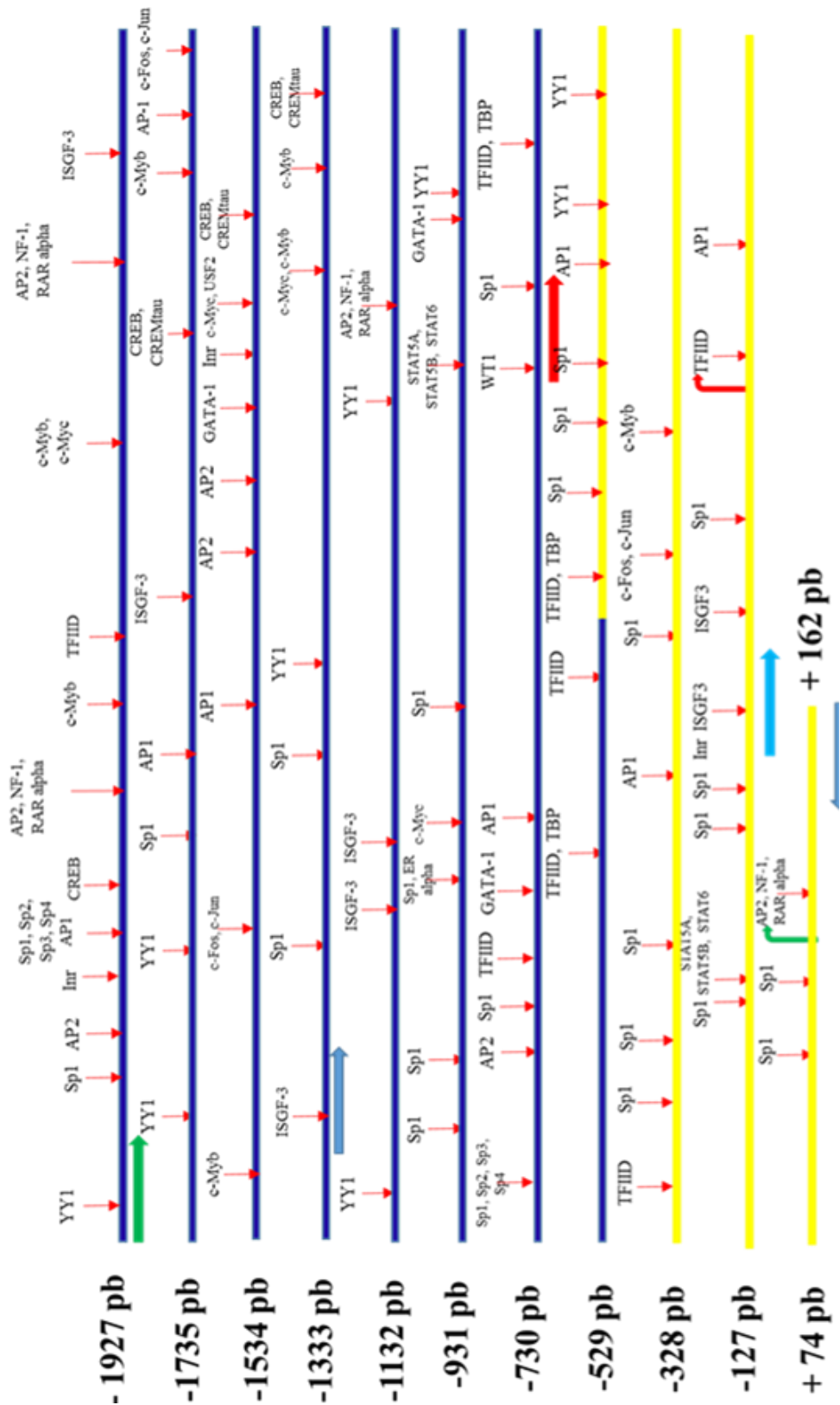


Figura 5.- Representación esquemática de los sitios de unión a factores de transcripción en la posible región promotora del gen *LGALS9*. Las flechas de colores son los primers utilizados.

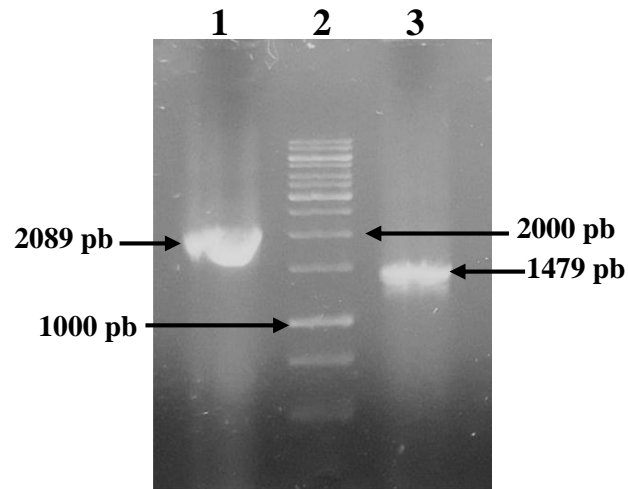


Figura 6.- Electroforesis en gel de agarosa: carril 1) fragmento de 2089 pb (posible región promotora), 2) marcador de peso molecular y 3) fragmento de 1479 pb (posible región promotora).

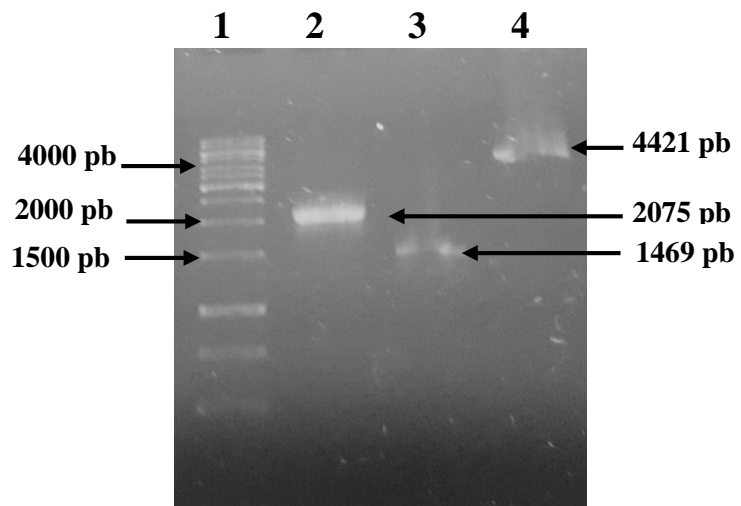


Figura 7.- Electroforesis en gel de agarosa: carril 1) marcador de peso molecular 1 Kb, 2) fragmento de 2075 pb, 3) fragmento de 1469 pb y 4) vector pGL 4.12 (4421 pb)

En la figura 7 se muestra la electroforesis de la restricción para los fragmentos de 2089 pb, 1479 pb y el vector pGL4.12. Del mismo modo se realizó la amplificación para los fragmentos de 727 pb y 222 pb de la posible región promotora.

Los fragmentos de 2089 y 1479 pb fueron digeridos por las enzimas *HindIII* y *XhoI*, además, los fragmentos de 727 y 222 pb fueron digeridos por las enzimas *HindIII* y *KpnI*. Cada fragmento fue ligado junto con el vector pGL4.12, con los sitios de corte de las enzimas correspondientes, posteriormente fueron introducidos en la cepa *E. coli* Top 10 One Shot (véase material y métodos transformación). Se eligieron los sitios de corte de la enzima *PstI* para confirmar que el fragmento de cada versión del promotor se haya insertado dentro del vector pGL 4.12.

En las figuras 8, 9, 10 y 11 se muestran de manera esquemática las construcciones con los diferentes fragmentos del promotor:

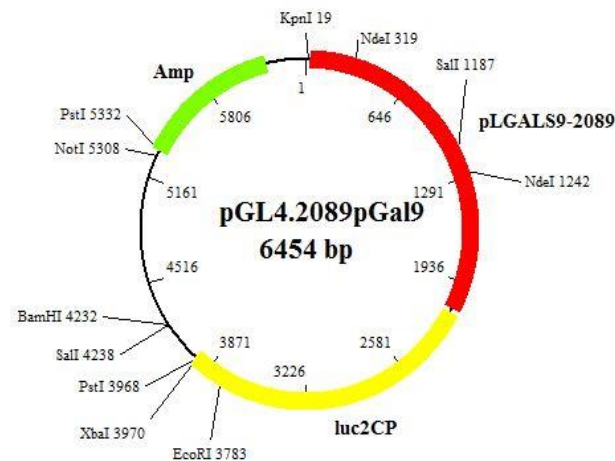


Figura 8.- Representación esquemática de la construcción pGL4.2089pGal9, (fragmento 2089 pb).

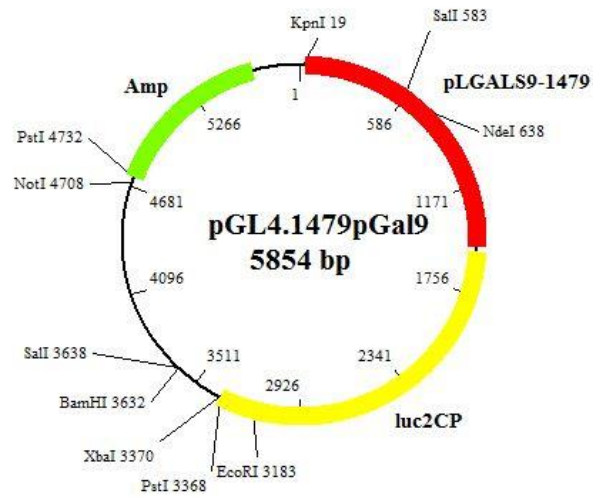


Figura 9.- Representación esquemática de la construcción pGL4.1479pGal9, (fragmento de 1479 pb).

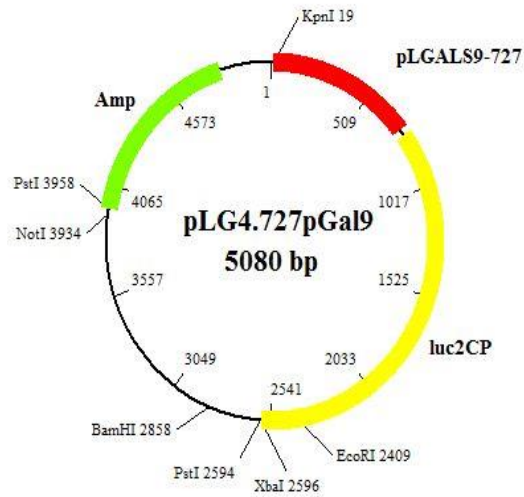


Figura 10.- Representación esquemática de la construcción pGL4.727pGal9 (fragmento de 727 pb).

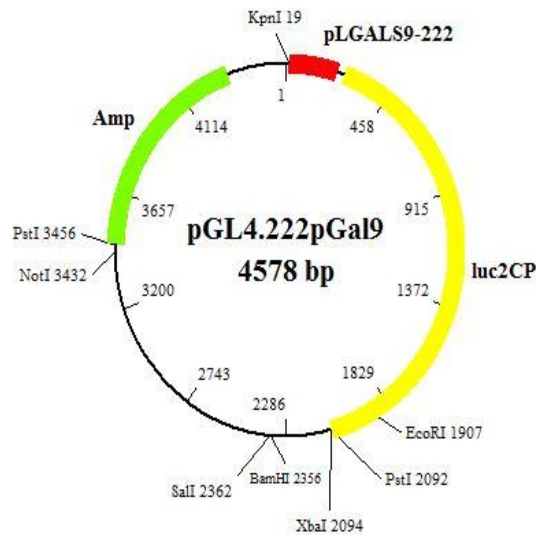
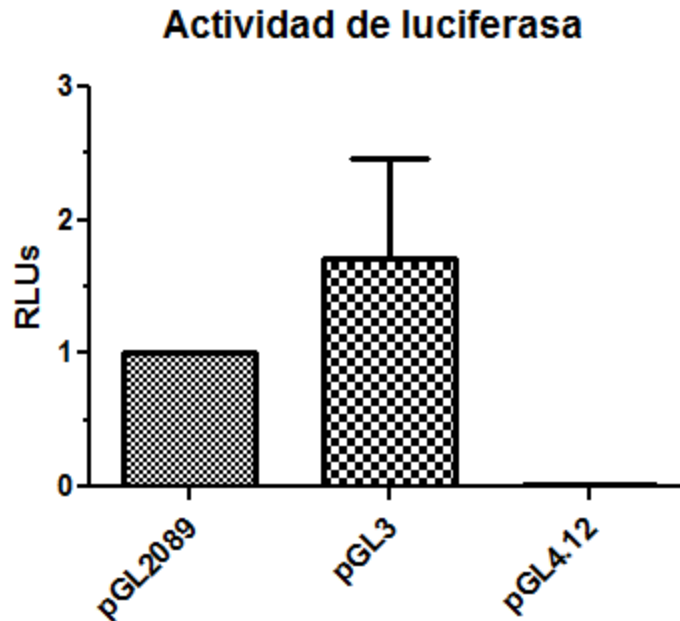


Figura 11.- Representación esquemática de la construcción pGL4.222pGal9 (fragmento 222pb)

Una vez obtenidas las construcciones se continuó con los ensayos de transfección en la línea celular HaCat. Después de realizar los ensayos de transfección se realizó la actividad de luciferasa para cada construcción.

Se evaluó la actividad de la región de 2089 (pGL4.2089pGal9); utilizando como control positivo el vector pGL3 y como control negativo se utilizó el vector pGL4.12, los resultados se muestran en la gráfica 1.

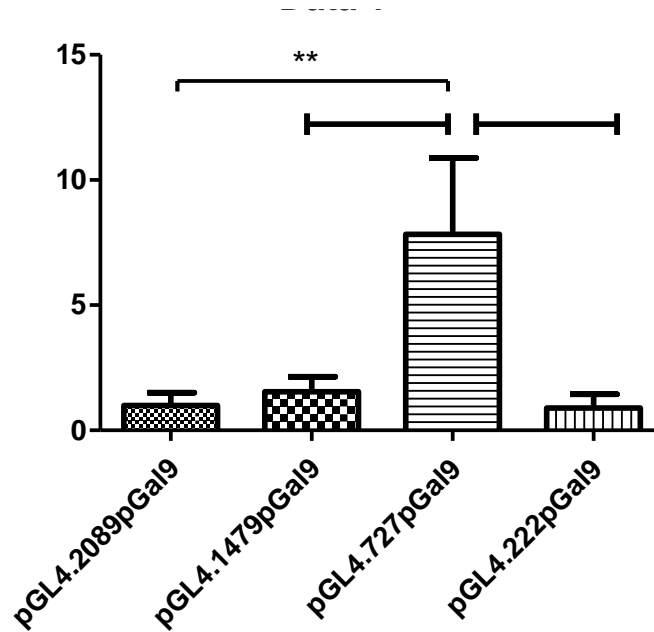


Grafica 1.- Resultados de la actividad de luciferasa comparando la construcción pGL4.2089pGal9, control positivo pGL3 y control negativo pGL4.12. Los resultados se obtuvieron de tres experimentos independientes, aplicando la prueba una ANOVA de una vía con una $p < 0.05$.

Se observó que el fragmento de 2089 pb (pGL4.2089pGal9) mostró actividad transcripcional respecto al vector pGL4.12 que no contiene promotor, posterior a esto nos dimos a la tarea de analizar fragmentos más cortos del promotor (1479, 727 y 222 pb).

En la gráfica 2 se muestran las actividades de luciferasa, el ensayo se realizó por triplicado en tres ensayos diferentes, el análisis estadístico considero una $p < 0.05$. Los análisis nos indican que los fragmentos clonados presentan actividad promotora, se observa una significancia en tres construcciones, pGL4.2089pGal9, pGL4.1479pGal9 y pGL4.222pGal9, con respecto a la construcción pGL4.727pGal9 la cual presenta una mayor actividad. Para los fragmentos de 2089pb y 1479 pb la menor actividad con respecto a la de 727, podría ser resultado de que contiene sitios de unión para el factor CREM que funciona como represor de la transcripción. También se observa que el fragmento de 222 pb presenta actividad

semejante a los fragmentos de 2089 y 1479 pb, por lo que este fragmento puede representar el promotor central del gen *LGALS9*.



Grafica 2.- Resultados de la actividad de luciferasa para cada fragmento de la posible región promotora del gen *LGALS9*. Los resultados se obtuvieron de tres experimentos independientes, aplicando la prueba una ANOVA de una vía con una $p < 0.05$.

Discusión de resultados

La estructura de los promotores de genes humanos suele ser compleja, comúnmente consisten de múltiples elementos reguladores transcripcionales (Kouzine, Levens & Baranello, 2014). La presencia de múltiples elementos reguladores dentro de los promotores confiere un control combinado de la regulación por lo cual incrementa el número de patrones de expresión (Maston, Evans & Green, 2006). Los factores involucrados en la transcripción correcta y precisa de los genes por la ARN pol II pueden ser clasificados en tres grupos: factores de transcripción generales (básicos), proteínas activadoras específicas del promotor (activadores) y coactivadores (Maston, Evans & Green, 2005). Los factores de transcripción generales (FTG) son esenciales además de suficientes para un inicio de la transcripción preciso (Orphanides, Lagrange & Reinberg, 1996), dentro de estos factores se incluyen a la propia ARN pol II además de una variedad de componentes auxiliares como TFIIA, TFIIB, TFIID, TFIIE, TFIIF y TFIIH (Malik & Roeder, 2005).

Los FTG se ensamblan en el promotor central de manera ordenada para formar el Complejo de Preiniciación de la Transcripción (CPT), el cual orienta a la ARN pol II hacia el sitio de inicio de la transcripción (SIT), el primer paso en el ensamble es la unión de TFIID un complejo de múltiples subunidades, una de ellas es la proteína de unión a caja TATA (TBP), después ocurre la unión de la ARN pol II para la elongación, después de que la ARN pol II se haya separado del promotor una estructura compuesta por TFIID, TFIIE y TFIIH permanece en el promotor central para la reiniciación de la transcripción para la cual solo se requiere el reclutamiento de TFIIB, TFIIF y la ARN pol II (Hanh, 2004).

Con el objetivo de identificar si el gen *LGALS9* presenta una región promotora se utilizó el programa EPD (Eukaryotic Promoter Database) el cual es una base de datos de promotores que han sido identificados de manera experimental, tomando en cuenta las cajas TATA y sitios Inr los cuales son característicos en los promotores (Dreos, Ambrosini, Cavin & Bucher, 2013).

El programa EPD proporciono dos posibles regiones promotoras, cada una con un tamaño de 600 pb, en un fragmento de 2089 pb del gen *LGALS9*, la primera localizada en la región -18222 y -1222 pb y la segunda en la región -466 y +134 pb con respecto al sitio de inicio de la transcripción del gen *LGALS9*.

Con el propósito de conocer que otros sitios de unión a factores de transcripción localizados en la posible región promotora se utilizó el programa Patch1.0 (Public 1.0 Pattern Search for Transcription Factor Binding Sites), es un programa que predice los sitios de unión a factores de transcripción.

En esta región fueron encontradas ocho cajas TATA que son reconocidas por TFIID y TBP, la caja TATA es un elemento basal del promotor que ha sido bien caracterizado, es una secuencia rica en A/T, la secuencia consenso generalmente es TATAAA, aunque existen algunas variaciones, se encuentra localizada entre 20 a 30 nucleótidos río arriba del sitio de inicio de la transcripción, esta secuencia es reconocida por TBP (TATA binding protein) que es una subunidad del complejo TFIID (Kutach & Kadonaga, 2000; Yella & Bansal, 2017). Como ya se mencionó no existen estudios sobre las regiones promotoras del grupo de las galectinas, solo se reportan análisis *in silico*. Para el caso de la región promotora de la galectina 1, 2, 6, 10 y 11 se reporta la presencia de una sola caja TATA y nosotros identificamos 8 cajas (en las posiciones -1836, -677, -536, -464, -436, -416, -327 y +6), el resto de los promotores de las otras galectinas no presentan caja TATA. Otro elemento basal del promotor es Inr (Iniciador) es una secuencia que engloba el sitio de inicio de la transcripción la cual es capaz de iniciar la transcripción en ausencia de caja TATA, los elementos Inr pueden estar presente o ausente en promotores que contienen caja TATA (Kadonaga & Bansal, 2000), en la región promotora del gen *LGALS9* se encontraron tres sitios Inr (en las posiciones -1861, -1372 y -46), los cuales podrían cooperar con las cajas TATA localizadas en esta región. Estos sitios Inr solo se identificaron en los promotores de las galactinas 1 y 10. De manera interesante no se logró identificar elementos BRE (TFIIB Recognition Element) y DPE (Downstream Promoter Element), los cuales también forman parte de la estructura general de un promotor, pero estos elementos tampoco se describen para los promotores de las otras galectinas (Butler & Kadonaga, 2002).

Además, fueron encontrados veinticuatro sitios de unión para el factor de transcripción Sp1, distribuidos a lo largo de la posible región, Sp1 recluta la maquinaria basal de la transcripción, interactuando con TFIID. Sp1 reconoce secuencias consenso como GC y GT/CACC. Al igual que Sp1, los factores Sp2, Sp3 y Sp4, localizados en dos sitios (-1877 y -724) de la posible región promotora del gen *LGALS9* también son capaces de regular la

actividad transcripcional, se ha observado que estos factores, Sp2, Sp3 y Sp4, unidos a Sp1 son potencializadores de la actividad transcripcional (O'Connor, Gilmour & Bonifer, 2016). Haciendo una comparación con el promotor de *LGALS3* (Galectina 3) el cual presenta cinco sitios para Sp1 y por los cuales mantienen la actividad de este promotor (Kadrofske, Openo & Wang, 1997), el promotor de *LGALS9* presenta un mayor número de sitios de unión para Sp1 por los cuales pueden potencializar su actividad transcripcional.

También fueron localizados tres de sitios de unión para el factor GATA 1 (-1379, -751 y -667), el cual se observado que participa en la regulación del desarrollo de eritrocitos, la secuencia consenso que reconocen es A/TGATA/G, se le ha observado como activador de la transcripción (Tremblay, Sanchez-Ferras & Bouchard, 2018).

Hasta ahora no existe algún reporte sobre el análisis de la región promotora del gen *LGALS9* en humanos, los estudios actuales solo están enfocados en los promotores de rata y ratón, encontrando sitios de unión a AP1 para rata, AP2, TCF2 α , GATA 1/2/3 para ratón (Mishra, Grzybek, Niki, *et al.*, 2010). En el análisis realizado en la posible región promotora encontramos cajas TATA, las cuales son reconocidas por la subunidad TBP del complejo TFIID, otro elemento es el sitio Inr, el cual contiene secuencias ricas en pirimidinas, estos son factores esenciales para la unión de la maquinaria basal de la transcripción (Smale & Kadonaga, 2003), estos sitios nos proporcionan información de que este promotor se activa de forma clásica, es decir por medio de cajas TATA y sitios Inr. Analizando la posible región promotora del gen *LGALS9* se encontraron sitios de unión para AP2, AP1 y GATA 1 como los hallados en ratón y rata (Mishra, Grzybek, Niki, *et al.*, 2010).

También fueron localizados un par de sitios que reconoce c-Jun, el cual se ha reportado como un activador de la transcripción y que Steelman, Smith, Welsh y Li en 2013 hallaron en astrocitos.

Al analizar los diferentes fragmentos de la región promotora se observa una menor actividad en los fragmentos de 2089, 1479 y 222 pb comparados con el fragmento de 727 pb. Dentro de los fragmentos de 2089 y 1479 pb se encuentran 4 sitios de unión a CREM-tau el cual se ha observado ser un represor en la actividad promotora, la actividad represora de CREM-tau se ha observado en testículos de ratón (Nantel, Monaco, Foulkes, *et al.* 1996) y humano (Bher & Wenbauer. 2000) observando un efecto de arresto del desarrollo de las espermatogonias,

en humano se ha identificado que CREM-tau actúa como un represor (Bher and Wenbauer. 2000).

Al observar una actividad menor en los fragmentos de 2089 y 1479 pb se realizó una nueva consulta en el programa EPD, se encontró que la base de datos de dicho programa había sido actualizada ya que solo proporciono una solo posible región promotora, -466 y +134 pb, la cual está presente en el fragmento de 727 pb. En este fragmento de 727 pb ya no contiene estos sitios de unión, lo que podría estar relacionado con el aumento de la actividad en el fragmento, manteniéndose los sitios de unión a TFIID, TBP, Sp1, Inr, AP1. completar

La disminución de la actividad en este fragmento. de 222 pb, en relación al fragmento de 727 pb, puede ser atribuida a la eliminación de dos sitios de unión de TFIID y TBP, dejando solo un sitio de unión a TFIID y TBP, además de un sitio Inr junto con el sitio de inicio de la transcripción, por lo cual a esta región la podremos llamar el promotor central del gen *LGALS9* ya que en esta región (222 pb) se observa una disminución de la luciferasa en relación al fragmento de 727 pb.

Conclusiones

Por medio de análisis *in silico* se encontró, al inicio de la investigación, se encontraron dos posibles regiones promotoras pero la realizar una segunda revisión en el programa EPD solo se encontró una sola región promotora del gen *LGALS9*, que comprende una extensión de 600 pb, la cual se encuentra ubicada entre las posiciones -426 y +162 respecto al sitio de inicio de la transcripción y que por medio de ensayos de luciferasa se comprobó que tiene actividad promotora, proponiendo que esta región es el promotor del gen *LGALS9*.

Esta región promotora presenta cajas TATA, así como también sitios Inr, por lo que se propone que es un promotor clásico.

El fragmento de 222 pb mantiene una actividad menor por lo que se propone como el promotor mínimo del gen *LGALS9*, sin embargo, es necesario realizar construcciones más pequeñas para poder hallar la región en la cual la actividad aún se conserva.

Perspectivas

Realizar ensayos de inmunoprecipitación de cromatina (ChIP) con diferentes factores de unión para confirmar los sitios de unión identificados con el programa Patch 1.0.

Realizar la construcción de regiones más cortas del promotor para poder encontrar el promotor mínimo que mantenga actividad.

Mutar los sitios de unión a TFIID, TBP, Inr, Sp1 y AP1 para observar si algún otro sitio de unión puede mantener la actividad transcripcional del promotor en ausencia de ellos.

Referencias

1. Alam, S., Li, H., Margariti, A. *et al.* (2011). Galectin-9 protein expression in endothelial cells is positively regulated by histone deacetylase 3. *The journal of biological chemistry.* 51, 44211-44217.
2. Ahmed, H. and Bandyopadhyaya, G. B.(2015). Chapter 4: Examination of the reStowell, S. R. and Cummings, R. D.
3. Barondes, S. H., Castronovo, V., Cooper, D. N. *et al.* (1994). Galectins: matricellular glycan-binding proteins linking cell adhesion, migration, and survival. *Galectins: A family of animal beta-galactoside-binding lectins. Cell.* 76, 597-598.
4. Bher, R. and Weinbauer, G. F. (2000). CREM activator and repressor isoforms in human testis: sequence variations and inaccurate splicing during impaired spermatogenesis. *Molecular Human Reproduction.* 11, 967-972.
5. Brinchman, F. B., Patel, D. M. and Iversen, M. H. (2018). The Role of Galectins as Modulators of Metabolism and Inflammation. *Hindawi Mediators of Inflammation.* 1, 1-11.
6. Caporal, L. (2019). Tesis: Análisis de la expresión de glicogenes modulados por la oncoproteína E6 del virus de papiloma humano de alto y bajo riesgo.
7. Chalkley, G. E. and Verrijzer, C. P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAFII250-TAFII150 complex recognizes the initiator. *EMBO Journal.* 18, 4835-4845.
8. Chiariotti, L., Salvatore, P., Frunzio, R. and Bruni, C. B. (2004). Galectin genes: Regulation of expression. *Glycoconjugate Journal.* 19, 441-449.
9. De Gregorio, E., Chiariotti, L. and Di Nocera P. P. (2001). The overlap of Inr and TATA elements sets the use of alternative transcriptional start sites in the mouse galectin-1 gene promoter. *Gene.* 268, 215-223.
10. Dreos, R., Ambrosini, G., Cavin Périer, R. and Bucher, P. (2013). EPD and EPDnew, high-quality promoter resources in the next-generation sequencing era. *Nucleic Acids Research.* 41, 157-164.
11. Dyer, K. D. and Rosenberg, H. F. (2001). Transcriptional regulation of galectin-10 (eosinophil Charcot-Leyden crystal protein): A GC box (-44 to -50) controls butyric acid induction of gene expression. *Life Science Journal.* 69, 201-212.

12. Elola, M. T., Wolfenstein-Todel, C., Troncoso, M. F., Vasta, G. R. and Rabinovich, G. A. (2007). Cellular and Molecular Life Sciences. 64, 1679-1700.
13. Fujita, K., Iwama, H., Oura, K. *et al.* (2017). Cancer therapy due to apoptosis: galectin-9. International journal of molecular sciences. 18, 1-15.
14. Gitt, M. A., Jordan, E. T. and Leffler H. Galectin-2, Galectin-5 and -9, and galectins-4 and -6. (1997) Trends Glycosci Glycotech. 9, 87-93.
15. Gitt, M. A., Xia, Y. R., Atchison, R. E., Lusic, A. J., Baromdes, S. H. and Leffler H. (1998). Sequence, structure and chromosomal mapping of the mouse Lgals6 gene, encoding galectin-6. The Journal of Biological Chemistry. 273, 2961-2970.
16. Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nature Structural & Molecular Biology. 11, 394-403.
17. Heusschen, R., Griffioen, A. W. and Thijssen, V. L. (2013). Galectin-9 in tumor biology: a jack of multiple trades. Biochimica et Biophysica acta. 1836, 177-185.
18. Hirabayashi, J., Hashidate, W. E., *et al.* (2002). Oligosaccharide specificity of galectins: A search by frontal affinity chromatography. Biochimica et Biophysica Acta. 1572, 232-254.
19. Houzelstein, D., Goncalves, I. R., Orth, A. *et al.* (2004). Phylogenetic analysis of the vertebrate galectin family. Molecular Biology and Evolution. 21, 1177-1187.
20. Hughes, R. C. (2001). Galectins as modulators of cell adhesion. Biochimie. 83, 667-670.
21. Irie, A., Yamauchi, A., Kontani, K., Kihara, D. *et al.* (2005). Galectin-9 as a prognostic factor with antimetastatic potential in breast cancer. Clinical Cancer Research. 11, 2962-2968.
22. Kadrofske, M. M., Openo, K. P. and Wang, J. L. (1997). The human LGALS3 (Galectin-3) gene: determination of the gene structure and functional characterization of the promoter. Archives of Biochemistry and Biophysics. 349, 7-20.
23. Kouzine, F., Levens, D. and Baranello, L. (2014) DNA and transcription. Nucleus. 5:3, 195-202.
24. Kutach, A. K. and Kadonaga, J. T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. Molecular Cell Biology. 20(13) 4757-4764.

25. Lahm, H., Andre, S., Hoeflich, A., Fischer, J.R. (2001). Comprehensive galectin fingerprinting in a panel of 61 human tumor cell lines by RT-PCR and its implications for diagnostic and therapeutic procedures. *Journal Cancer Research Clinical*. 127, 375-386.
26. Leffler, H., Carlsson, S., Hedlund, M., Qian, Y. & Poirier, F. (2004). Introduction to galectins. *Glycoconjugate Journal*. 19, 433-440.
27. Liang, M., Ueno, M., Oomizu, S., Arikawa, T. (2008). Galectin-9 expression links to malignant potential of cervical squamous cell carcinoma. *Journal of Cancer Research and Clinical Oncology*. 134 (8), 899-907.
28. Liu, F. T. and Rabinovich, G. A. (2010). Galectins: regulators of acute and chronic inflammation. *Annals of the New York Academy of Sciences*. 1183, 158-182
29. Lutterbach, B. and Hiebert, S. W. (2000). Role of the transcription factor AML-1 in acute leukemia and hematopoietic differentiation. *Gene*. 245, 223-235.
30. Ma, C. J., Li, G.Y., Cheng, Y. Q. *et al.* (2013). *Cis* Association of galectin-9 with Tim-3 differentially regulates IL-12/IL-23 expressions in monocytes via TLR signaling. *Plos One*. 8, 1-15.
31. Malik, S. and Roeder, R.G. (2005). Dynamic regulation of pol transcription by the mammalian mediator complex. *Trends in Biochemical Sciences*. 30, 256-263.
32. Maston, G. A., Evans, S. K. and Green, M. R. (2006). Transcriptional Regulatory Elements in the Human. *Annual Review of Genomics and Human Genetics*. 7, 29-59.
33. Matsuura, A., Tsukada, J., Mizobe, T., et al. (2009). Intracellular galectin-9 activates inflammatory cytokines in monocytes. *Genes to Cells*. 14, 511-521.
34. Mishra, R., Grzybek, M., Niki, T., Hirashima, M. and Simons, K. (2010). Galectin-9 trafficking regulates apical-basal polarity in Madin-Darby canine kidney epithelial cells. *Proceedings National Academy of Sciences USA*. 107, 17633-17638.
35. Nakamura, S. and Raz, A. (2007). Regulation of cancer-related gene expression by galectin-3 and the molecular mechanism of its nuclear import pathway. *Cancer Metastasis Review*. 26, 605-610.

36. Nantel, F., Monaco, L., Foulkes, N. S. *et al.* (1996). Spermiogenesis deficiency and germ-cell apoptosis in CREM-mutant mice. *Nature*. 380, 159-162.
37. O'Connor, L., Gilmour, J. and Bonifer C. (2016). The role of the ubiquitously expressed transcription factor Sp1 in tissue-specific transcriptional regulation and in disease. *Yale Journal of Biology and Medicine*. 89, 513-525.
38. Ogden, A. T., Nunes, I., Ko, K., Wu, S., Hines, C. S., Wang, A. F. Hegde, R. S. and Lang, R. A. (1998). GRIFIN, a novel lems-specific protein related to the galectin family. *The Journal of biological chemistry*. 273, 28889-28896.
39. Orphanides, G., Lagrange, T. and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & Development*. 10, 2657-2683.
40. Park W. S., Jung, W-K., Park, S. K. *et al.* (2011). Expression of galectin-9 by IFN-g stimulated human nasal polyp fibroblast through MAPK, P13K, and JAK/STAT signaling pathways.
41. Rabinovich, G. A., Baum, L. G., Tinari, N. *et al.* (2002). Galectins and their ligands: Amplifiers, silencers or turners of the inflammatory response?. *Trends in Immunology*. 23, 313-320. *Biochemical and Biophysic Research Communications*. 411, 259-264.
42. Steelman, A. J., Smith III R., Welsh, C. J. and Li, J. (2013). Galectina-9 protein is up-regulated in astrocytes by tumor necrosis factor and promotes encephalitogenic T-cell apoptosis. *The Journal of biological chemistry*. 32 23776-23787.
43. Smale, S. T. and Kadonaga, J. T. (2003). The RNA polymerase II core promoter. *Annual Review of Biochemistry*. 72, 449-479.
44. Tremblay, M., Sanchez-Ferras, O. and Bouchard, M. (2018). GATA transcription factors in development and disease. *The company of biologists*. 145, 1-20.
45. Wu, C., Thalhamer, T., Franca, R. F. *et al.* (2014). Galectin-9-CD44 interaction enhances stability and function of adaptive regulatory T cells. *National Institute of Health*. 41, 270-282.
46. Yang, R. Y., Rabinovich, G. A. and Liu, F. T. (2008). Galectins: structure, function and therapeutic potential. *Experts Review in Molecular Medicine*. 10, 1-24.
47. Yella, V. R. and Bansal, M. (2017). DNA structural features of eukaryotic TATA-containing and TATA-less promoters. *FEBS Open Bio*. 7(3), 324-334.

48. Zhu, C., Anderson, A. C., Schubart, A. *et al.* (2005). The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nature Immunology*. 6, 1245-1252.

Anexos



5^o Congreso Latinoamericano de Glicobiología
Sociedad Latinoamericana de Glicobiología A.C.



Centro de Investigación en Dinámica Celular



The Latin American Society of Glycobiology and the Cell Dynamics Research Center of the Morelos State Autonomous University

CERTIFICATE OF ATTENDANCE

We hereby certify that:

Jesús Aparicio Flores

has attended the 5th Latin American Congress of Glycobiology and 2nd Meeting of the Glycoscience in Health Thematic Network held on October 2-4, 2019 in Mexico City, Mexico.

Iván Martínez Duncker R.
President of the Latin American Society for Glycobiology
Head of the Glycoscience in Health Thematic Network CONACYT

Carla Asteggiano
Organizing committee

Edgar Centeno Galindo
Organizing committee

Héctor Mora Montes
Organizing committee

Roberta Salinas Marín
Organizing committee

Tonatiuh Ramírez Reivich
Organizing committee

Verónica Vallejo Ruiz
Organizing committee





Centro de
Investigación en
Dinámica Celular



The Latin American Society of Glycobiology and the Cell Dynamics
Research Center of the Morelos State Autonomous University

CERTIFICATE OF ATTENDANCE

We hereby certify that:

Aparicio Flores J., Milflores Flores L., Vallejo Ruiz V.

Presented the poster:

Identification, characterization and evaluation of the LGALS9 gene promoter.

At the 5th Latin American Congress of Glycobiology and 2nd Meeting
of the Glycoscience in Health Thematic Network held on October 2-4,
2019 in Mexico City, Mexico.

Iván Martínez Duncker R.
President of the Latin American
Society for Glycobiology
Head of the Glycoscience in Health Thematic
Network CONACYT

Carla Asteggiano
Organizing committee

Edgar Centeno Galindo
Organizing committee

Héctor Mora Montes
Organizing committee

Roberta Salinas Marin
Organizing committee

Tonatiuh Ramírez Revitch
Organizing committee

Verónica Vallejo Ruiz
Organizing committee

