



FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

**EVALUACIÓN DEL POTENCIAL DE LA SECUENCIA SEÑAL DEL FACTOR VIII
PARA SECRETAR PROTEÍNAS RECOMBINANTES EN LÍNEAS CELULARES DE
MAMÍFEROS**

Línea de investigación:

Genética, bioquímica y biotecnología

Tesis para optar el título profesional de Licenciado en Biología

Autor:

Quispe Pomatay, Luis Alberto

Asesor:

Marcelo Rodríguez, Álvaro Julián

(ORCID: 0000-0003-2031-6806)

Jurado:

Rodrigo Rojas, Maria Elena

Riveros Ramirez, Maribel Denise

Mayanga Herrera, Ana Lucia

Lima - Perú

2024



EVALUACIÓN DEL POTENCIAL DE LA SECUENCIA SEÑAL DEL FACTOR VIII PARA SECRETAR PROTEÍNAS RECOMBINANTES EN LÍNEAS CELULARES DE MAMÍFEROS

INFORME DE ORIGINALIDAD

13%

INDICE DE SIMILITUD

13%

FUENTES DE INTERNET

6%

PUBLICACIONES

0%

TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	www.teses.usp.br Fuente de Internet	4%
2	patents.google.com Fuente de Internet	3%
3	hdl.handle.net Fuente de Internet	2%
4	bibliotecadigital.exactas.uba.ar Fuente de Internet	<1%
5	repositorio.unfv.edu.pe Fuente de Internet	<1%
6	idoc.pub Fuente de Internet	<1%
7	pesquisa.bvsalud.org Fuente de Internet	<1%
8	Mateo Lozano, Silvia. "Sarcoma de Ewing : nuevas aproximaciones terapéuticas y búsqueda de dianas biológicas del oncogén"	<1%



FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

**EVALUACIÓN DEL POTENCIAL DE LA SECUENCIA SEÑAL DEL FACTOR VIII
PARA SECRETAR PROTEÍNAS RECOMBINANTES EN LÍNEAS CELULARES DE
MAMÍFEROS**

Línea de investigación:

Genética, bioquímica y biotecnología

Tesis para optar el Título Profesional de Licenciado en Biología

Autor:

Quispe Pomatay, Luis Alberto

Asesor:

Marcelo Rodríguez, Álvaro Julián

(ORCID: 0000-0003-2031-6806)

Jurado:

Rodrigo Rojas, Maria Elena

Riveros Ramirez, Maribel Denise

Mayanga Herrera, Ana Lucia

**LIMA-PERU
2024**

AGRADECIMIENTOS

Me gustaría aprovechar estas líneas para poder agradecer a las diferentes personas que me brindaron todo tipo de apoyo durante el desarrollo de este trabajo.

A mi amada madre, Alejandra, y a mi amado padre, Pedro, por alentar siempre mi amor por el conocimiento y la investigación, sin su ayuda incondicional, no hubiera llegado tan lejos. A Mis hermanos Alfredo, Lidia y Magdalena, que al igual que mis padres, siempre han estado conmigo, ayudándome de diferentes formas durante mis 5 años en la universidad, de poco sirve tratar de expresarles la magnitud de mi agradecimiento.

Al PhD. Álvaro Marcelo, mi asesor de tesis, quien me incentivo y me ha ayudado para el culmino de este trabajo. Junto con los demás profesores de la universidad han hecho posible el desarrollo de mi ansia por el conocimiento.

A los que han permitido el desarrollo y financiamiento de este trabajo. A la PhD. Mari Cleide Sogayar, al PhD. Marcos Almeida Demasi y todo el grupo NUCEL, quienes con su dirección, conocimiento y docencia han permitido el éxito en el desarrollo de este trabajo.

A la Universidad de São Paulo y a las agencias de financiación del Brasil, CNPq y FAPESP

A cada persona q no mencione pero que de alguna forma u otra me ha ayudado para el desarrollo de este trabajo y mi formación como biólogo. Muchas gracias

INDICE

RESUMEN.....	1
ABSTRACT.....	2
I. INTRODUCCIÓN.....	3
1.1. Descripción y formulación del problema	5
1.2. Antecedentes	8
1.3. Objetivos	12
□ Objetivo general	12
□ Objetivos específicos	12
1.4. Justificación	13
1.5. Hipótesis	14
II. MARCO TEÓRICO	15
2.1. Bases teóricas sobre el tema de investigación	15
2.1.1. Proteínas recombinantes.....	15
2.1.2. Proteínas fluorescentes	15
2.1.3. Proteína amarilla fluorescente	16
2.1.4. Células CHO-DG44.....	16
2.1.5. Células HEK2933T.....	17
2.1.6. Secuencias señal.....	18
2.1.7. Factor VIII de coagulación humana.....	18
2.1.8. Regulación genética en eucariontes	19

2.1.9.	Transfección.....	20
III.	MÉTODO.....	22
3.1.	Tipo de investigación.....	22
3.2.	Ámbito temporal y espacial.....	22
3.3.	Variables	22
3.4.	Población y muestra	22
3.5.	Instrumentos	23
3.6.	Procedimientos	23
3.6.1.	Línea celular y manutención	23
3.6.2.	Construcción del vector de expresión de una variante secretable de la proteína YFP (pLV/spYFP_DHFR)	24
3.6.3.	Transformación de bacterias <i>Escherichia coli</i> , cepa XL1 con el vector pLV/spYFP_IRES_DHFR	26
3.6.4.	Caracterización del vector de expresión de una variante secretable de la proteína EFYP (pLV/spYFP_IRES_DHFR).....	27
3.6.5.	Transfección transitoria en las células HEK 293T con el vector pLV/spYFP_DHFR.....	28
3.6.6.	Transfección estable de las células CHO-DG44 con el vector pLV/spYFP_DHFR.....	29
3.6.7.	Evaluación de los niveles de secreción de la proteína YFP en las poblaciones celulares mixtas.....	30
3.6.8.	Análisis de datos	31

IV.	RESULTADOS.....	32
4.1.	Construcción del vector de expresión de una variante secretada de la proteína YFP.....	32
4.2.	Caracterización del vector de expresión de una variante secretada de la proteína YFP (pLV/spYFP_DHFR)	34
4.3.	Generación de las poblaciones celulares mixta a través del metodo de transfección transitoria	38
4.4.	Generación de las células CHO-DG44 con expresión estable de la proteína YFP.....	41
4.5.	Evaluación de los niveles de secreción de la proteína YFP en las poblaciones celulares mixtas generadas por transfección transitoria y estable.....	43
V.	DISCUSIÓN DE LOS RESULTADOS.....	48
VI.	CONCLUSIÓN.....	55
VII.	RECOMENDACIONES.....	56
VIII.	REFERENCIAS	57
IX.	ANEXOS	71
9.1.	Anexos 1	71
9.2.	Anexos 2	82

LISTA DE FIGURAS

Figura 1: Línea celular derivada de ovario de hámster chino (CHO DG44) observadas al microscopio compuesto con un aumento total de 200X.	17
Figura 2: Línea celular derivada de riñón embrionario humano (HEK 293T) observadas al microscopio compuesto con un aumento total de 50X.	17
Figura 3: Representación esquemática de los principales elementos del vector lentiviral bicistrónico de expresión pLV/spYFP_DHFR.	32
Figura 4: Análisis por electroforesis en gel de agarosa al 1% de los fragmentos de DNA	34
Figura 5: Selección de 11 colonias bacterianas recombinantes por PCR obtenidas después de la transformación de las bacterias E. coli (cepa XL-1 Blue) con el producto de ligación pLV/spYFP_DHFR.....	35
Figura 6: Análisis por electroforesis en gel de agarosa 1% (m/v) del perfil de restricción con las enzimas indicadas del clon CR5 del vector recombinante pLV/spYFP_DHFR.	36
Figura 7: Representación esquemática del alineamiento realizado a través de la herramienta BLAST de las secuencias consenso de los clones recombinantes obtenidas por secuenciación por el método de Sanger, con la secuencia teórica del elemento spYFP_IRES_DHFR.....	37
Figura 8: Evaluación de la expresión transitoria de la proteína spYFP en la línea celular HEK 293T por microscopia de fluorescencia.	39
Figura 9: Evaluación de la secreción de la proteína YFP (~27KDa) en la línea celular HEK293T por SDS-PAGE 10% en condiciones no reductoras, seguida por detección de fluorescencia por escaneamiento a laser.....	40

Figura 10: Evaluación de la expresión estable de la proteína YFP en las células CHO-DG44 por microscopia de fluorescencia.....	42
Figura 11: Secreción de la proteína YFP (~27KDa) en las células CHODG44 evaluadas por SDS-PAGE 10% en condiciones no reductoras, seguida por detección de fluorescencia por escaneamiento a laser.....	43
Figura 12: Evaluación de los niveles de secreción de la proteína YFP de las poblaciones celulares mixtas generadas por transfección transitoria en las células HEK 293T.....	45
Figura 13: Evaluación de los niveles de secreción de la proteína YFP de las poblaciones celulares mixtas generadas por transfección estable en la línea celular CHO-DG44.....	47
Figura 14: Alineamiento del ensamblaje de la secuencia del clon bacteriano CR1 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST.....	71
Figura 15: Alineamiento del ensamblaje de la secuencia del clon bacteriano CR5 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST.....	73
Figura 16: Alineamiento del ensamblaje de la secuencia del clon bacteriano CR20 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST.....	74
Figura 17: Alineamiento del ensamblaje de la secuencia del clon bacteriano CR36 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST.....	76
Figura 18: Alineamiento del ensamblaje de la secuencia del clon bacteriano CR60 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST.....	78
Figura 19: Alineamiento del ensamblaje de la secuencia del clon bacteriano CR82 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST.....	80

LISTA DE TABLAS

Tabla 1: Evaluación de los niveles de secreción de la proteína EYFP en las poblaciones celulares mixtas generadas por transfección transitoria en las células HEK293T..... 82

Tabla 2: Evaluación de los niveles de secreción de la proteína EYFP en las poblaciones celulares mixtas generadas por transfección estable en las células CHO-DG44..... 84

RESUMEN

Esta tesis determino la capacidad de secreción de la secuencia señal del factor VIII de coagulación humana para secretar proteínas heterólogas en las líneas celulares de mamíferos. Inicialmente, se generó por PCR, una variante secretada de la proteína amarilla fluorescente (YFP) a través de la fusión de la secuencia señal del factor VIII a la región codificante de esta proteína, formando así, el vector pLV/spEYFP_IRES_DHFR. En cada uno de los clones 6 generados de este vector, fue confirmado su fidelidad y similaridad a través de secuenciación, de los cuales, el clon CR5 fue seleccionado por presentar la mayor similaridad con la secuencia teórica. Para evaluar la capacidad de secreción, fueron usados los vectores que contenían una variante secretada de la proteína YFP modificado por la secuencia señal de la interleucina-2 y la albumina, caracterizados por su alta capacidad de secreción. Junto al clon CR5, estos vectores fueron usados para generar las poblaciones celulares mixta por transfección transitoria y estable, en las líneas celulares HEK293T y CHODG44, respectivamente. A partir de electroforesis SDS-PAGE y por microscopia de fluorescencia fue posible la detección de la proteína YFP en el medio condicionado. En los ensayos de los niveles de secreción de la proteína YFP por fluorometria en todas las poblaciones celulares generadas, no se encontró ninguna diferencia estadística entre el clon CR5 con los otros vectores evaluados. Demostrando así, que la secuencia señal del factor VIII presenta una capacidad de secreción similar al de esas secuencias señal ampliamente usadas en procesos biotecnológicos.

Palabras claves: Proteínas recombinantes, células HEK, células CHO, Secuencia señal, Factor VIII

ABSTRACT

This thesis determined the secretion capacity of the signal sequence derived from light chain of human coagulation factor VIII to secrete heterologous proteins in mammalian cell lines. Initially, a secretable variant of the yellow fluorescent protein (YFP) fused to the signal sequence of factor VIII was generated by PCR, thus forming the vector pLV/spEYFP_IRES_DHFR. I identified 6 candidate clones, in each of the clones generated from this vector, their fidelity and similarity were confirmed through sequencing, of which clone CR5 was selected for presenting the greatest identity with the theory sequence. To evaluate the secretion capacity, vectors containing a secretable variant of the YFP protein modified by the signal sequence of interleukin-2 and albumin were used, characterized by their high secretion capacity. Together with the CR5 clone, these vectors were used to generate mixed cell populations by transient and stable transfection, in the HEK293T and CHODG44 cell lines, respectively. Trough SDS-PAGE electrophoresis and fluorescence microscopy, the presence of YFP protein in the condition medium was detected. In the assays of the secretion levels of the YFP protein by fluorometry in both the mixed cell populations generated by transient and stable transfection, no statistical difference was found between the CR5 clone with the other vectors evaluated. Thus, demonstrating that the signal sequence of factor VIII has a secretion capacity similar to that of those signal sequences widely used in biotechnological processes.

Keywords: Recombinant proteins, HEK cells, CHO cells, Signal sequence, Factor VIII

I. INTRODUCCIÓN

Las proteínas recombinantes terapéuticas o biofármacos constituyen un grupo heterogéneo de moléculas con actividad terapéutica, aislados a partir de organismos vivos, ya sea directamente o mediante producción biotecnológica (Food and Drug Administration[FDA], 2020).

Los primeros biofármacos fueron autorizados en la década del 80, surgiendo así, una nueva alternativa para el tratamiento de diferentes enfermedades(Kim *et al.*, 2012). Estas nuevas terapias demostraron una alta potencia y especificidad en comparación con el tratamiento de drogas de pequeñas moléculas (Altamirano *et al.*, 2013), lo que aceleró su expansión e impulsó el desarrollo de la industria biofarmacéutica, llegando a alcanzar ventas de \$224 billones en 2019 (La Merie, 2020). Ante el constante crecimiento y expansión de esta industria, surge un mayor interés por el aumento gradual en los niveles de producción de los biofármacos (Wurm y Wurm, 2017).

El desarrollo de una plataforma de producción de un biofármaco comienza a nivel genético, donde la secuencia codificante de la proteína de interés se aísla y se clona a un vector de expresión, que presenta elementos como, la secuencia del promotor, sitio de unión al ribosoma, codón terminador y marcadores de selección, ya que el ADN clonado carece de estas regiones reguladoras, y posteriormente, es insertado en una célula huésped para su expresión a través de un método de transferencia genómica aleatoria (Mathew, 2022).

Inicialmente, las células procariotas fueron usadas como plataforma de expresión para la producción biotecnológica de estas proteínas, debido a características como, su alta tasa de crecimiento, bajo costo de manutención y fácil manipulación genética(Stern *et al.*, 2007) . Sin embargo, estas células presentaron algunas desventajas al producir proteínas de mamíferos y su uso limitado para ser usado con fines terapéuticos en humanos (Knappskog *et al.*, 2007).

Ante estas características, las células de mamíferos surgieron como una opción viable por poder realizar las modificaciones postraduccionales necesarias para el producto objetivo, por presentar proteínas accesorias necesarias para un óptimo plegamiento de la proteína y una poderosa maquinaria que permite la producción de proteínas solubles (Román *et al.*, 2016). De todas las líneas celulares de mamíferos, las más utilizadas son aquellas células derivadas de ovario de hámster chino (*Chinese Hamster Ovary*, CHO), mieloma murino (NSO), riñón de hámster bebe (*Baby Hamster Kidney*, BHK), riñón embrionario humano (*Human Embryo Kidney*, HEK) y de retina humana (PER.C6) (Wurm, 2004).

En las líneas celulares de mamíferos es posible producir las proteínas recombinantes a través de sistema de expresión transitorio o estable (Baser *et al.*, 2015). La diferencia entre estos dos sistemas radica en que en el sistema de expresión transitorio es posible obtener con rapidez una gran cantidad del producto, pero por un corto periodo de tiempo. Por el contrario, el sistema de expresión estable expresa de forma constitutiva la proteína, permitiendo obtener un producto reproducible, homogéneo y con capacidad de escalonamiento en grandes volúmenes, solo que en menor proporción que con que el sistema de expresión transitorio. Hoy en día, ambos sistemas se pueden desarrollar en las líneas celulares HEK293 y CHO adaptadas por suspensión (Bandaranayake y Almo, 2014).

Los avances provenientes de estudios “omicos” y del campo de la ingeniería de células de mamíferos han permitido la optimización de los procesos de creación de líneas celulares de mamífero a través de múltiples modificaciones genéticas para aumentar la estabilidad y los niveles de expresión de proteínas recombinantes (Kuystermans *et al.*, 2007). De la misma forma, la aplicación de herramientas de biología de sistemas se ha mostrado muy relevante para la comprensión de las complejas alteraciones fenotípicas sufridas por las células de mamíferos durante el proceso de obtención de clones superproductores (Birch y Onakunle, 2005).

1.1. Descripción y formulación del problema

Para la producción biotecnológica de las proteínas recombinantes se han utilizado diferentes sistemas de expresión, tanto procariotas como eucariotas (Ferrer-Miralles *et al.*, 2009). Las bacterias usadas como plataforma de expresión, presentaron una serie de desventajas asociadas a deficiencias en su maquinaria celular, como su incapacidad de realizar modificaciones postraduccionales, un plegamiento incorrecto de las proteínas y a su agregación en cuerpos de inclusión, lo que dificulta su posterior purificación. (Knappskog *et al.*, 2007). Además, su uso con fines terapéutico en humanos está restringido, debido a que muchas bacterias tienen un alto contenido de endotoxinas, que pueden ser de naturaleza patogénica (Stern *et al.*, 2007).

Antes estos problemas, las células de mamíferos se tornaron una buena alternativa por ser capaces de realizar las modificaciones postraduccionales necesarias para el producto objetivo como, por ejemplo, un patrón de glicosilación similar al de células humanas, presentar las chaperonas necesarias para un adecuado plegamiento de las proteínas y una poderosa maquinaria de secreción, que evita el uso de técnicas costosas de ruptura y purificación celular (Román *et al.*, 2016; Walsh, 2018). Sin embargo, el cultivo de estas células es complicado en comparación con las otras plataformas de expresión, porque presentar una tasa de crecimiento lenta, requieren un medio de cultivo caro y una capacidad de producción escalonada limitada (Attallah *et al.*, 2017). Aun presentando estas desventajas, la necesidad de obtener proteínas recombinantes completamente funcionales y terapéuticamente seguras, torna su uso prioritario (Stern *et al.* 2007).

De las líneas celulares de mamíferos, las más utilizadas en la industria biofarmacéuticas son: la línea celular proveniente de riñón embrionario humano, la línea celular originada de células de mieloma murino y la línea celular derivada de células de ovarios de hámster chino,

y sus líneas celulares mutantes derivadas a partir de este linaje, deficiente para la expresión de la enzima dihidrofolato reductasa (DHFR) (Hacker y Balasubramanian, 2016).

Comúnmente, el desarrollo de una plataforma de producción de proteínas recombinantes usando células de mamíferos, requiere la modificación genética de una línea celular a través de la inserción en su genoma de vectores de expresión conteniendo el gen recombinante de interés, asociado a elementos regulatorios de la expresión génica, y marcadores de selección apropiados para la obtención de clones celulares con una productividad elevada y estable de la proteína recombinante de interés (Derouazi *et al.*, 2006). La eficiencia en cada una de estas etapas define los niveles de expresión celular de una determinada proteína (Lewis *et al.*, 2013). Eventos como la integración genómica aleatoria del vector de expresión del gen de interés obtenida a través de esta metodología, evita predecir el sitio cromosómico de inserción, el número de copias integradas, y el nivel de expresión de las células generadas (Bushman, 2007). Adicionalmente, el empleo de esta metodología está asociado a una baja probabilidad de generar clones celulares altamente productores, siendo necesario la utilización de diversas alternativas para superar estas desventajas (Kim *et al.*, 2012).

El procedimiento de producción de biofármacos usando células de mamíferos está regulado por el procesamiento de las proteínas, que inicia con la transcripción, seguido de la traducción, el procesamiento postraducciona l y finalmente la secreción (Stern *et al.* 2007). Esta última etapa, está conformado por factores que regulan el flujo de transporte de las proteínas por los varios estadios de la vía secretoria y que pueden impactar en el control general de la producción, trayendo consigo una reducción, pérdida y modificación de la proteína producida y, por consiguiente, una gran pérdida económica (Harter y Wieland, 1996). Esos factores pueden incluir los procesamientos efectuados en el retículo endoplasmático (ER) y en el complejo de Golgi, los cuales desempeñan una función clave en la segregación de las proteínas,

en los procesos de glicosilación y otras modificaciones postraduccionales (Gutierrez *et al.*, 2020).

Por otra parte, como los promotores y las secuencias señal suelen formar parte del casete de expresión que es insertado en la célula hospedera, su elección correcta tiene un gran impacto en el rendimiento de la producción final de la proteína recombinante (Dalton y Barton, 2014). Las secuencias señal en procariotas y eucariotas comparten una arquitectura general común, por lo cual es posible usar una con otro (Driessen y Van der Does, 2002). Sin embargo, no todas las secuencias señal funcionan de manera eficiente en otros organismos (Kojima y Kitamura, 1999). Por consiguiente, es casi imposible predecir de antemano que secuencia señal funcionará mejor en el contexto de una proteína específica y una plataforma de expresión en particular (Dalton y Barton, 2014). Varios registros han demostrado una reducción drástica en los niveles de secreción de la proteína heteróloga al usarse una secuencia señal y célula diferente de la nativa (Stern *et al.* 2007). Por lo cual, la elección de la mejor secuencia señal para una proteína determinada es una de las etapas más críticas para una producción eficiente. Estos problemas, complican aún más el desarrollo rápido de una plataforma de producción de proteína recombinante terapéutica en células de mamíferos, debido al hecho de que las secuencias señal no solo afectan la translocación de las proteínas diana, sino que también influyen en la biosíntesis de las proteínas precursoras y en los eventos que ocurren en la célula (Freudl, 2018). Antes estas dificultades, surge la cuestión, ¿cuál será la secuencia señal que permita la secreción más eficiente de las proteínas recombinantes en las líneas celulares de mamíferos?

1.2. Antecedentes

Las secuencias señal de procariotas y eucariotas son funcionalmente intercambiables, incluso entre diferentes especies (Young y Rance, 2008). En base a estas observaciones, se han podido usar secuencias señal de diferentes especies para expresar un gen de interés en un tipo celular definido (Nagano y Masuda, 2014).

La determinación del rendimiento de una secuencia señal no puede necesariamente ser inferido a partir de las características de la proteína de origen, sino que requiere su evaluación empírica (Stern *et al.*, 2007). Siguiendo esta lógica, se ha intentado identificar secuencias señal potentes que se puedan fusionar con las proteínas recombinantes para mejorar su eficiencia de secreción (Kober *et al.*, 2013).

En el caso de las bacterias, el uso de secuencia señal ha sido ampliamente usado debido a su falta de capacidad de secretar las proteínas que sintetizan. Por ejemplo, a través del uso de la secuencia señal de la fosfatasa alcalina se logró secretar satisfactoriamente la proteína quimérica del citocromo b5 de hígado de rata en bacterias *E. coli* (Kaderbhai *et al.*, 2008).

De la misma forma, en las células eucarióticas ha sido evidenciado que el uso de secuencia señal de especies heterólogas han podido aumentar la producción a niveles comercialmente significativos (Negahdaripour *et al.*, 2017).

Por ejemplo, Stern *et al.* (2007) demostraron la importancia vital de la selección de la secuencia señal, cuando se pretende producir cantidades máximas de una proteína recombinante en una plataforma de expresión de células de mamífero. Ellos evaluaron en las células CHO el impacto en la síntesis y la secreción de la proteína luciferasa de *Gaussia sp.*, al reemplazar su secuencia señal nativa por las secuencias señal de interleucina-2, albúmina, tripsinógeno-2, quimotripsinógeno de mamífero, la secuencia señal de *Metridia sp.*, siete

secuencias señal derivados de *Oikopleura sp* y dos secuencias señal de procariontes. La secuencia señal nativa produjo altos niveles de síntesis y secreción de esta proteína, variando en un factor de más de 100 con respecto a las otras secuencias señal evaluadas y pudo operar igualmente en una línea celular de origen humano. Por lo tanto, es evidente que la secuencia señal de origen marino también se puede usar con éxito para la producción de proteínas de mamíferos. Siguiendo esta lógica, experimentaron con la secuencia señal de un procarionte, la secuencia señal de Slmj 1 de *Methanococcus jannaschii*, un Archaeobacterium que vive a niveles de aguas profundas en condiciones de alta presión, que fue considerado un candidato interesante para este estudio por su capacidad de secreción proteínas en estas condiciones tan extremas, sin embargo, el resultado fue todo lo contrario, mostrando una secreción ineficiente. Sorprendentemente, las secuencias señal de mamíferos, la interleucina 2 humana y la albúmina, demostraron ser muy inferiores a las secuencias señal derivado de organismos marinos.

De la misma forma, Kober *et al.* (2013), mediante una búsqueda extensa de la literatura, identificaron 16 secuencias señal prometedoras de diferentes organismos (como virus, bacterias y mamíferos), para aumentar la cantidad extracelular de un anticuerpo monoclonal expresado en células CHO-K1 y CHO-DG44. Los resultados mostraron que la mayoría de las secuencias señal derivadas de hongos o virus tuvieron una productividad específica muy baja. Por el contrario, la secuencia señal derivado de la albúmina humana mostró el mejor rendimiento de todas las secuencias señal analizados, aunque ya había sido descrito anteriormente como una secuencia señal débil en células CHO (Knappskog *et al.*, 2007). En conclusión, las secuencias señal de la Preproteína de albúmina sérica y de la preproteína azurocidina son adecuados para aumentar considerablemente la eficiencia de producción de una amplia gama de proteínas biofarmacéuticas, que incluyen tanto anticuerpos como también proteínas que no son anticuerpos, con tasas de producción claramente mejoradas, adecuadas para fines comerciales.

En el caso de Neill *et al.* (2023), usaron como plataforma de expresión las células CHO K1 y herramientas bioinformáticas predictivas basadas en modelos matemáticos de aprendizaje automático y profundo, para identificar y seleccionar secuencias señal que permitan la secreción óptima de una proteína específica. Para este objetivo, ellos diseñaron una biblioteca de 37 secuencias señal, que contenía 17 secuencias señal homólogas de CHO, 15 secuencias señal verificados experimentalmente y 5 secuencias señal diseñados sintéticamente. Las secuencias señal del grupo homólogo de CHO fueron seleccionadas por presentar una mayor abundancia de ARNm, suponiendo que genes altamente expresados deben contener secuencias señal con una alta tasa de translocación. Las secuencias señal extraídos de la literatura fueron seleccionados bajo el criterio de que estas moléculas tienen una mayor facilidad de expresar una proteína recombinante. Las formas sintéticas de las secuencias señal fueron creados a partir del análisis de factores tales como el tamaño promedio y la identificación de regiones de aminoácidos conservados de diferentes secuencias señal, verificados experimentalmente. La capacidad de secreción de estas secuencias señal se probaron en la producción de tres productos biofarmacéuticos industrialmente relevantes a través del modelo de regresión XGBoosting, el cual fue entrenado usando las variables del punto isoeléctrico, la estabilidad del dipéptido, la flexibilidad, el índice alifático, la energía libre de Gibbs, el promedio general del índice de hidropaticidad y el porcentaje de residuos de glicina y prolina en la secuencia señal. El modelo creado demostró que es funcional y que las construcciones sintéticas presentaron un rendimiento mayor de secreción que las secuencias evolucionadas naturalmente. Aunque, ninguna construcción facilitó un aumento en los títulos de las tres proteínas analizadas y a pesar de esta imprevisibilidad, el estudio demostró que es posible reducir significativamente el tiempo de prueba de vectores necesario para identificar combinaciones de secuencias señal que aumenten el título de una proteína recombinante específica.

Con respecto a la línea celular HEK 293, Wen *et al.* (2011), intentaron mejorar la secreción de las glicoproteínas del ectodominios E1 y E2 del virus de hepatitis C en esta misma línea celular, a través de la sustitución de su secuencia señal nativo por tres secuencias señal heterólogas. Las secuencias señal modificados del activador de plasminógeno tisular o luciferasa de *Gauss sp*, mostraron una expresión significativamente mayor que la secuencia señal nativa. Demostrando así, que el reemplazo *upstream* de las secuencias señal de los genes E1 y E2 aumentan los niveles de secreción y de expresión, y lo más importante, el reemplazo no afectó a la glicosilación. Esta estrategia también podría aplicarse a la expresión y producción de otras glicoproteínas en sistemas basados en líneas celulares de mamíferos.

De la misma forma, Güler-Gane *et al.* (2016), usaron las células HEK293 para superar el mal plegamiento y pérdida de la actividad del interferón α -2 humano cuando es expresado en bacterias (Babu *et al.*, 2000). La estrategia utilizada se basó en la modificación de diferentes secuencias señal heterólogas mediante el uso de aminoácidos espaciadores. Demostrando, que el impacto de tales modificaciones parece ser específico del contexto, ya que en algunos casos estas modificaciones mejoraron significativamente los niveles de expresión.

Por otra parte, Román *et al.* (2016), evaluaron el efecto de la combinación del promotor con la secuencia señal en la secreción de una proteína recombinante. Para esto, fueron evaluados los promotores del promotor temprano del citomegalovirus humano (CMV), del factor de elongación humano-1 y del potenciador temprano de CMV y de la β -actina de pollo con las secuencias señal de interleucina-2 humana, interferón α -2 humano, albúmina de suero humano, cadena pesada de IgG humana y luciferasa de *Gaussia princeps*, en las células HEK293, para determinar su capacidad de secretar la proteína verde fluorescente, usada como proteína modelo. Las secuencias señal de la albúmina y *G. princeps*, mostraron los más altos niveles de secreción, como se ha informado anteriormente en la literatura en otras líneas celulares (Kober *et al.* 2013). Estas secuencias señal junto con la del interferón α -2 humano

demonstraron ser buenos candidato para la secreción eficiente de otras proteínas heterólogas. Finalmente, la mejor combinación fue del promotor del CMV junto con el interferón α -2 humano para una secreción eficaz de la proteína verde fluorescente en células HEK 293, llegando hasta producir 2 veces más que con los otros promotores y secuencias señal.

1.3. Objetivos

- Objetivo general
 - Evaluar la capacidad de la secuencia señal de la cadena leve del factor VIII de coagulación humana para secretar proteínas recombinantes en las líneas celulares de mamíferos, CHO-DG44 y HEK293T

- Objetivos específicos
 - Construir y caracterizar el vector de expresión de una variante secretada de la proteína amarilla fluorescente(pLV/spYFP_DHFR).
 - Evaluar funcionalmente y determinar la secreción del vector de expresión por transfección transitoria en células HEK 293T con el vector pLV/spYFP_DHFR.
 - Determinar la secreción de la proteína amarilla fluorescente a través de la transfección estable en las células CHO-DG44 con el vector pLV/spYFP_DHFR.

1.4. Justificación

Las células de mamíferos se han utilizado ampliamente en la producción de proteínas recombinantes de importancia biomédica (Stern *et al.* 2007). La capacidad de realizar modificaciones postraduccionales complejas, poseer una poderosa maquinaria de secreción, y la presencia de proteínas accesorias necesarias para un adecuado plegamiento y formación de complejo proteicos, han hecho que el 60-70% de todas las proteínas recombinantes terapéuticas sean desarrolladas en estas células (Walsh, 2018).

En los últimos años, la diversidad de linajes y la optimización de bioprocesos en el cultivo de las células CHO ha permitido un aumento sustancial en la productividad de las proteínas terapéuticas, tornando así a las células CHO, la plataforma de expresión más utilizada para la producción biofarmacéutica, siendo responsable por el 70% de la producción de las proteínas recombinantes utilizadas comercialmente (Birch y Onakunle, 2005).

Para superar las dificultades asociados al desarrollo de una plataforma de producción de una proteína recombinante terapéutica usando células de mamíferos y las ineficiencias de cada una de las etapas del procesamiento de proteínas, se han ido desarrollado una serie de estrategias para acelerar y reducir los costos de producción (Kober *et al.*, 2012). En este sentido, se ha realizado la optimización de los parámetros de crecimiento, las condiciones de transfección y de los componentes de los medios de cultivo, la utilización de promotores/*enhancer* específicos de la transcripción génica asociado a niveles de expresión más elevados, el empleo de protocolos de amplificación génica y el uso de estrategias de ingeniería celular para la generación de células de mamíferos de alta producción (Lee *et al.*, 1996; Wurm, 2004; Bollin *et al.*, 2011).

Un enfoque nuevo y diferente, es la identificación de secuencia señal específicos que permitan un alto nivel de productividad de la proteína de interés. Esta nueva alternativa aún no

se ha explorado ampliamente en la producción de proteínas recombinantes usando células de mamíferos (Stern *et al.*, 2007). Entonces, el uso de secuencias señal de proteínas asociadas a una elevada secreción, es una idea atractiva para su uso en la producción de proteínas heterólogas con la finalidad de mejorar su rendimiento (Güler-Gane *et al.*, 2016). Sin embargo, sin un ensayo empírico no es posible determinar los niveles finales de la proteína secretada (Dalton y Barton, 2014). Por lo tanto, la forma más prometedora de encontrar una secuencia señal óptima para una proteína deseada, es probando la mayor diversidad posible de secuencias señal, la modificación aleatoria de su secuencia señal nativa o por el uso de aquellas asociadas a una alta actividad de secreción celular como, por ejemplo, la secuencia señal del factor VIII de coagulación humano (Freudl, 2018).

Dentro de la secuencia de aminoácidos del factor VIII, incluye una secuencia señal de 19 aminoácidos que direcciona todo el complejo proteico de un peso molecular de 300 kDa (Orlova *et al.*, 2017). Al transportar a este enorme complejo proteico de forma eficiente y de manera constitutiva, torna esta secuencia señal un buen candidato a ser utilizado para la secreción de proteínas recombinantes en líneas celulares de mamíferos.

1.5. Hipótesis

La secuencia señal del factor VIII de coagulación humana presenta una capacidad similar o superior de secretar proteínas recombinantes que otras secuencias señal usadas ampliamente en las líneas celulares de mamíferos.

II. MARCO TEÓRICO

2.1. Bases teóricas sobre el tema de investigación

2.1.1. *Proteínas recombinantes*

Las proteínas recombinantes son formas modificadas de proteínas naturales que se producen de diversas maneras para aumentar la producción de proteínas, alterar las secuencias genéticas y producir productos útiles (Mathew, 2022). Estas proteínas han sido usadas ampliamente en estudios biofísicos y estructurales de proteínas, en el diagnóstico de enfermedades y con fines terapéuticos (Jozala *et al.*, 2016).

2.1.2. *Proteínas fluorescentes*

Las proteínas fluorescentes son un grupo de proteínas quimioluminiscente, siendo la proteína verde fluorescente (*Green Fluorescent Protein*, GFP) de la medusa *Aequorea jellyfis*, la primera descubierta por Shimomura *et al.* (1962). Esta proteína genera su propia fluorescencia por una espontánea ciclización y oxidación de la secuencia interna de Ser65-Tyr66-Gly67 (Tsien, 1998). A partir de esta proteína se originaron las proteínas fluorescentes azul, cian y amarillo, a través de un proceso de mutagénesis (Remington, 2011).

La GFP de la medusa *A. jellyfis* se ha utilizado ampliamente en biología molecular y celular (Elslinger *et al.*, 1999). La independencia de esta proteína de los sustratos enzimáticos ha permitido su uso en embriones y animales transgénicos (Gagneten *et al.*, 1997), y para monitorear la efectividad de la transferencia genética (Cheng *et al.*, 1996).

2.1.3. Proteína amarilla fluorescente

La proteína amarilla fluorescente es una proteína mutante de la proteína verde fluorescente, en la cual fue sustituido el aminoácido treonina 203 por tirosina, generando un desplazamiento de su espectro de emisión en dirección al rojo, con un pico de 528 nm de longitud de onda, y las sustituciones adicionales en los aminoácidos, serina 65 por glicina, valina 68 por leucina y serina 72 por alanina, que mejoraron el brillo de la proteína en células vivas (Jayaraman *et al.*, 2000).

2.1.4. Células CHO-DG44

Las células CHO fueron creadas a partir de células de ovario de hámster chino en el laboratorio del Dr. Theodore Puck al final de la década del 50 (Puck, 1985). A través de mutagénesis *in vitro* de la línea parental, se generaron mutaciones en el gen de la enzima dihidrofolato reductasa, dando origen a las líneas celulares conocidas como CHO-DUKX-B11 e CHO-DG44 (Urlaub y Chasin, 1980). La mutación en esta enzima inhabilita la síntesis de purinas e pirimidinas en estas líneas celulares, por esta enzima participar en la síntesis *de novo* de estas bases nitrogenadas (Kaufman, 1990).

Las células CHO son de particular interés para la fabricación industrial de productos biofarmacéuticos, debido a ventajas como, la viabilidad de su manipulación genética, la capacidad de crear patrones de glicosilación similares a los humanos y la capacidad de crecimiento en altas densidades celulares en cultivos en suspensión sin suero (Attallah *et al.*, 2017) (Figura 1).

Figura 1

Línea celular derivada de ovario de hámster chino (CHO DG44) observadas al microscopio compuesto con un aumento total de 200X.

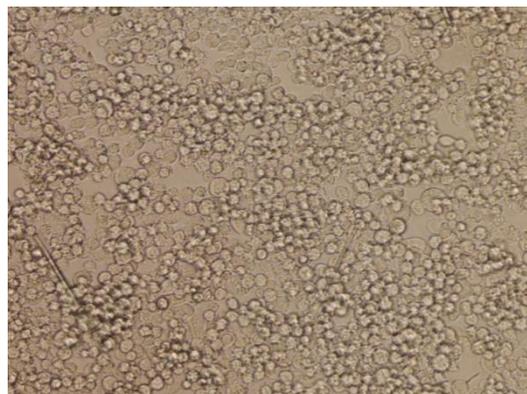
**2.1.5. Células HEK293T**

La línea celular HEK 293T fue formada a partir de las células de riñón embrionario humano, que posteriormente fue transformada con la proteína “large” T del virus de simio 40, y denominada inicialmente como 293tsA1609neo(Kim *et al.*, 2011).

Inicialmente, el uso de células HEK293 se limitó a aplicaciones de producción de adenovirus (Gálvez *et al.*, 2012). Hoy en día, HEK293 es una de las líneas celulares de origen humano más importantes utilizadas para la producción de bioterapéuticos industriales (Román *et al.*, 2016)(Figura 2).

Figura 2

Línea celular derivada de riñón embrionario humano (HEK 293T) observadas al microscopio compuesto con un aumento total de 50X.



2.1.6. *Secuencias señal*

Las secuencias señal son secuencias cortas de aminoácidos presentes en la región amino terminal de muchas proteínas recién sintetizadas, y que dirigen las proteínas por sus respectivos sistemas de exportación presentes en la membrana citoplasmática (Freudl, 2018). Esta secuencia está presente tanto en células procariontes como en eucariontes, donde desempeña un papel importante en la regulación fisiológica de la secreción proteica (Nagano y Masuda, 2014; Almagro Armenteros *et al.*, 2019). Las secuencias señal presentan una estructura primaria conservada, ordenada y dividida en 3 dominios (Von Heijne, 1990). El tamaño de estas secuencias varía entre 18 a 30 aminoácidos, estas variaciones reflejan las diferencias en el grosor de las bicapas lipídicas en las que residen las proteínas secretadas (Paetzel, 2014).

2.1.7. *Factor VIII de coagulación humana*

El factor VIII de coagulación humana (FVIII) es una glicoproteína plasmática grande, que es esencial para la vía intrínseca de la coagulación sanguínea. Está presente en el torrente sanguíneo, principalmente conformando el complejo no covalente con el factor de Von Willebrand (Mei *et al.*, 2006). La disminución del nivel de FVIII o la pérdida de su función procoagulante debido a anomalías genéticas da como resultado el trastorno hemorrágico de la hemofilia A (Orlova *et al.*, 2017).

El gen FVIII se encuentra en la parte distal del brazo largo del cromosoma X y comprende 26 exones, que codifican una cadena polipeptídica de 2351 aminoácidos (Wood *et al.*, 1984). Este gen incluye una secuencia señal de 19 aminoácidos y 2332 aminoácidos de proteína madura (Mei *et al.*, 2006). El FVIII se sintetiza como una sola cadena de seis dominios estructurales con una masa molecular que va desde 190 a 300 kDa (Toole *et al.*, 1984).

2.1.8. Regulación genética en eucariontes

La regulación genética en eucariontes se inicia con la abertura de la estructura de la cromatina por los factores de remodelación de la cromatina, permitiendo así, que los factores de transcripción (*Transcription Factor*, TF) y la ARN polimerasa II puedan unirse al promotor del gen para iniciar su transcripción (Darzacq *et al.*, 2005). Junto a los promotores, el gen presenta otros elementos reguladores *upstream* y *downstream*, como los activadores, potenciadores, represores y silenciadores, que pueden influir en el establecimiento y mantenimiento de la transcripción (Verrijzer *et al.*, 1995). Después del reconocimiento y la unión de TF y el RNA Polimerasa II, se forma el complejo de preiniciación (*Pre-initiation Complex*, PIC), lo que inicia, la síntesis y extensión de la cadena de ARN mensajero (ARNm), seguido de la adición de una capa de guanosina metilada al extremo 5' de la transcripción naciente, para evitar la degradación del ARN por exonucleasas (Stern *et al.*, 2007). Además de otros eventos conocido como, el *splicing*, que es el corte de los intrones intermedios y empalme de los segmentos exónicos, que produce una molécula de ARNm traducible, y finalmente, la adición de una cola de poliadeninas, que permite su liberación (Saguez *et al.*, 2005).

Todas las moléculas de ARNm maduras se transportan desde el núcleo al citosol para su traducción a través del complejo de poros nucleares (Rout y Aitchison, 2001). Una vez en el citosol, el ARNm es dirigido hasta los ribosomas para iniciar su traducción. En el caso de las proteínas secretadas, a medida que emerge la secuencia señal del ribosoma, esta es reconocida por la partícula de reconocimiento de señales (*Signal Recognition Particle*, SRP)(Güler-Gane *et al.*, 2016).

La SRP eucariota es una ribonucleoproteína citoplasmática compuesta por seis polipéptidos y un ARN 7S(Clerigo *et al.*, 2008). Después de su unión con la secuencia señal, la traducción es detenida, evitando así la liberación de las proteínas secretadas en el citosol(Belin *et al.*, 1996). Todo este complejo se transporta a la cara externa del retículo

endoplasmático(RE) donde se une al receptor SRP(Güler-Gane *et al.*, 2016). Tras esta interacción, SRP y su receptor (ambos con dominios de unión de GTP homólogos) estimulan mutuamente sus actividades de hidrólisis de GTP y direcciona la cadena nascente del ribosoma para el lumen, a través del poro interno del canal de translocación, denominado de "translocón", y se libera la SRP unido a GDP del complejo para iniciar un nuevo ciclo de translocación (Clerigo *et al.* 2008). A medida que el polipéptido nascente entra en el ER, él se une por chaperonas moleculares que evitan la agregación de la proteína y auxilia en su plegamiento (Stern *et al.*, 2007). Cualquier glicosilación inicial sucede en cuanto la proteína transita por el translocón, mientras hace esto, la secuencia señal es reconocida por una peptidasa señal y es cribada para generar la proteína madura (Hegde y Bernstein, 2006). Parcialmente glicosilada, esta es empaquetada en gránulos secretorios inmaduros y almacenados en el citosol hasta que la célula recibe una señal accionando su transporte para la membrana plasmática (Murshid y Presley, 2004).

2.1.9. Transfección

La transfección es un proceso mediante el cual se administran ácidos nucleicos extraños a una célula eucariota para modificar la composición genética de la célula huésped (Chow *et al.*, 2016). Generalmente, la transfección se puede clasificar en dos tipos, transfección estable y transitoria (Stepanenko y Heng, 2017).

La transfección estable se refiere al mantenimiento de la expresión a largo plazo de un transgén mediante la integración de ADN extraño en el genoma nuclear del huésped (Fus-kujawa *et al.*, 2021). Luego, el transgén puede expresarse constitutivamente incluso con la replicación de las células (Kim y Eberwine, 2010). Este método suele ser usado para la producción de proteínas a gran escala y a largo plazo (Elgundi *et al.*, 2017).

Por el contrario, la transfección transitoria no requiere la integración de ácidos nucleicos en el genoma de la célula huésped, sino que solo se expresa de forma episomal (Lufino *et al.*, 2008). Por lo tanto, la expresión del transgén eventualmente se perderá a medida que las células huésped se repliquen (Nejepinska *et al.*, 2012). La transfección transitoria generalmente se aplica en estudios a corto plazo y pequeña escala (Kim y Eberwine, 2010).

III. MÉTODO

3.1. Tipo de investigación

La investigación es de tipo descriptiva, debido a que se caracterizan los niveles de secreción de la proteína amarilla fluorescente en los cultivos celulares, además es cuantitativa porque se representan los resultados de forma numérica, es de tipo prospectivo y de corte longitudinal dado que las variables se evaluarán en un periodo definido.

3.2. Ámbito temporal y espacial

El proyecto fue desarrollado en los laboratorios del Núcleo de Terapia molecular y Celular (NUCEL), localizado en el campus Butantã de la Universidad de São Paulo, São Paulo, Brasil. Durante el año 2018.

3.3 Variables

- Variable dependiente: Nivel de secreción de la proteína amarilla fluorescente.
- Variable independiente: Concentración de las células plaqueadas, intervalo de tiempo de los cultivos celulares y procedencia de la secuencia señal.

3.4 Población y muestra

Las poblaciones utilizadas para el desarrollo de este proyecto fueron las poblaciones de las líneas celulares HEK 293T y CHO-DG44, seleccionándose diferentes muestras de acuerdo a los requerimientos de los experimentos realizados.

3.5 Instrumentos

- Para la etapa inicial, el instrumento de validación fue la secuenciación de los vectores generados, para verificar la integridad del constructo realizado.
- La comprobación de la funcionalidad de los vectores construidos, fue efectuada a través de las fotografías de los cultivos celulares obtenidos por microscopia de fluorescencia y de los geles de electroforesis SDS-PAGE.
- Finalmente, como instrumento para validar los niveles de secreción para cada secuencia señal, se utilizó los niveles de absorbancia detectadas de la proteína fluorescente amarilla por fluorometría en cada uno de los experimentos realizados.

3.6 Procedimientos

3.6.1 *Línea celular y manutención*

Las líneas celulares usadas para la realización de este trabajo fueron, las células CHO-DG44 adherentes, para la producción estable de una variante secretable de la proteína amarilla fluorescente (YFP), y las células derivadas de riñón embrionario humano, HEK 293T, para la producción transitoria de esta misma proteína.

Las células HEK 293T fueron mantenidas en el medio de cultivo RPMI suplementado con 10% de suero fetal bovino (SFB). Para el cultivo de las células CHO-DG44 adherentes fue utilizado el medio de cultivo HAM's F12 suplementado con 10% de SFB. Las poblaciones celulares derivadas a partir de la transfección estable de las células CHO-DG44 con el vector

de expresión pLV/spYFP_DHFR fueron mantenidas en medio RPMI sin nucleósidos ni desoxinucleótidos (Invitrogen) suplementado con 7% SFB dializado.

Las células fueron mantenidas en frascos de cultivo T descartables, conteniendo el medio adecuado para cada línea celular, y almacenada en atmosfera humidificada con 2% de CO₂ a 37°C. Las células fueron subcultivadas cuando alcanzaron el 80% de confluencia, conforme descrito en Freshney(2010). Los estoques celulares fueron mantenidos en medios de cultivos conteniendo 10% SFB y 10% dimetilsulfóxido (DMSO) y almacenados en un depósito conteniendo nitrógeno líquido a -190°C.

3.6.2 Construcción del vector de expresión de una variante secretable de la proteína YFP (pLV/spYFP_DHFR)

La estrategia usada para la obtención del vector de expresión bicistronico de la proteína fusión entre la secuencia señal del factor VIII (sp), nucleótidos 172 a 228 de la secuencia de referencia del GenBank NM_000132.3, y la proteína YFP, pLV/spYFP_DHFR, envolvió dos ciclos de PCR. El primer ciclo se basó en la utilización del vector de expresión que contenía una forma variante de la cadena leve del factor VIII humano recombinante, pCXNIID/Lch, como molde para la síntesis del *primer forward* spYFP (5'CCCGGGATCCACCGGTCGCCACCATGCAAATAGAGCTCTCCACCTGCTTCTTTC TGTGCCTTTTGCATTCTGCTTTAGTATGGTGAGCAAGGGCGAGGA3'). Para este primer ciclo de PCR, el medio reaccional fue constituido por el vector pCXNIID/Lch (400 ng), tampón Phusion HF Buffer 1X (Invitrogen), MgCl₂ 2 mM, dNTPs 0,2 mM, *primer forward* spLCE(5'CCCGGGATCCACCGGTCGCCACCATGCAAATAGAGCTCTCCAC3') y el *primer reverse* pLCR(5'TCCTCGCCCTTGCTCACCATACTAAAGCAGAATCGCAAAAGG3'), en la concentración de 0,5 µM, Taq polimerasa (Invitrogen) 1U y agua milliQ® suficiente para un volumen final de 50 µl. El programa usado en el termociclador fue:

98°C	30 s	35 ciclos
98°C	10 s	
65°C	30 s	
72°C	30 s	
72°C	10 min	

El fragmento de DNA correspondiente al *primer forward* spYFP, de aproximadamente 100pb, que contenía la secuencia codificadora de la secuencia señal del Factor VIII humano fusionada al inicio de la región codificadora de la proteína YFP, fue purificado a través de una electroforesis en gel de agarosa 2%, utilizando el kit QIAquick Gel Extraction (Quiagen), siguiendo las instrucciones del fabricante. El *primer forward* spEYFP purificado fue concentrado en el Rotoevaporador a vacío Speed-Vac®, y usado en el segundo ciclo de PCR para la síntesis completa del elemento spYFP_IRES_DHFR. En el segundo ciclo de PCR, se usó el vector pUC19/YFP_DHFR, que fue cedido por nuestro grupo de laboratorio, este vector presenta la secuencia codificadora de la proteína YFP, el elemento IRES y del DHFR. Para esta PCR fue usado 400ng de este vector y el medio reaccional fue constituido por, tampón Phusion HF Buffer 1X (Invitrogen), MgCl₂ 2 mM, dNTPs 0,2 mM, *primer forward* spYFP (8 ng), *primer reverse* complementario a una porción anterior de la secuencia señalizadora de poliadenilación del vector pUC19/YFP_DHFR (2222R) (5'CCCGTTCGACCCAGCCACCCACCTTCTGATA3') en la concentración de 0,5μM, Taq polimerasa (Invitrogen) 1U y agua milliQ® suficiente para un volumen final de 20 μl. El programa usado en el termociclador será:

98°C	30 s	35 ciclos
98°C	10 s	
65°C	30 s	
72°C	120 s	
72°C	10 min	

El producto resultante de esta reacción de PCR, el fragmento de DNA spYFP_IRES_DHFR, fue purificado a partir de electroforesis en gel de agarosa al 1% y concentrado como descrito anteriormente. Seguidamente, el fragmento spYFP_IRES_DHFR fue codigerido con las enzimas de restricción Sall y BamHI, de acuerdo con las orientaciones del fabricante (ThermoScientific).

En paralelo, fue digerido el vector pLV/GFP con las mismas enzimas para la eliminación del elemento GFP_IRES. Este vector, es un vector lentiviral de tercera generación que presenta la secuencia codificadora de la proteína GFP, cedido por nuestro grupo de laboratorio. El fragmento de DNA correspondiente al vector pLV fue purificado a través de una electroforesis en gel de agarosa al 1%, y purificado como descrito anteriormente. A partir del elemento spYFP_IRES_DHFR y de la porción del vector pLV codigeridos con las enzimas Sall y BamHI, fue realizado la reacción de ligación de estos fragmentos, utilizando la enzima T4 DNA Ligasa (Invitrogen), siguiendo las indicaciones del fabricante.

3.6.3 Transformación de bacterias *Escherichia coli*, cepa XLI con el vector pLV/spYFP_IRES_DHFR

Una alícuota de la reacción de ligación fue utilizada para la transformación por electroporación (cubetas de 2mm, y un pulso de 2500V/25µF/200Ω - Electroporador Gene

Pulser XCELL – Biorad) de las bacterias *E. coli*, cepa XL1, electrocompetentes. Las bacterias electroporadas fueron inoculadas en medio LB a 37°C sobre agitación por una hora, plaqueadas en medio LB-Agar conteniendo ampicilina (100µg/ml), e incubadas por 16 horas a 37°C, lo que permitió la selección de clones bacterianos que recibieron el vector de expresión.

3.6.4 *Caracterización del vector de expresión de una variante secretable de la proteína EFYP (pLV/spYFP_IRES_DHFR)*

La búsqueda de las colonias bacterianas recombinantes fue realizada por PCR a partir de los inóculos bacterianos, empleándose los tres pares de *primers*:

- spLC F (5'TACCGCGGGCCCGGGATCCAGTCGACCACCATGCAAATAG3') / 2222R, con amplicon esperado de 2140pb
- pUC seq5F(5'AGTTGTGGAAAGAGTCAAATGG3') / WPRE R(5'GCCACAACCTCCTCATAAAGAGACAG3'), con amplicon esperado de aproximadamente 1Kpb
- pUC seq6F (5'GAACTCAAAGAACCACCACG3')/ WPRE R, con amplicon esperado de aproximadamente 550pb).

Por medio de esta estrategia de selección, los clones bacterianos asociados a la detección de los fragmentos de DNA esperados en las tres reacciones de PCR fueron inoculados en medio LB + Ampicilina (100µg/ml) a 37°C por 16 horas, para su posterior recuperación y purificación de los plásmidos recombinantes de interés, através de una Minipreparación de DNA plasmidial (Miniprep), usando el plasmind Miniprep Kit (Qiagen), siguiendo orientaciones del fabricante. A partir de las Miniprep obtenidas, fue evaluado la integridad de los vectores aislados a través de la reacción de digestión por las enzimas:

- XhoI (fragmentos de DNA esperados de 820pb e 7,9Kpb),
- BamHI/Sall (fragmentos de DNA esperados 2140pb e 6,6Kpb),
- PstI (fragmentos esperados de 300pb e 8,6Kpb) y
- EcoRV (fragmentos esperados de 830pb e 7,9Kpb),

Aquellos clones bacterianos con el perfil de restricción esperado fueron evaluados por secuenciación de DNA. Los análisis de la integridad de la secuencia del elemento spYFP_IRES_DHFR fueron realizadas por secuenciación de DNA (método de Sanger) a través de la utilización del kit Big Dye (ThermoScientific), empleándose los *primers* flanqueadores del sitio de clonaje y siguiendo las orientaciones del fabricante. Las reacciones de secuenciación fueron analizadas en un secuenciador de DNA modelo ABI 3700, en el laboratorio GATE Lab, del Instituto de Biología de la USP. Los montajes de las secuencias consenso fueron realizadas a través del programa Seqman (DNASTAR), y los alineamientos con la secuencia teórica del fragmento spEYFP_IRES_DHFR fueron realizados a través de la herramienta bioinformática BLAST(<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

3.6.5 Transfección transitoria en las células HEK 293T con el vector pLV/spYFP_DHFR

Los experimentos de transfección transitoria en las células HEK 293T fueron realizados mediante la técnica de lipofección, y por tanto se utilizó el reactivo Lipofectamina 2000[®](Invitrogen), conforme las orientaciones del fabricante. Además, se realizó este mismo procedimiento con el vector pLV/GFP, que contenía la proteína verde fluorescente (GFP), y los vectores pLV/ilYFP_DHFR y pLV/abYFP_DHFR, que presentan una forma secretable de la proteína YFP modificada por las secuencias señal de la interleucina-2(il) y la albumina

humana(ab), respectivamente; los cuales fueron cedidos por el grupo de laboratorio y usados como referencias.

La funcionalidad del casete de expresión del vector construido fue evaluada en la monocapa celular 72 horas después del momento de la transfección, a través de la observación de la fluorescencia de las células en el microscopio de fluorescencia Nikon, modelo TE-300, con el conjunto de filtros de excitación/emisión y espejo dicróico para el fluoróforo isotiocianato de fluoresceína (FITC) (excitación 460 – 510 nm y emisión 515 – 565 nm). Adicionalmente, la secreción de la proteína YFP fue determinada a partir de la detección de esta proteína en el medio condicionado colectado, evaluados a través de electroforesis en gel de poliacrilamida conteniendo SDS (SDS-PAGE 10%), seguido de escaneamiento a laser del gel para la detección de fluorescencia en el equipo Typhoon Trio, utilizando como longitud de onda de excitación un láser de 532 nm. Posteriormente, los geles fueron fijados y coloreados con Coomassie blue G-250 coloidal.

3.6.6 Transfección estable de las células CHO-DG44 con el vector pLV/spYFP_DHFR

Para la transfección estable de las células CHO-DG44 con los vectores pLV/spYFP_DHFR, pLV/abYFP_DHFR y pLV/ilYFP_DHFR, fueron plaqueados cerca de 2×10^5 células en un pozo de una placa de 24 pozos en medio HAM's F12 suplementado con 10% de SFB. La transfección fue realizada por lipofección (Lipofectamina 2000[®] - Invitrogen), siguiendo orientaciones del fabricante, se empleó 500 ng del vector plasmidial. Cerca de 48 horas después a la transfección, la monocapa celular fue tripsinizada. Seguidamente, las células fueron subdivididas para 4 pozos de una placa de 6 pozos, y se realizó el cambio de medio por el medio de cultivo de selección (RPMI conteniendo 7% SFB dializado sin nucleósidos ni desoxinucleótidos) basado en la expresión de la DHFR. El medio de cultivo fue renovado periódicamente, y las poblaciones celulares mixta resultantes, fueron mantenidas

en cultivo hasta alcanzar un número de células suficiente (cerca de 4×10^6 células) para congelamiento y un porcentaje de viabilidad celular mayor al 90%, evaluado a través de la utilización del colorante azul de tripano al 0,4%(m/v). Las poblaciones mixtas fueron observadas al microscopio de fluorescencia Nikon modelo TE-300, con el conjunto de filtros y espejo dichroico para el fluoróforo isotiocianato de fluoresceína (FITC) (excitación 460 - 510nm y emisión 515 -565nm) y se determinó la secreción por fluorometría y electroforesis SDS- PAGE, como ha sido descrito anteriormente.

3.6.7 Evaluación de los niveles de secreción de la proteína YFP en las poblaciones celulares mixtas

La determinación de los niveles de secreción de la proteína YFP se realizó en cada uno de las poblaciones celulares generadas. Esta evaluación fue realizada a través de la cantidad acumulada de la proteína YFP secretadas por las células en el medio de cultivo condicionado colectado. En placas de seis pozos se sembraron 1×10^5 células/ml con un volumen final de 3 ml, incubados a 37 ° C en un ambiente humidificado y una atmósfera conteniendo 2% de CO₂. Después de 72 h, se colectaron los medios condicionados, se clarificaron por centrifugación a 200xg durante 5 min y se almacenaron a -20 ° C hasta el momento del análisis.

Para la evaluación de los niveles de secreción de la proteína YFP, las muestras de los medios condicionados colectados se concentraron en un factor de 10X mediante ultrafiltración en microcolumnas Amicon de 10 KDa. Seguidamente, estas muestras fueron evaluadas por fluorometría en el equipo SpectraMax® Paradigm® con filtro de excitación 510 nm y filtro de emisión 530 nm, los valores de absorbancias detectadas fueron normalizadas con la relación del número de células iniciales y el número de células finales. Estos experimentos fueron realizados por triplicado.

3.6.8 *Análisis de datos*

Los análisis estadísticos usados fueron los testes de ANOVA de una vía y posteriormente el teste de Tukey de múltiple comparación, todas las conclusiones fueron tomadas al nivel de significancia de 5%. Se realizo un análisis descriptivo de los resultados en tablas y gráficos utilizando el software GraphPad Prism 5.0 (GraphPad software, Inc., San Diego, CA, USA).

IV. RESULTADOS

4.1. Construcción del vector de expresión de una variante secretada de la proteína YFP

El vector de expresión construido presenta en el primer cistron, un gen reporte, una forma secretable de la proteína amarilla fluorescente (*Yellow Fluorescent Protein*, YFP) generada a través de la adición de la secuencia señal de la cadena leve del factor VIII de coagulación humana, para evaluar su potencial de secretar proteínas recombinantes en líneas celulares de mamíferos. Esta capacidad de secreción de esta secuencia señal fue comparada con la capacidad de secreción de las secuencias señal de la interleucina-2 y de la albumina humana, utilizadas ampliamente en la industria e investigación de bioprocesos (Komada *et al.*, 1999; Suzuki *et al.*, 2001; Knappskog *et al.*, 2007; Kober *et al.*, 2013). En el segundo cistron, se localiza la región codificadora de la enzima DHFR, un marcador de selección, que permitió la formación de una población celular mixta secretora establemente de la proteína YFP (Figura 3).

Figura 3

Representación esquemática de los principales elementos del vector lentiviral bicistronico de expresión pLV/spYFP_DHFR.



Nota. Una variante secretable de la proteína amarilla fluorescente (spYFP) generada por la fusión de las secuencias codificadoras de la secuencia señal del facto VIII de coagulación humano(sp) y de la proteína YFP. LTR: repetición terminal larga derivada del virus HIV-1; CMV: región promotora derivada del citomegalovirus humano; IRES: Sitio interno de entrada del ribosoma; DHFR: región codificadora de la enzima DHFR de ratón

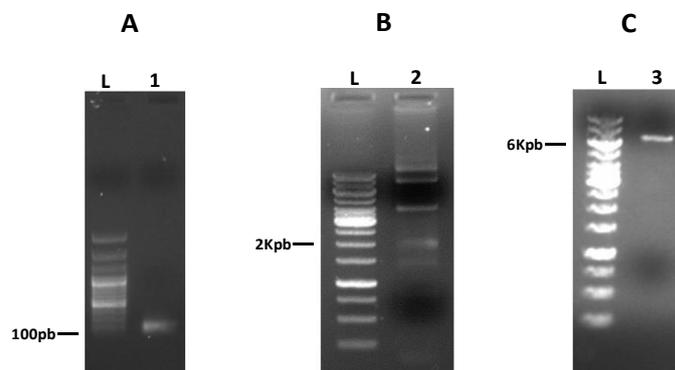
Para la construcción del mismo fue requerido dos ciclos de PCR. En la primera reacción de PCR fue generado un megaprimer que contenía la región codificadora de la secuencia señal de la cadena leve del factor VIII de coagulación humano subclonado a partir del vector pCXNIID/Lch, que fue cedido por el grupo de investigación. El producto generado fue purificado del gel de agarosa y usado en la segunda PCR.

El producto obtenido de la segunda reacción de PCR, spYFP-IRES-DHFR, fue purificado a partir de electroforesis en gel de agarosa, y posteriormente, codigerido con las enzimas de restricción Sall y BamHI. Paralelamente, el vector bicistrónico pLV/GFP y pUC19/YFP_DHFR fueron codigeridos con las mismas enzimas, y el fragmento correspondiente al vector pLV y el constructo spYFP_IRES_DHFR fueron purificados a partir de una electroforesis en gel de agarosa.

En la figura 4, se observa el fragmento amplificado por PCR correspondiente a la secuencia señal del FVIII, con un tamaño similar al esperado para la secuencia teórica (aproximadamente 150 pb) al ser evaluado en electroforesis en gel de agarosa. De la misma forma, el fragmento generado a partir de la amplificación por PCR del casete spYFP_IRES_DHFR, de cerca de 2.2Kpb, correspondió al tamaño esperado, el cual fue observado juntamente a diferentes productos inespecíficos o posibles formas de doblamientos del plásmido pUC19/GFP_DHFR, utilizado como molde (Figura 4B). La digestión del vector pLV/GFP con las enzimas de restricción Sall y BamHI tuvo una alta eficiencia y especificidad, evidenciadas por la detección, después de la purificación del fragmento de DNA correspondiendo al vector pLV sin el elemento GFP_IRES, de una única banda de DNA definida en el gel de agarosa, y presentando el tamaño esperado de 6,9Kpb (Figura 4C).

Figura 4

Análisis por electroforesis en gel de agarosa al 1% de los fragmentos de DNA



Nota. spYFP (Figura 4A, carril 1, tamaño aproximado 100pb), spYFP-IRES-DHFR (Figura 4B, carril 2, tamaño aproximado 2.2kpb) y pLV (Figura 4C, carril 3, tamaño aproximado 6.9kpb). L: Padrón de peso molecular.

4.2. Caracterización del vector de expresión de una variante secretada de la proteína YFP (pLV/spYFP_DHFR)

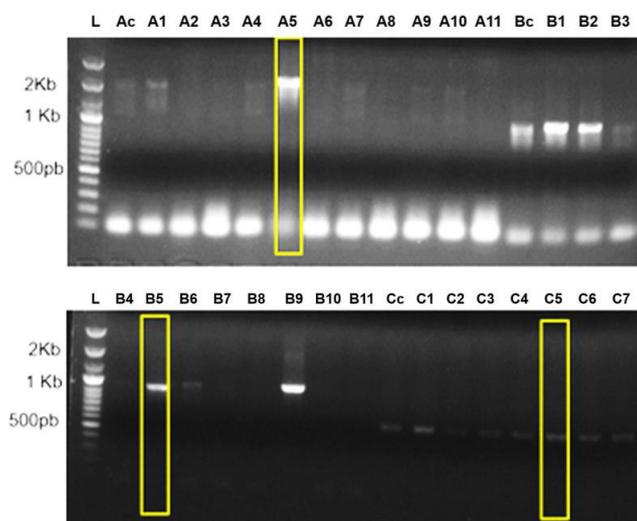
Se siguió con la reacción de ligación entre el fragmento spYFP-IRES-DHFR, digerido con las enzimas de restricción Sall y BamHI, y el vector pLV, digeridos con las mismas enzimas. Una alícuota de la reacción de ligación fue utilizada para la transformación por electroporación de la cepa XL1 Blue de las bacterias *E. coli*. En las placas de LB crecieron alrededor de 70 colonias por placa, así que fue necesario seleccionar aquellas colonias que serían usadas para la minipreparación de DNA plasmidial. Por este motivo, se implementó la estrategia de selección por PCR, usándose los pares de *primers* spLC F/ 2222R (el inserto), pUC seq5F/WPRE R (región consenso del gen IRES), y pUC19 seq6F/WPRE R (región consenso del gen DHFR) y como control positivo fue usado el vector pLV/GFP_IRES y pUC19/YFP_DHFR.

Las colonias bacterianas recombinantes obtenidas fueron inoculadas en pozos de una placa de 96 pozos conteniendo 200 µl de medio LB y ampicilina (100 µg/ml), y puestas a

incubar por 24 horas a 37°C, para ser usadas en la PCR. Al ser corrida las reacciones de PCR en gel de agarosa se evidencio los amplicones correspondiente al tamaño esperado, inclusive nuestros controles positivos amplificaron, como puede observarse en la Figura 5, con este experimento se demostró que esta estrategia es eficiente como método de evaluación y se procedió a realizar con las demás colonias obtenidas.

Figura 5

Selección de 11 colonias bacterianas recombinantes por PCR obtenidas después de la transformación de las bacterias E. coli (cepa XL-1 Blue) con el producto de ligación pLV/spYFP_DHFR.



Nota. Alícuotas de las reacciones de PCR fueron fraccionadas por electroforesis en gel de agarosa al 1% (m/v). los carriles indicados como Ac a A11 están relacionadas al PCR realizada con el par de iniciadores spLC F/ 2222R (fragmento esperado 2140pb), las indicadas como Bc a B11 al par de iniciadores pUC seq5F/WPRE R (fragmento esperado 1Kpb) y las indicadas como Cc a C7 al par de iniciadores pUC seq6F/WPRE R (fragmento esperado 550pb). Ac y Bc: muestras del vector pLV/GFP_IRES con el par de primer correspondiente. Cc: muestra del vector pUC19/YFP_DHFR. L: Marcador de peso molecular. En destaque (rectángulo amarillo) está indicado el perfil de fragmentos para un de los clones positivos obtenidos (clone bacteriano recombinante CR5).

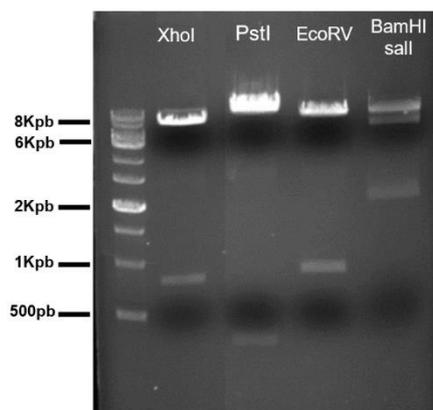
En la figura 6, se muestran los análisis por electroforesis en gel de agarosa de las reacciones de PCR utilizadas para la selección de los clones bacterianos recombinantes obtenidos. Observase en destaque el perfil de fragmentos de DNA obtenidos para un clon

bacteriano positivo (clon bacteriano CR5, carriles A5, B5 y C5). Además del clon CR5, fueron obtenidos otros clones positivos, empleándose el mismo procedimiento de selección (clones bacterianos CR1, CR20, CR36, CR60 y CR82) (datos no mostrados).

Todos estos clones asociados a la detección de los fragmentos de DNA esperados en las tres reacciones de PCR fueron seleccionados para la extracción del DNA plasmidial, y posteriormente se les determinó su perfil de restricción con las enzimas XhoI, BamHI/SalI, PstI y EcoRV para la evaluación de la integridad de los vectores aislados (Figura 6).

Figura 6

Análisis por electroforesis en gel de agarosa 1% (m/v) del perfil de restricción con las enzimas indicadas del clon CR5 del vector recombinante pLV/spYFP_DHFR.



Nota. Los fragmentos de DNA esperados eran 7.8Kpb y 800pb (XhoI), 6.3Kpb y 2.3Kpb (BamHI/SalI), 8.4Kpb y 200pb (PstI) y 7.9Kpb y 900pb (EcoRV). L: Padrón de peso molecular.

La Figura 6 representa el análisis del perfil de restricción de uno de los clones bacterianos obtenidos del vector pLV/spYFP_DHFR, el CR05. Como puede ser observado, el perfil de fragmentos de DNA obtenidos está de acuerdo con el tamaño esperado, indicando la integridad estructural del vector de expresión pLV/spYFP_DHFR. De la misma forma, fueron evaluados los otros clones bacterianos (clones CR1, CR20, CR36, CR60 y CR82), los cuales también presentaron el perfil de restricción esperado (datos no mostrados). Todos los clones

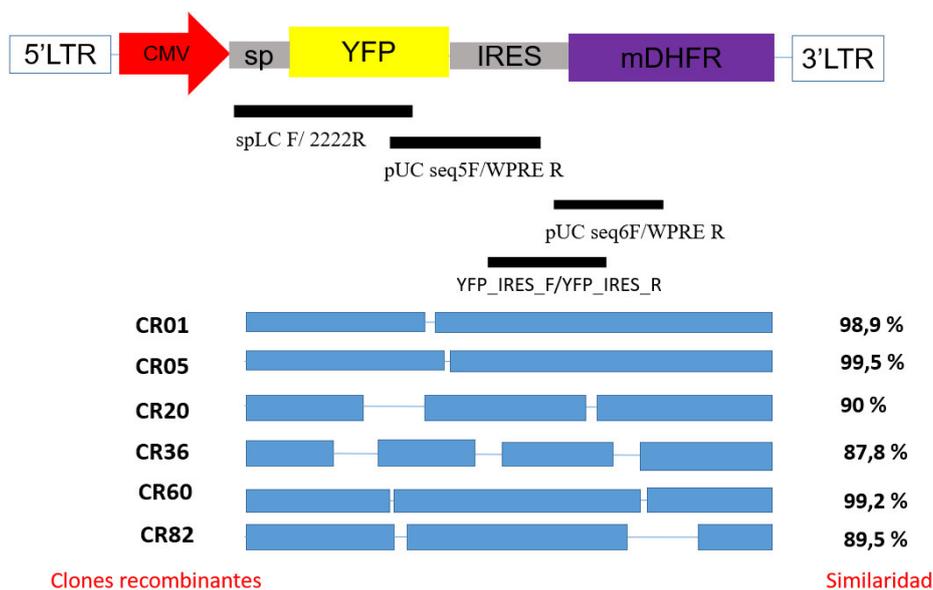
positivos confirmados por PCR y perfil de restricción fueron secuenciados con el objetivo de evaluar la identidad y la integridad de las regiones codificadoras de la proteína de fusión spYFP y DHFR, y del elemento IRES, conforme descrito en Materiales y Métodos.

Los alineamientos entre las secuencias de DNA de los clones CR1, CR20, CR36, CR60 y CR82, con la secuencia teórica del elemento spYFP_IRES_DHFR, realizados a través de la herramienta BLAST(Anexo 1), mostraron la presencia de *gaps* en las secuencias evaluadas, localizados en diferentes porciones del casete de expresión. Estos *gaps* pueden estar relacionados a deleciones o inserciones en esas regiones, con lo que representa una alteración en su secuencia y consecuentemente posibles modificaciones en la funcionalidad de la proteína. En el caso del clon CR82, presento *gaps* en la extremidad 3' del constructo evaluado, alteraciones en esta región podría generar la falta de expresión de la proteína DHFR, con lo que no sería posible usarlo como marcador de selección para generar la población celular por transfección estable. De la misma forma en el clon CR20, que presento varios *gaps* en el extremo 5' que corresponde a la secuencia codificadora de la proteína YFP, la falta de expresión de esta proteína no permitiría evaluar la fluorescencia ni realizar los experimentos posteriores. También fue detectado clones con la presencia de *gaps* en ambos extremos de la secuencia consenso, como en el clon CR36, con lo que comprometería la funcionalidad de ambas proteínas (Figura 7).

De todos los clones evaluados, apenas las secuencias consenso de los clones CR01(98.9%), CR05(99.5%) y CR60(99.2%) presentaron una similaridad casi completa con la secuencia teórica en las secciones evaluadas, complementando y confirmando los datos obtenidos en la caracterización parcial por PCR y por el perfil de restricción. De estos 3 clones, fue elegido el clon CR05 para continuar con los próximos experimentos, debido a que fue el que presento mayor similaridad y fidelidad con la secuencia teórica

Figura 7

Representación esquemática del alineamiento realizado a través de la herramienta BLAST de las secuencias consenso de los clones recombinantes obtenidas por secuenciación por el método de Sanger, con la secuencia teórica del elemento *spYFP_IRES_DHFR*.



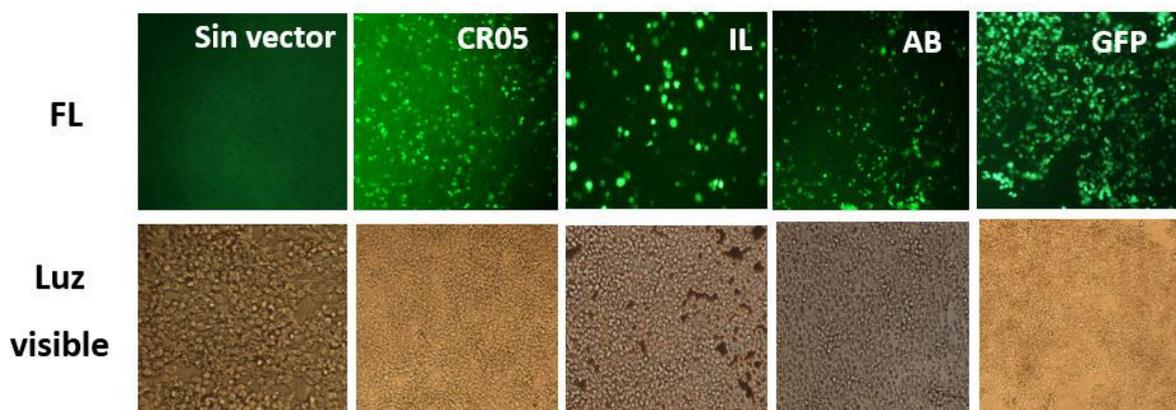
Nota. Las barras abarcan las regiones donde presentaron similitud, y los espacios vacíos la presencia de *gaps* en las regiones evaluadas.

4.3 Generación de las poblaciones celulares mixta a través del método de transfección transitoria

Los experimentos de transfección transitoria por el método de lipofección con el clon CR05 del vector pLV/spYFP_DHFR en las células HEK293T, permitió evaluar de forma rápida la expresión y secreción de la proteína YFP. Adicionalmente, se realizó el mismo procedimiento con los vectores pLV/ilYFP_DHFR y pLV/abYFP_DHFR (Figuras 8 y 9).

Figura 8

Evaluación de la expresión transitoria de la proteína spYFP en la línea celular HEK 293T por microscopia de fluorescencia.



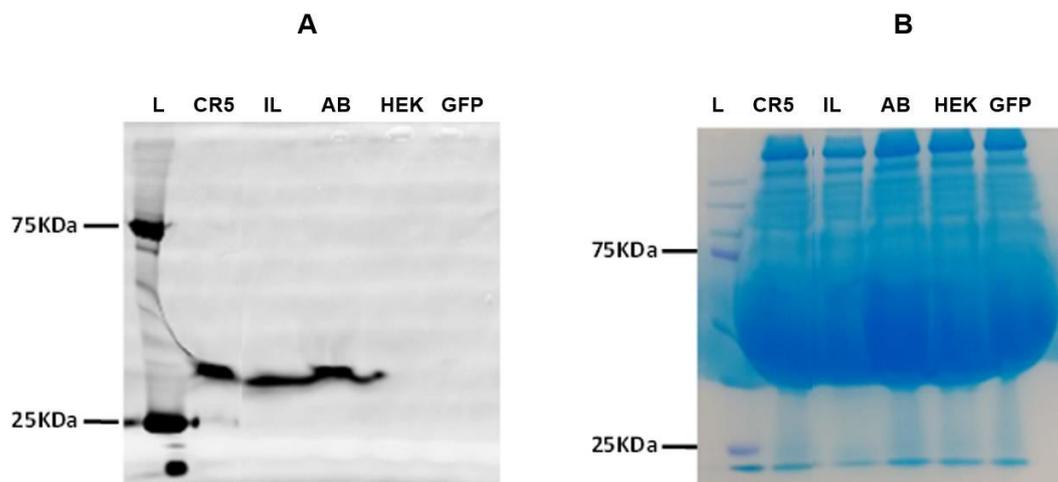
Nota. Las células HEK-293T fueron transfectadas transitoriamente por lipofección con el vector pLV/spYFP_DHFR (clon CR5), y evaluadas por microscopia de campo claro (luz visible) y fluorescencia (FL) 48 horas después de la transfección. Como control positivo de la detección de la fluorescencia de la proteína, se transfecto las células HEK-293T con el vector de expresión plasmidial pLV/GFP. Como control negativo, se emplearon las células HEK-293T no sometidas al procedimiento de transfección.

En la evaluación de las monocapas celulares por microscopia de fluorescencia (Figura 4 FL), se puede observar que el clon CR05 del vector pLV/spYFP_DHFR es funcional en cuanto a la capacidad de expresión de la variante secretable de la proteína fluorescente YFP modificada por la secuencia señal del factor VIII, en experimentos de transfección transitoria de las células HEK-293T. Este mismo resultado fue encontrado en las poblaciones celulares mixtas de la variante secretable de la proteína YFP modificada por las secuencias señal de la interleucina 2 y la albumina humana. En estos experimentos de transfección transitoria fue posible detectar la emisión de la fluorescencia intracelular de la variante secretable de la YFP con cada una de las secuencias señal evaluadas, posiblemente por el tránsito de estas moléculas en la vía secretoria. Las células HEK-293T no sometidas al procedimiento de transfección transitoria no presentaron emisión de fluorescencia en las condiciones de detección utilizadas, demostrando que este tipo celular no pose componentes celulares que emitan fluorescencia en estas condiciones de detección. En las células HEK-293T transfectadas con el vector de

expresión pLV/GFP, utilizadas como control positivo, se observó la presencia de fluorescencia que indicaba el acúmulo intracelular de la proteína GFP, como era lo esperado.

Figura 9

Evaluación de la secreción de la proteína YFP (~27KDa) en la línea celular HEK293T por SDS-PAGE 10% en condiciones no reductoras, seguida por detección de fluorescencia por escaneamiento a laser.



Nota. A) Muestras concentradas (10x) por ultrafiltración, del medio de cultivo condicionado por 48 horas por las células HEK-293T transfectadas transitoriamente por lipofección con los vectores pLV/spYFP_DHFR (clon CR5 - carril CR5), pLV/iYFP_DHFR(carril IL) y pLV/abYFP_DHFR(carril AB), fueron fraccionadas por SDS-PAGE 10% en condiciones no reductoras, seguida por detección de la emisión de fluorescencia a 540 nm por escaneamiento a laser (530nm). Como control negativo, se empleó el medio de cultivo condicionado por 48h por las células HEK-293T no sometidas al procedimiento de transfección (carril HEK), o por las células HEK-293T transfectadas con el vector de la proteína fluorescente no secretada GFP, pLV/GFP (carril GFP). (B) Coloración con Coomassie blue coloidal del gel de acrilamida después del escaneamiento a laser. L: Padron de peso molecular.

Se siguió con la detección de la proteína YFP en los medios de cultivos colectados por electroforesis SDS-PAGE, conforme descrito en Métodos. A través de esta técnica fue posible constatar que la proteína YFP es secretada para el medio de cultivo de las células HEK-293T transfectadas transitoriamente con el clon CR5 del vector pLV/spYFP_DHFR y las otras poblaciones celulares mixtas generadas, a partir de la detección de una doble banda proteica fluorescente con tamaños correspondientes a la proteína YFP (~27KDa), sin embargo el

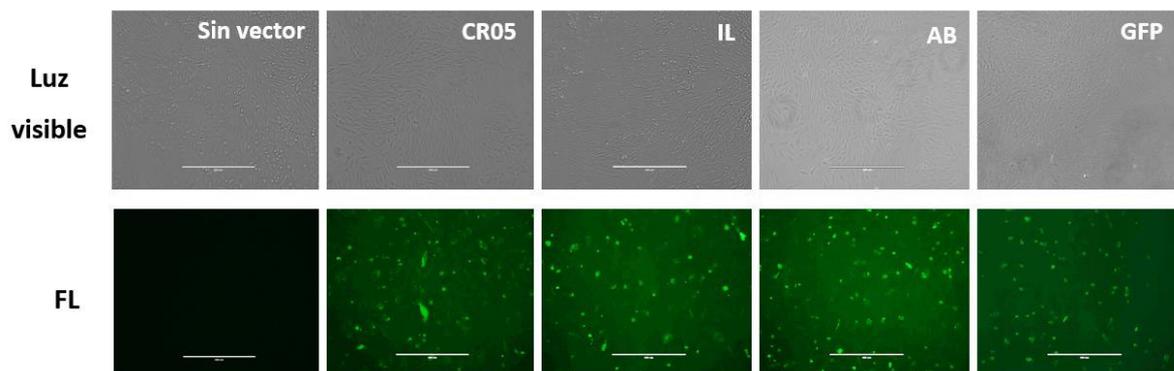
motivo de la existencia de estas dos bandas referente a esta proteína son desconocidas. La posibilidad de la detección de la proteína YFP en el medio de cultivo ser debido a la infiltración del contenido intracelular asociado a eventos de muerte celular fue descartado frente al resultado obtenido con un de los controles negativos utilizados, el medio condicionado colectado de las HEK293T sin transfectar y de las células HEK-293T transfectadas con el vector de expresión de la proteína fluorescente GFP, pLV/GFP, la cual, por no presentar una secuencia señal, es acumulada en el medio intracelular (Figura 9A, carril GFP). Estos experimentos demostraron la funcionalidad y la secreción de la proteína YFP expresada por el clon CR5 del vector pLV/spYFP-DHFR.

4.4 Generación de las células CHO-DG44 con expresión estable de la proteína YFP

Al igual que en la transfección transitoria, el clon CR5 del vector pLV/spYFP_DHFR junto con los vectores pLV/ilYFP_DHFR y pLV/abYFP_DHFR fueron utilizados para la transfección estable en la línea celular de ovario de hámster chino, conforme descrito en Métodos. Las poblaciones celulares mixtas obtenidas fueron evaluadas de la misma forma que con la línea celular humano HEK 293T, los niveles de secreción por fluorometría, la detección de la proteína YFP por electroforesis en SDS-PAGE y microscopia de fluorescencia (Figura 10 y 11), usándose también como controles, la línea celular parental y células transfectadas con el vector pLV/GFP_IRES, para evidenciar posibles proteínas siendo secretada en el medio de cultivos diferentes a la proteína evaluada.

Figura 10

Evaluación de la expresión estable de la proteína YFP en las células CHODG44 por microscopia de fluorescencia.



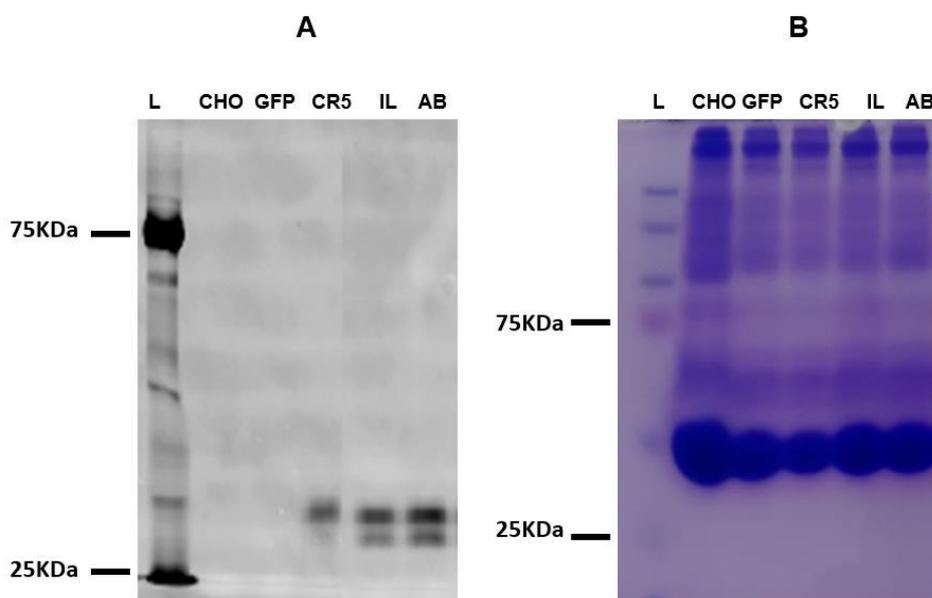
Nota. Las células CHO-DG44 fueron transfectadas establemente por lipofección con el vector pLV/spYFP_DHFR (clon CR5), y evaluadas por microscopia de campo claro (luz visible) y fluorescencia (FL). Como control positivo de la detección de la fluorescencia de la proteína, se transfecto las células CHO-DG44 con el vector de expresión plasmidial pLV/GFP. Como control negativo, se emplearon las células CHO-DG44 no sometidas al procedimiento de transfección.

En las células de las poblaciones celulares CHO generada por la transfección estable con el clon CR05 del vector pLV/speYFP_DHFR se evidencio la presencia de fluorescencia intracelular. Al igual que en las poblaciones celulares mixtas generadas por la transfección estable con los vectores que contenían la secuencias señal de la interleucina-2 y la albumina humana, y aquellas transfectadas con el vector pLV/GFP_IRES, usado como control positivo, ya que expresa intracelularmente la proteína GFP. En la línea celular parental no se observó la presencia de ninguna molécula que presente una fluorescencia en el espectro de emisión evaluado, demostrando así, que la detección de fluorescencia correspondía a esta proteína en específico. La evaluación de la secreción por cromatografía SDS-PAGE determino que la molécula detectada en la microscopia de fluorescencia correspondía a la variante secretable de la proteína YFP, al presentar las mismas bandas ya encontradas en la línea celular humana HEK293T. Además, de la carencia de estas bandas en la línea celular parental y en la población

celular mixta transfectada establemente por el vector pLV/GFP, corroboraron que no hay ningún tipo de infiltración de la proteína YFP por algún evento de muerte celular u otra molécula que está siendo sintetizada por este tipo celular que emita algún tipo de fluorescencia.

Figura 11

Secreción de la proteína YFP (~27KDa) en las células CHODG44 evaluadas por SDS-PAGE 10% en condiciones no reductoras, seguida por detección de fluorescencia por escaneamiento a laser.



Nota. A) Muestras concentradas (10x) por ultrafiltración, del medio de cultivo condicionado por 48 horas por las células CHO-DG44 transfectadas establemente por lipofección con los vectores pLV/spYFP_DHFR (clon CR5 - carril CR5), pLV/ilYFP_DHFR(carril IL) y pLV/abYFP_DHFR(carril AB), fueron fraccionadas por SDS-PAGE 10% en condiciones no reductoras, seguida por detección de la emisión de fluorescencia a 540nm por escaneamiento a laser (530nm). Como control negativo, se empleó el medio de cultivo condicionado por 48h por las células CHO-DG44 no sometidas al procedimiento de transfección (carril HEK), o por las células CHO-DG44 transfectadas con el vector de la proteína verde fluorescente no secretada GFP, pLV/GFP (carril GFP). (B) Coloración con Coomassie blue coloidal del gel de acrilamida después del escaneamiento a laser. L: Padrón de peso molecular.

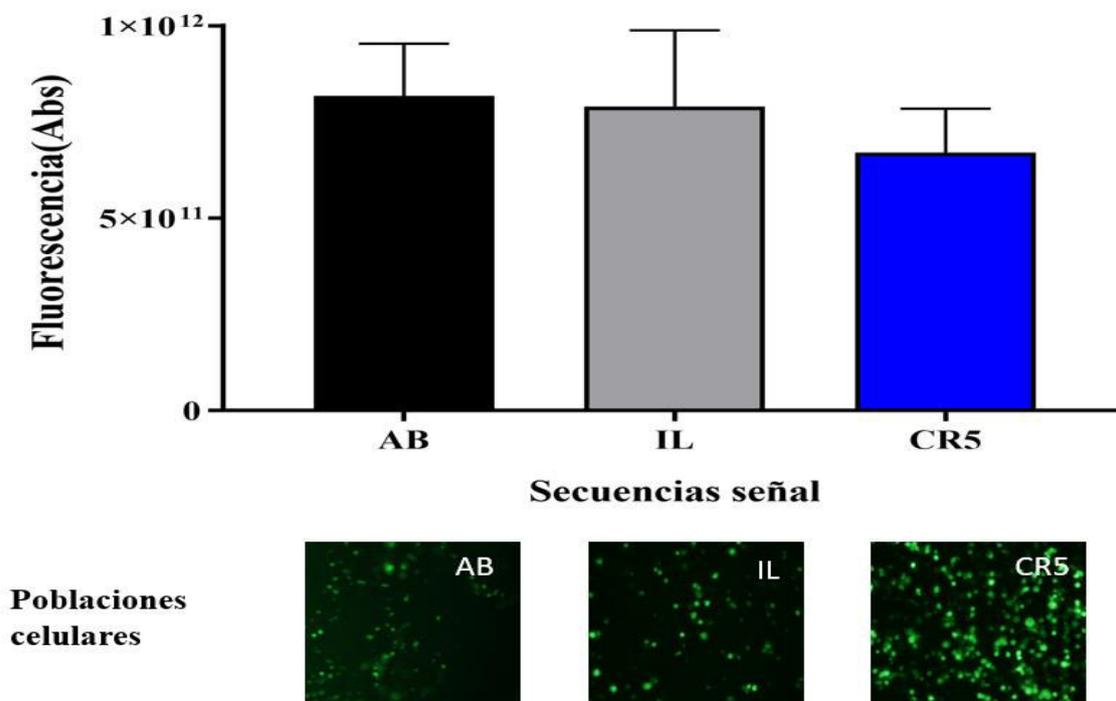
4.5 Evaluación de los niveles de secreción de la proteína YFP en las poblaciones celulares mixtas generadas por transfección transitoria y estable.

A partir de los medios de cultivo colectados en los ensayos de secreción de las poblaciones celulares mixtas generadas por transfección transitoria y estable, se analizó los

niveles de secreción de la proteína YFP por fluorometría (Anexo 2). En el caso de las muestras colectadas por la transfección transitoria, como se muestra en la Figura 12, no se observó ninguna diferencia estadísticamente significativa en los niveles de secreción de la proteína YFP entre las muestras proveniente del clon CR5 del vector pLV/spYFP_DHFR, modificado por la secuencia señal del factor VIII de coagulación humana y las muestras proveniente de las poblaciones celulares transfectadas transitoriamente con los vectores que contenían las secuencias señal de la interleucina 2 y la albumina. De las muestras analizadas aquella correspondiente a la población celular generada por la transfección con el vector pLV/abYFP_DHFR, que contenía la secuencia señal de la albumina, presento un ligero aumento en los niveles de secreción al respecto de las otras dos poblaciones celulares evaluadas (absorbancia de 8.2×10^{11}), cabe resaltar que al observar las monocapas celulares por microscopia de fluorescencia al momento de colectar el medio de cultivo condicionado, estas células mostraron los menores niveles de fluorescencia respecto a las otras poblaciones celulares. Por el contrario, los menores niveles de secreción de la proteína YFP fueron de la población celular generada por el clon CR5 del vector pLV/spYFP_DHFR (absorbancia de 6.7×10^{11}), y la fluorescencia de sus células evaluadas por microscopia de fluorescencia, mostraron la mayor intensidad al ser comparada con las otras dos poblaciones celulares. Aun con estas ligeras variaciones de una población celular a otra, no generaron una diferencia estadísticamente significativa al ser analizadas por el teste de ANOVA de una vía y posteriormente el teste de Tukey de múltiple comparación.

Figura 12

Evaluación de los niveles de secreción de la proteína YFP de las poblaciones celulares mixtas generadas por transfección transitoria en las células HEK 293T.



Nota. Los niveles de secreción de cada población celular mixta generada por transfección transitoria a través del método de lipofectamina, fue evaluada a partir del medio de cultivo condicionado colectado después de 72 horas y concentrado en un factor de 10X mediante ultrafiltración en microcolumnas Amicon de 10 KDa. Seguidamente, estas muestras fueron evaluadas por fluorometría en el equipo SpectraMax® Paradigm® con filtro de excitación 510 nm y filtro de emisión 530 nm, siendo los valores de la absorbancia normalizados por la relación del número de celulares finales e iniciales. Estos valores fueron analizados por el teste de ANOVA de una vía y posteriormente el teste de Tukey de múltiple comparación. En la parte inferior se muestran las monocapas celulares de estas poblaciones celulares al momento de colectar las muestras evaluadas a través del microscopio de fluorescencia. AB: Población celular mixta generada por la transfección transitoria del vector que contenía la secuencia señal de la albumina, pLV/abYFP_DHFR. IL: Población celular mixta generada por la transfección transitoria del vector que contenía la secuencia señal de la interleucina 2, pLV/ilYFP_DHFR. CR5: Población celular mixta generada por la transfección transitoria del vector que contenía la secuencia señal del factor VIII de coagulación humana, clon CR5 del vector pLV/spYFP_DHFR. Las barras indican los niveles de secreción de la proteína YFP cuantificada a partir de su absorbancia (Abs) y las barras de error la desviación estándar (n = 3). Significación estadística: ** p < 0.05 o * p < 0.001 (ANOVA seguido de Tukey).

Datos similares fueron encontrados en la evaluación de los niveles de secreción de la proteína YFP en las poblaciones celulares generadas por la transfección estable en las células

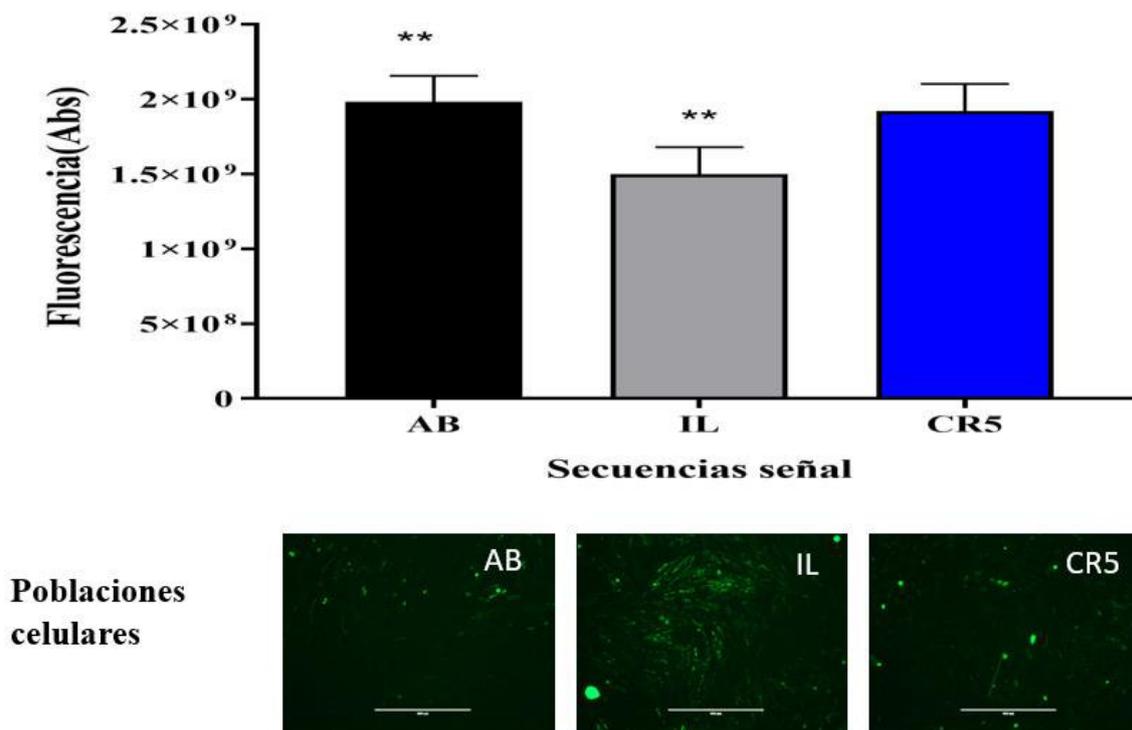
CHO DG44, donde el clon CR5 del vector pLV/spYFP_DHFR no mostro ninguna diferencia estadísticamente significativa al ser comparado con las muestras de los medios de cultivo condicionado de las otras poblaciones celulares evaluadas, aunque con este análisis fue posible detectar una diferencia estadística entre los niveles de secreción de la secuencia señal de la interleucina 2 y la albumina, que no fue posible detectar en el ensayo de transfección transitoria. Al igual que en la población transfectada transitoriamente, las células generadas por la transfección estable con el vector pLV/abYFP_DHFR, que contenía la secuencia señal de la albumina, mostro un ligero aumento en los niveles de secreción de la proteína YFP(absorbancia de 1.9×10^9), y la intensidad de la fluorescencia de sus células evaluadas a través de microscopia de fluorescencia, fue similar a la emitida por las células generadas por la transfección estable con el clon CR5 del vector pLV/spYFP_DHFR, que además presento niveles de secreción de la proteína YFP bastante similares a esta población celular. Por el contrario, la población celular generada por la transfección estable con el vector pLV/ilYFP_DHFR, que contenía la secuencia señal de la interleucina 2, mostro una disminución en los niveles de secreción de la proteína YFP (absorbancia de 1.5×10^9), y una intensidad mayor en la fluorescencia emitida por sus células con respecto las otras poblaciones celulares, como es observado en Figura 13.

Al comparar los valores de absorbancia de la fluorescencia de la proteína YFP en las poblaciones celulares mixtas generadas por ambos métodos de transfección, es notorio una disminución en los niveles de secreción de la proteína YFP en las poblaciones celulares generadas por el método de transfección estable, de alrededor de dos órdenes de grandeza, que los obtenidos por las poblaciones celulares mixtas generada por transfección transitorias, como era de esperarse. Además, los niveles de secreción de la secuencia señal de la albumina fueron ligeramente mayores que las otras secuencias señal en ambas metodologías de transfección. La evaluación de la intensidad de la fluorescencia en las monocapas celulares por microscopia de fluorescencia de ambos métodos de transfección, evidencio una disminución en la intensidad

de fluorescencia de la proteína YFP de las poblaciones celulares transfectadas establemente, similar a lo detectado en los niveles de secreción de la proteína YFP.

Figura 13

Evaluación de los niveles de secreción de la proteína YFP de las poblaciones celulares mixtas generadas por transfección estable en la línea celular CHO-DG44.



Nota. Los niveles de secreción de cada población celular mixta generada por transfección estable a través del método de lipofectamina, fue evaluada a partir del medio de cultivo condicionado colectado después de 72 horas y concentrado en un factor de 10X mediante ultrafiltración en microcolumnas Amicon de 10 KDa. Seguidamente, estas muestras fueron evaluadas por fluorometría en el equipo SpectraMax® Paradigm® con filtro de excitación 510 nm y filtro de emisión 530 nm, siendo los valores de la absorbancia normalizados por la relación del número de celulares finales e iniciales. Estos valores fueron analizados por el teste de ANOVA de una vía y posteriormente el teste de Tukey de múltiple comparación. En la parte inferior se muestran las monocapas celulares de estas poblaciones celulares al momento de colectar las muestras, evaluadas a través de microscopia de fluorescencia. AB: Población celular mixta generada por la transfección estable del vector que contenía la secuencia señal de la albumina, pLV/abYFP_DHFR. IL: Población celular mixta generada por la transfección estable del vector que contenía la secuencia señal de la interleucina 2, pLV/ilYFP_DHFR. CR5: Población celular mixta generada por la transfección estable del vector que contenía la secuencia señal del factor VIII de coagulación humana, clon CR5 del vector pLV/spYFP_DHFR. Las barras indican los niveles de secreción de la proteína YFP cuantificada a partir de su absorbancia (Abs) y las barras de error la desviación estándar (n = 3). Significación estadística: ** p < 0.05 o * p < 0.001 (ANOVA seguido de Tukey).

V. DISCUSIÓN DE RESULTADOS

Los desafíos asociados al desarrollo de una plataforma de expresión usando células de mamíferos involucra procedimientos laboriosos y extensos, de búsqueda de eventos tanto aleatorios como favorables de la inserción genómica, así como de la elección de fenotipos celulares compatibles con la manutención de las características estructurales y funcionales de la proteína de interés, cuando es producida en grandes cantidades(Chusainow *et al.*, 2009). En este sentido, se torna relevante el establecimiento de nuevas metodologías que permitan desarrollar en menor tiempo los clones celulares superproductores y estables(Kim *et al.*, 2012).

Cuando se requiere obtener los máximos niveles de productividad de una proteína recombinante, es necesario la optimización en cada uno de los siguientes procesos: control transcripcional, control de procesamiento del ARNm, exportación del ARNm desde el núcleo, transporte y dirección de ARNm, control de la estabilidad del ARNm, control de la traducción y secreción(Stern *et al.*, 2007). Como alternativa para reducir las ineficiencias asociadas en las etapas del procesamiento de la proteína, específicamente en la etapa de secreción de la proteína emergente, es que nuestro objetivo fue demostrar la capacidad de la secuencia señal de la cadena leve del factor VIII de coagulación humana, para poder secretar proteínas heterólogas en la plataforma de expresión de las células de mamíferos, para lo cual, se creó una variante secretoria de una proteína repórtor, la proteína amarilla fluorescente, que fue usada como proteína modelo para este estudio y cuya expresión estaba ligada a la expresión del marcador de selección, la proteína DHFR. La fusión de la región codificadora de estas proteínas fue realizado a través de mutación sitio-dirigida por PCR. Inicialmente, se generaron una gran cantidad de clones bacterianos recombinantes, por este motivo que se realizó una búsqueda inicial por PCR, para analizar la mayor cantidad de muestras, y así, seleccionar los clones más prominentes. Las colonias bacterianas recombinantes analizadas por PCR directamente de la

placa de LB mostraron muchos productos inespecíficos, por este motivo se dejaron incubando por 24 horas en medio LB + ampicilina. Siguiendo esta estrategia se obtuvo un resultado satisfactorio y en aquellas colonias bacterianas recombinante que dieron positivo en las PCR, se realizó el análisis de la fidelidad e integridad de su secuencia de DNA a través del perfil con enzimas de restricción, fue así que se seleccionaron los 6 clones recombinantes del vector pLV/spYFP_DHFR (CR01, CR05, CR20, CR36, CR60 y CR82) y que posteriormente fueron secuenciados por el método de Sanger. Los datos obtenidos en la secuenciación de estos clones recombinantes fueron alineados en la herramienta BLAST, demostrando que varios clones presentaron *gaps* en diferentes regiones de su secuencia de ADN, estas particularidades explicarían porque algunos clones bacterianos recombinantes no dieron positivo en la PCR, posibles pérdidas o adición de nucleótidos en las regiones donde los *primers* se alinean a su DNA molde ocasionaría la falta de amplificación. La presencia de las colonias en la placa de LB aun con la presencia de estos *gaps*, sería debido a que el elemento usado para seleccionar las bacterias es la ampicilina y estos vectores al conservar la funcionalidad de resistencia a este antibiótico permitiría su supervivencia. De los clones evaluados, solo 3 presentaron la mayor similitud con la secuencia teórica, y por presentar el mayor valor, se seleccionó el clon CR5. Este clon recombinante fue usado en los próximos experimentos que evalúan la capacidad de secreción de la proteína YFP de la secuencia señal del factor VIII con respecto a los otros vectores que contenían las secuencias señal usadas como comparativo.

Se sabe que el nivel de expresión relativo de un determinado transgen de interés es variado entre los diferentes subclones que componen una determinada población celular mixta (Girod & Mermoud, 2003). Esta heterogeneidad con respecto a los niveles de expresión se pudo evidenciar, aunque parcialmente, en los diferentes experimentos de microscopía de fluorescencia (Figuras 8 y 10), donde al observar las monocapas celulares es posible detectar pequeñas variaciones en la intensidad de fluorescencias de una célula a otra. Tal vez estas

pequeñas variaciones estén relacionada a eventos como, inserciones múltiples, muchas veces organizadas como *loci* complejos, o inserciones transgénicas reordenadas, y la localización genómica de integración, como descrito en la literatura (Barnes *et al.*, 2003; Frenzel *et al.*, 2013b)

Estos resultados fueron respaldados por la detección de la proteína YFP a través del fraccionamiento de las muestras de medio condicionado por SDS-PAGE 10%, seguido por detección de la fluorescencia por escaneamiento a laser del gel. Con el empleo de estas metodologías, se constató que la proteína YFP era secretada para el medio de cultivo de las células HEK-293T y CHODG44 transfectadas tanto transitoriamente como establemente con el vector pLV-spEYFP-DHFR (Figuras 9 y 11).

La vía de secreción de las proteínas es particularmente importante en la biotecnología y la industria biofarmacéutica, ya que la mayoría de las proteínas terapéuticas se producen en células de mamíferos (Gutierrez *et al.*, 2020). Estudios anteriores en células de mamíferos han identificado la vía secretora como un factor limitante en los rendimientos de producción de la proteína recombinante, debido a que la tasa de translocación del retículo endoplasmático es un factor crítico de control en la ruta biosintética del producto (Haryadi *et al.*, 2015). Por lo tanto, en estas células productoras de proteínas recombinantes es importante garantizar que el ARNm que codifica la proteína recombinante en cuestión, sea direccionado de manera eficiente a la población de polisomas correspondiente para su mayor eficiencia en la secreción del producto de la célula, función que es encargada por la secuencia señal (Stern *et al.*, 2007). La eficacia de la secreción para el caso de las proteínas heterólogas no está asegurada por la secuencia señal natural, por lo cual es necesario la identificación de la secuencia señal ideal para cada proteína objetivo, y así, obtener los mayores niveles de secreción en el sobrenadante de la cultura celular (Zhang y Leng, 2005). La manera más promisor de encontrar la secuencia señal ideal para una proteína recombinante deseada es rastreando la mayor diversidad posible de

secuencias señal, para eso se han ideado varias estrategias, como la variación de su secuencia genética usando grandes bibliotecas de secuencia señal o por estrategias de modificación de manera direccionada o aleatoriamente (Freudl, 2018). Un puñado de secuencias señal muy eficientes han podido ser descrito a través de estas estrategias, dentro de este grupo, tenemos a las secuencias señal de la interleucina-2, la cadena ligera de inmunoglobulina Kappa, tripsinógeno, albúmina, prolactina y muchas otras que también han demostrado ser beneficiosas (Stern *et al.*, 2007; Frenzel *et al.*, 2013; Kober *et al.*, 2013). Una reciente propuesta para encontrar nuevas secuencias señal altamente eficientes, es a través de la experimentación con secuencias señal de proteínas asociadas a una alta capacidad de secreción, para su uso en el mejoramiento de los rendimientos de la producción de proteínas recombinantes (Güler-Gane *et al.*, 2016).

Los ensayos iniciales para determinar los niveles de secreción de la proteína YFP por fluorometría a partir de los medios de cultivo condicionados colectados utilizando los medios de cultivo recomendado para cada línea celular, el medio D-MEM para las células HEK 293T y el medio α -MEM para las células CHO-DG44 (usado para el seleccionamiento de la población transfectada establemente a través del marcador DHFR), los cuales no permitieron la detección de la fluorescencia de la proteína YFP en las muestras analizadas. Se desconoce los motivos exactos que dificultaron la detección de la fluorescencia de la proteína YFP directamente en el medio de cultivo condicionado, pero se presume que sea debido a una probable supresión (“quenching”) de la señal fluorescente de esta proteína por algún compuesto presente en el medio de cultivo DMEM y α -MEM (Lakowicz, 1983). La detección por fluorometría de la secreción de la proteína fluorescente YFP solo se tornó posible después de una adaptación de las células usadas en los experimentos al medio de cultivo RPMI, que carece de coloración, y concentrándolos previamente en el equipo Speed Vac, antes de su análisis.

En los experimentos de secreción fue evidenciado las variaciones en cuanto a los niveles de secreción de la proteína YFP, entre el clone CR5 de pLV/speYFP_DHFR y las otras dos secuencias señal usadas de referencia (Figuras 11 y 12). Esto está relacionado a las diferencias entre las secuencias de aminoácidos que presentan cada una de ellas, lo que genera una afinidad diferente con el SRP y, por lo tanto, una variación en la eficiencia con la que el polipéptido naciente entra en la vía secretoria (Zheng y Gierasch, 1996). Sin embargo, el análisis comparativo de los niveles de secreción de la proteína YFP a través del test de ANOVA de una vía y posteriormente el test de Tukey de múltiple comparación, entre las secuencias señal del factor VIII, la interleucina 2 y la albumina, no presentaron ninguna diferencia estadísticamente significativa, tanto por el método de transfección transitoria como estable. Al comparar los niveles de secreción entre estas estrategias de transfección es notorio la diferencia entre ambas, siendo de alrededor de dos órdenes de magnitud. Como el método de transfección estable requiere una integración exitosa del ADN en el genoma del huésped, a menudo es mucho más difícil de lograr que la transfección transitoria y, por lo general, tiene una menor eficiencia de transfección (Fus-kujawa *et al.*, 2021). Además, en este ensayo de secreción se evidenció un ligero aumento en la secreción de la proteína YFP por parte de la población celular que fue transfectada con el vector que contenía la secuencia señal de la albumina con respecto a las otras secuencias señal analizadas, tanto en los experimentos de transfección transitoria como estable. Esta secuencia señal ya ha sido descrita anteriormente por poseer una alta capacidad de secreción para otras líneas celulares (Kober *et al.* 2013; Knappskog *et al.* 2007), aunque también fue descrito como una secuencia señal débil en las células CHO (Hesketh *et al.*, 2005). Este relativo aumento tal vez sea debido a que la región codificante de la secuencia señal de la albumina no afecta a la estabilidad del ARNm cuando está presente en una construcción quimérica (Partridge *et al.* 1999). En el caso de la proteína albúmina, esta se expresa de manera constitutiva y su secreción se produce como un flujo constante hacia la

sangre desde las células hepáticas para reemplazar la eliminada de la circulación (Knappskog *et al.*, 2007). La proteína interleucina-2, por otro lado, se produce y secreta de una manera "similar a una explosión" tras la estimulación apropiada de los leucocitos (Barash *et al.*, 2002). Estas características en la síntesis de estas proteínas fueron determinantes a la hora de seleccionar estas secuencias señal como padrones comparativos para nuestro estudio, ya que esta alta eficiencia debe estar asociada de cierta forma a la capacidad de su secuencia señal de poder dirección de manera eficiente a las proteínas al medio extracelular (Knappskog *et al.*, 2007).

Aunque, los métodos para el aumento de los niveles de producción del factor VIII recombinante se basan en el uso de estrategias como, la amplificación del número de copias del gen, la adición de agentes estabilizantes en el medio de cultivo, la maximización de la densidad celular por crecimiento en biorreactor en un sistema continuo en batelada o el uso de vectores lentivirales para aumentar la eficiencia de transferencia genética (Saenko *et al.*, 2003; Spencer *et al.*, 2011). Estos estudios no se han enfocado en la eficiencia de secreción de su secuencia señal natural, sin embargo, se ha demostrado que su uso mejora los niveles de producción al comparado con los métodos de extracción del factor VIII a partir de plasma (Yonemura *et al.*, 1993). Por otra parte, al comparado con otros sistemas de expresión heterólogos, los niveles de secreción han resultado ser de 2 de 3 órdenes de magnitud menor que las observadas con otras proteínas recombinantes que utilizan las mismas estrategias de expresión en células de mamífero (Kaufman *et al.*, 1988). Esta disminución en los niveles de secreción, posiblemente esté relacionada a que en estos sistemas de expresión esta forma recombinante presenta un peso molecular ligeramente superior al de su forma natural (Kaufman *et al.*, 1997). A pesar de todo esto, por todos los resultados de los experimentos realizados, se pudo definir que la secuencia señal de la cadena leve del factor VIII presenta en las líneas celulares de HEK 293T y CHO DG44 una eficiencia de secreción similar al de la

secuencia señal de albumina y de la interleucina 2, reconocidas por su capacidad de secretar con gran eficacia en la naturaleza y en sistema de expresión heterólogos.

VI. CONCLUSIÓN

El presente trabajo demostró que la secuencia señal de la cadena leve del factor VIII es una alternativa interesante para su uso en la expresión de proteínas recombinantes en las líneas celulares de mamíferos, HEK 293T y CHO DG44, al tener una capacidad de secreción similar al de las secuencias señal de la interleucina 2 y la albumina humana.

VII. RECOMENDACIONES

- Baja eficiencia y la formación de productos inespecíficos en la PCR se pueden superar aumentando la temperatura de alineamiento con su DNA molde.
- Productos de PCR extraídos a partir del gel de agarosa es recomendable concentrarlos al vacío antes de usarlos en un nuevo experimento.
- Al no obtener amplificación de una PCR de una colonia extraída directamente de una placa de LB, es posible superar algunas interferencias de la reacción de PCR incubando la colonia por 24 horas medio LB con su antibiótico correspondiente.
- En los cultivos celulares es mejor mantenerlos sin antibiótico para evitar arrastrar alguna contaminación previa, que puede estar camuflada por el uso del mismo y que puede afectar en los resultados de los próximos experimentos.
- Es recomendable realizar análisis constantes de la presencia de micoplasma por PCR nested en todos los cultivos de células utilizados, para evitar su contaminación y que puedan interferir en el metabolismo de las células y por consiguiente en los resultados de los análisis.
- El congelamiento de las células para su conservación debe ser realizada de forma gradual y su descongelamiento debe ser rápida para evitar el estrés de las células.
- El análisis de absorbancia de la fluorescencia de una proteína fluorescente a partir de un medio de cultivo, es recomendable centrifugar la muestra previamente para evitar algún contaminante que pueda interferir en la lectura

VIII. REFERENCIAS

- Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, *37*(4), 420–423. <https://doi.org/10.1038/s41587-019-0036-z>.
- Altamirano, C., Berrios, J., Vergara, M., & Becerra, S. (2013). Advances in improving mammalian cells metabolism for recombinant protein production. *Electronic Journal of Biotechnology*, *16*(3). <https://doi.org/10.2225/vol16-issue3-fulltext-2>.
- Attallah, C., Etcheverrigaray, M., Kratje, R., & Oggero, M. (2017). A highly efficient modified human serum albumin signal peptide to secrete proteins in cells derived from different mammalian species. *Protein Expression and Purification*, *132*, 27–33. <https://doi.org/10.1016/j.pep.2017.01.003>.
- Babu, K. R., Swaminathan, S., Marten, S., Khanna, N., & Rinas, U. (2000). Production of interferon- α in high cell density cultures of recombinant *Escherichia coli* and its single step purification from refolded inclusion body proteins. *Applied Microbiology and Biotechnology*, *53*(6), 655–660. <https://doi.org/10.1007/s002530000318>.
- Bandaranayake, A. D., & Almo, S. C. (2014). Recent advances in mammalian protein production. En *FEBS Letters* (Vol. 588, Número 2, pp. 253–260). <https://doi.org/10.1016/j.febslet.2013.11.035>.
- Barash, S., Wang, W., & Shi, Y. (2002). Human secretory signal peptide description by hidden Markov model and generation of a strong artificial signal peptide for secreted protein expression. *Biochemical and Biophysical Research Communications*, *294*(4), 835–842. [https://doi.org/10.1016/S0006-291X\(02\)00566-1](https://doi.org/10.1016/S0006-291X(02)00566-1).

- Barnes, L. M., Bentley, C. M., & Dickson, A. J. (2003). Stability of protein production from recombinant mammalian cells. *Biotechnology and Bioengineering*, 81(6), 631–639. <https://doi.org/10.1002/bit.10517>
- Baser, B., Spehr, J., Büssow, K., & van den Heuvel, J. (2015). A method for specifically targeting two independent genomic integration sites for co-expression of genes in CHO cells. *Methods*, 95, 3–12. <https://doi.org/10.1016/j.ymeth.2015.11.022>.
- Belin, D., Bost, S., Vassalli, J. D., & Strub, K. (1996). A two-step recognition of signal sequences determines the translocation efficiency of proteins. *EMBO Journal*, 15(3), 468–478. <https://doi.org/10.1002/j.1460-2075.1996.tb00379.x>.
- Birch, J. R., & Onakunle, Y. (2005). Biopharmaceutical Proteins: Opportunities and Challenges. En C. M. Smales & D. C. James (Eds.), *Therapeutic Proteins* (pp. 001–016). Humana Press. <https://doi.org/10.1385/1-59259-922-2:001>.
- Bollin, F., Dechavanne, V., & Chevalet, L. (2011). Design of Experiment in CHO and HEK transient transfection condition optimization. *Protein Expression and Purification*, 78(1), 61–68. <https://doi.org/10.1016/j.pep.2011.02.008>.
- Bushman, F. (2007). Retroviral integration and human gene therapy. *The Journal of clinical investigation*, 117(8), 1–4. <https://doi.org/10.1172/JCI32949.based>.
- Cheng, L., Fu, J., Tsukamoto, A., & Hawley, R. G. (1996). Use of green fluorescent protein variants to monitor gene transfer and expression in mammalian cells. *Nature Biotechnology*, 14(5), 606–609. <https://doi.org/10.1038/nbt0596-606>.
- Chow, Y. T., Chen, S., Wang, R., Liu, C., Kong, C. W., Li, R. A., Cheng, S. H., & Sun, D. (2016). Single cell transfection through precise microinjection with quantitatively controlled injection volumes. *Scientific Reports*, 6(March), 1–9.

<https://doi.org/10.1038/srep24127>.

Clérico, E. M., Maki, J. L., & Gierasch, L. M. (2008). Use of synthetic signal sequences to explore the protein export machinery. *Peptide Science*, *90*(3), 307–319.

<https://doi.org/10.1002/bip.20856>.

Chusainow, J., Yang, Y. S., Yeo, J. H. M., Ton, P. C., Asvadi, P., Wong, N. S. C., & Yap, M. G. S. (2009). A study of monoclonal antibody-producing CHO cell lines: What makes a stable high producer? *Biotechnology and Bioengineering*, *102*(4), 1182–1196.

<https://doi.org/10.1002/bit.22158>

Dalton, A. C., & Barton, W. A. (2014). Over-expression of secreted proteins from mammalian cell lines. *Protein Science*, *23*(5), 517–525. <https://doi.org/10.1002/pro.2439>.

Darzacq, X., Singer, R. H., & Shav-Tal, Y. (2005). Dynamics of transcription and mRNA export. *Current Opinion in Cell Biology*, *17*(3), 332–339.

<https://doi.org/10.1016/j.ceb.2005.04.004>.

Derouazi, M., Martinet, D., Besuchet Schmutz, N., Flaction, R., Wicht, M., Bertschinger, M., Hacker, D. L., Beckmann, J. S., & Wurm, F. M. (2006). Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. *Biochemical and Biophysical Research Communications*, *340*(4), 1069–1077.

<https://doi.org/10.1016/j.bbrc.2005.12.111>.

Driessen, A. J. M., & van der Does, C. (2002). Protein Export in Bacteria. En R. E. Dalbey & G. B. T.-P. T. von Heijne Transport, and Translocation (Eds.), *Protein Targeting, Transport, and Translocation* (pp. 47–73). Elsevier. <https://doi.org/10.1016/B978-012200731-6.50006-9>.

Elgundi, Z., Sifniotis, V., Reslan, M., Cruz, E., & Kayser, V. (2017). Laboratory scale

- production and purification of a therapeutic antibody. *Journal of Visualized Experiments*. <https://doi.org/10.3791/55153>.
- Elslinger, M. A., Wachter, R. M., Hanson, G. T., Kallio, K., & Remington, S. J. (1999). Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry*, *38*(17), 5296–5301. <https://doi.org/10.1021/bi9902182>.
- Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J., Vázquez, E., & Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories*, *8*(1), 17. <https://doi.org/10.1186/1475-2859-8-17>.
- Food and Drug Administration [FDA]. (2020). Basics - What is a biological product? [fda.gov](https://www.fda.gov/aboutfda/transparency/basics/ucm194516.htm). <https://www.fda.gov/aboutfda/transparency/basics/ucm194516.htm>.
- Frenzel, A., Hust, M., & Schirrmann, T. (2013a). Expression of recombinant antibodies. *Frontiers in Immunology*, *4*(JUL), 1–20. <https://doi.org/10.3389/fimmu.2013.00217>.
- Frenzel, A., Hust, M., & Schirrmann, T. (2013b). Identification of *Arabidopsis thaliana* transformants without selection reveals a high occurrence of silenced T-DNA integrations. *Frontiers in Immunology*, *4*(3), 464–477. <https://doi.org/10.1111/j.1365-313X.2004.02312.x>.
- Freshney, R. I. (2010). Culture of Animal Cells. En *Culture of Animal Cells: A Manual of Basic Techniques and Specialized Application* (Vol. 346). John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470649367>.
- Freudl, R. (2018). Signal peptides for recombinant protein secretion in bacterial expression systems. *Microbial Cell Factories*, *17*(1), 1–10. <https://doi.org/10.1186/s12934-018-0901-3>.
- Fus-kujawa, A., Prus, P., Bajdak-rusinek, K., Teper, P., Gawron, K., Kowalczyk, A., & Sieron,

- A. L. (2021). *An Overview of Methods and Tools for Transfection of Eukaryotic Cells in vitro*. 9(July), 1–15. <https://doi.org/10.3389/fbioe.2021.701031>.
- Gagneten, S., Le, Y., Miller, J., & Sauer, B. (1997). Brief expression of a GFPcre fusion gene in embryonic stem cells allows rapid retrieval of site-specific genomic deletions. *Nucleic Acids Research*, 25(16), 3326–3331. <https://doi.org/10.1093/nar/25.16.3326>.
- Gálvez, J., Lecina, M., Solà, C., Cairó, J. J., & Gòdia, F. (2012). Optimization of HEK-293S cell cultures for the production of adenoviral vectors in bioreactors using on-line OUR measurements. *Journal of Biotechnology*, 157(1), 214–222. <https://doi.org/10.1016/J.JBIOTEC.2011.11.007>.
- Girod, P., & Mermoud, N. (2003). Use of scaffold / matrix-attachment regions for protein production. *Gene Transfer and Expression in Mammalian Cells*, i, 359–379.
- Güler-Gane, G., Kidd, S., Sridharan, S., Vaughan, T. J., Wilkinson, T. C. I., & Tigue, N. J. (2016). Overcoming the refractory expression of secreted recombinant proteins in mammalian cells through modification of the signal peptide and adjacent amino acids. *PLoS ONE*, 11(5), 1–15. <https://doi.org/10.1371/journal.pone.0155340>.
- Gutierrez, J. M., Feizi, A., Li, S., Kallehauge, T. B., Hefzi, H., Grav, L. M., Ley, D., Baycin Hizal, D., Betenbaugh, M. J., Voldborg, B., Fastrup Kildegaard, H., Min Lee, G., Palsson, B. O., Nielsen, J., & Lewis, N. E. (2020). Genome-scale reconstructions of the mammalian secretory pathway predict metabolic costs and limitations of protein secretion. *Nature Communications*, 11(1), 1–10. <https://doi.org/10.1038/s41467-019-13867-y>.
- Hacker, D. L., & Balasubramanian, S. (2016). Recombinant protein production from stable mammalian cell lines and pools. *Current Opinion in Structural Biology*, 38(Figure 1), 129–136. <https://doi.org/10.1016/j.sbi.2016.06.005>.

- Harter, C., & Wieland, F. (1996). The secretory pathway: mechanisms of protein sorting and transport. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, 1286(2), 75–93. [https://doi.org/10.1016/0304-4157\(96\)00003-2](https://doi.org/10.1016/0304-4157(96)00003-2).
- Haryadi, R., Ho, S., Kok, Y. J., Pu, H. X., Zheng, L., Pereira, N. A., Li, B., Bi, X., Goh, L. T., Yang, Y., & Song, Z. (2015). Optimization of heavy chain and light chain signal peptides for high level expression of therapeutic antibodies in CHO cells. *PLoS ONE*, 10(2), 1–16. <https://doi.org/10.1371/journal.pone.0116878>.
- Hegde, R. S., & Bernstein, H. D. (2006). The surprising complexity of signal sequences. *Trends in Biochemical Sciences*, 31(10), 563–571. <https://doi.org/10.1016/j.tibs.2006.08.004>.
- Hesketh, J., Ravneberg, H., Gjerdrum, C., Tauler, A., Pryme, I., Stern, B. (2005). Protein expression system(U.S. Patent No: WO 2005/001099 A2). U.S. Patent and Trademark Office.
- Jayaraman, S., Haggie, P., Wachter, R. M., Remington, S. J., & Verkman, A. S. (2000). Mechanism and cellular applications of a green fluorescent protein- based halide sensor. *J Biol Chem*, 275(9), 6047-50. <https://doi.org/10.1074/jbc.275.9.6047>.
- Jozala, A. F., Geraldles, D. C., Tundisi, L. L., Feitosa, V. de A., Breyer, C. A., Cardoso, S. L., Mazzola, P. G., Oliveira-Nascimento, L. de, Rangel-Yagui, C. de O., Magalhães, P. de O., Oliveira, M. A. de, & Pessoa, A. (2016). Biopharmaceuticals from microorganisms: from production to purification. *Brazilian Journal of Microbiology*, 47(October), 51–63. <https://doi.org/10.1016/j.bjm.2016.10.007>.
- Kaderbhai, N. N., Harding, V., & Kaderbhai, M. A. (2008). Signal peptidase I-mediated processing of an engineered mammalian cytochrome b5 precursor is an exocytoplasmic post-translocational event in Escherichia coli. *Molecular Membrane Biology*, 25(5), 388–399. <https://doi.org/10.1080/09687680802154799>.

- Kaufman, R. J., Wasley, L. C., & Dorner, A. J. (1988). Synthesis, processing, and secretion of recombinant human factor VIII expressed in mammalian cells. *Journal of Biological Chemistry*, 263(13), 6352–6362. [https://doi.org/10.1016/s0021-9258\(18\)68793-9](https://doi.org/10.1016/s0021-9258(18)68793-9).
- Kaufman, R. J., Pipe, S. W., Tagliavacca, L., Swaroop, M., & Moussalli, M. (1997). Biosynthesis, assembly and secretion of coagulation factor VIII. *Blood Coagulation & Fibrinolysis : An International Journal in Haemostasis and Thrombosis*, 8 Suppl 2, S3-14.
- Kaufman, Randal J. (1990). [42] Selection and coamplification of heterologous genes in mammalian cells. In *Methods in enzymology* (Vol. 185, pp. 537–566). [https://doi.org/10.1016/0076-6879\(90\)85044-O](https://doi.org/10.1016/0076-6879(90)85044-O).
- Kim, J. Y., Kim, Y.-G. G., & Lee, G. M. (2012a). CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Applied microbiology and biotechnology*, 93(3), 917–930. <https://doi.org/10.1007/s00253-011-3758-5>.
- Kim, J. Y., Kim, Y. G., & Lee, G. M. (2012b). CHO cells in biotechnology for production of recombinant proteins: Current state and further potential. *Applied Microbiology and Biotechnology*, 93(3), 917–930. <https://doi.org/10.1007/s00253-011-3758-5>.
- Kim, T. K., & Eberwine, J. H. (2010). *Mammalian cell transfection : the present and the future*. 3173–3178. <https://doi.org/10.1007/s00216-010-3821-6>.
- Kim, Y.-G., Han, Y. K., Kim, J. Y., Lee, E. G., Lee, H. W., & Lee, G. M. (2011). Effect of constitutively active ras overexpression on cell growth in recombinant chinese hamster ovary cells. *Biotechnology Progress*, 27(2), 577–580. <https://doi.org/10.1002/btpr.567>.
- Knappskog, S., Ravneberg, H., Gjerdrum, C., Tröbse, C., Stern, B., & Pryme, I. F. (2007). The level of synthesis and secretion of Gaussia princeps luciferase in transfected CHO cells

- is heavily dependent on the choice of signal peptide. *Journal of Biotechnology*, 128(4), 705–715. <https://doi.org/10.1016/j.jbiotec.2006.11.026>.
- Kober, L., Zehe, C., & Bode, J. (2012). Development of a novel ER stress based selection system for the isolation of highly productive clones. *Biotechnology and Bioengineering*, 109(10), 2599–2611. <https://doi.org/10.1002/bit.24527>.
- Kober, L., Zehe, C., & Bode, J. (2013). Optimized signal peptides for the development of high expressing CHO cell lines. *Biotechnology and Bioengineering*, 110(4), 1164–1173. <https://doi.org/10.1002/bit.24776>.
- Kojima, T., & Kitamura, T. (1999). A signal sequence trap based on a constitutively active cytokine receptor. *Nature Biotechnology*, 17(5), 487–490. <https://doi.org/10.1038/8666>.
- Komada, F., Nishiguchi, K., Tanigawara, Y., Iwakawa, S., & Okumura, K. (1999). Effects of secreted SOD delivered by genetically modified cells on xanthine/xanthine oxidase and paraquat-induced cytotoxicity in vitro. *Biological & Pharmaceutical Bulletin*, 22(8), 846–853. <https://doi.org/10.1248/bpb.22.846>.
- Kuystermans, D., Krampe, B., Swiderek, H., & Al-Rubeai, M. (2007). Using cell engineering and omic tools for the improvement of cell culture processes. *Cytotechnology*, 53(1–3), 3–22. <https://doi.org/10.1007/s10616-007-9055-6>.
- Lakowicz, J. R. (1983). Quenching of Fluorescence BT - Principles of Fluorescence Spectroscopy (J. R. Lakowicz (ed.); pp. 257–301). Springer US. https://doi.org/10.1007/978-1-4615-7658-7_9
- La Merie Business Intelligence. (2020). Blockbuster Biologics 2018: Sales of Recombinant Therapeutic Antibodies & Proteins. lamerie.com.

<https://lamerie.com/report/blockbuster-biologics-2019-sales-of-recombinant-therapeutic-antibodies-proteins/>.

Lee, Y. C., Kim, C. H., & Tsuji, S. (1996). An efficient expression vector for extracellular secretion in mammalian cells. *Molecules and Cells*, 6(5), 552–556.

Lewis, N. E., Liu, X., Li, Y., Nagarajan, H., Yerganian, G., O'Brien, E., Bordbar, A., Roth, A. M., Rosenbloom, J., Bian, C., Xie, M., Chen, W., Li, N., Baycin-Hizal, D., Latif, H., Forster, J., Betenbaugh, M. J., Famili, I., Xu, X., ... Palsson, B. O. (2013). Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. *Nature biotechnology*, 31(8). <https://doi.org/10.1038/nbt.2624>.

Lufino, M. M. P., Edser, P. A. H., & Wade-Martins, R. (2008). Advances in high-capacity extrachromosomal vector technology: Episomal maintenance, vector delivery, and transgene expression. *Molecular Therapy*, 16(9), 1525–1538. <https://doi.org/10.1038/mt.2008.156>.

Mathew, J. (2022). *Production of Recombinant Proteins by R-DNA Technology*. 14(1000505), 1000505. <https://doi.org/10.35248/0974-8369.22.14.505>.

Mei, B., Chen, Y., Chen, J., Pan, C. Q., & Murphy, J. E. (2006). *FVIII Expression in HKB11 Cells 165 MOLECULAR BIOTECHNOLOGY Expression of Human Coagulation Factor VIII in a Human Hybrid Cell Line, HKB11. 34.* <https://link.springer.com/content/pdf/10.1385%2FMB%3A34%3A2%3A165.pdf>.

Murshid, A., & Presley, J. F. (2004). ER-to-Golgi transport and cytoskeletal interactions in animal cells. *Cellular and Molecular Life Sciences CMLS*, 61(2), 133–145. <https://doi.org/10.1007/s00018-003-3352-9>.

Nagano, R., & Masuda, K. (2014). Establishment of a signal peptide with cross-species

- compatibility for functional antibody expression in both *Escherichia coli* and Chinese hamster ovary cells. *Biochemical and Biophysical Research Communications*, 447(4), 655–659. <https://doi.org/10.1016/j.bbrc.2014.04.060>.
- Negahdaripour, M., Nezafat, N., Hajighahramani, N., & Soheil, S. (2017). *In Silico Study of Different Signal Peptides for Secretory Production of Interleukin-11 in Escherichia coli*. 1–10. <https://doi.org/10.2174/1570164614666170106110>.
- Nejepinska, J., Malik, R., Moravec, M., & Svoboda, P. (2012). Deep sequencing reveals complex spurious transcription from transiently transfected plasmids. *PLoS ONE*, 7(8). <https://doi.org/10.1371/journal.pone.0043283>.
- Orlova, N. A., Kovnir, S. V., Gabibov, A. G., & Vorobiev, I. I. (2017). Stable high-level expression of factor VIII in Chinese hamster ovary cells in improved elongation factor-1 alpha-based system. *BMC Biotechnology*, 17(1), 1–15. <https://doi.org/10.1186/s12896-017-0353-6>.
- Paetzel, M. (2014). Structure and mechanism of *Escherichia coli* type I signal peptidase. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1843(8), 1497–1508. <https://doi.org/10.1016/j.bbamcr.2013.12.003>.
- Partridge, K. A., Johannessen, A., Tauler, A., Pryme, I. F., & Hesketh, J. E. (1999). Competition between the signal sequence and a 3'UTR localisation signal during redirection of beta-globin mRNA to the endoplasmic reticulum: Implications for biotechnology. *Cytotechnology*, 30(1–3), 37–47. <https://doi.org/10.1023/a:1008079901508>.
- Puck, T.T. (1985). Development of the Chinese Hamster Ovary (CHO) Cell for Use in Somatic Cell Genetics. In *Molecular Cell Genetics*; Gottesman, M.M., Ed.; John Wiley and Sons: New York, NY, USA, 1985:37–64.

- Remington, S. J. (2011). Green fluorescent protein: A perspective. *Protein Science*, 20(9), 1509–1519. <https://doi.org/10.1002/pro.684>.
- Román, R., Miret, J., Scalia, F., Casablanco, A., Lecina, M., & Cairó, J. J. (2016). Enhancing heterologous protein expression and secretion in HEK293 cells by means of combination of CMV promoter and IFN α 2 signal peptide. *Journal of Biotechnology*, 239, 57–60. <https://doi.org/10.1016/j.jbiotec.2016.10.005>.
- Rout, M. P., & Aitchison, J. D. (2001). The Nuclear Pore Complex as a Transport Machine. *Journal of Biological Chemistry*, 276(20), 16593–16596. <https://doi.org/10.1074/jbc.R100015200>.
- Saenko, E. L., Ananyeva, N. M., Shima, M., Hauser, C. A. E., & Pipe, S. W. (2003). The future of recombinant coagulation factors. *Journal of Thrombosis and Haemostasis*, 1(5), 922–930. <https://doi.org/10.1046/j.1538-7836.2003.00196.x>
- Saguez, C., Olesen, J. R., & Jensen, T. H. (2005). Formation of export-competent mRNP: escaping nuclear destruction. *Current Opinion in Cell Biology*, 17(3), 287–293. <https://doi.org/10.1016/J.CEB.2005.04.009>.
- Spencer, H. T., Denning, G., Gautney, R. E., Dropulic, B., Roy, A. J., Baranyi, L., Gangadharan, B., Parker, E. T., Lollar, P., & Doering, C. B. (2011). Lentiviral vector platform for production of bioengineered recombinant coagulation factor VIII. *Molecular Therapy*, 19(2), 302–309. <https://doi.org/10.1038/mt.2010.239>
- Stepanenko, A. A., & Heng, H. H. (2017). Transient and stable vector transfection: Pitfalls, off-target effects, artifacts. *Mutation Research/Reviews in Mutation Research*, 773, 91–103. <https://doi.org/10.1016/J.MRREV.2017.05.002>.
- Stern, B., Olsen, L. C., Trösse, C., Ravneberg, H., & Pryme, I. F. (2007). *Improving*

mammalian cell factories : The selection of signal peptide has a major impact on recombinant protein synthesis and secretion in mammalian cells.

- Suzuki, K., Nakazato, H., Matsui, H., Hasumi, M., Shibata, Y., Ito, K., Fukabori, Y., Kurokawa, K., & Yamanaka, H. (2001). NK cell-mediated anti-tumor immune response to human prostate cancer cell, PC-3: immunogene therapy using a highly secretable form of interleukin-15 gene transfer. *Journal of Leukocyte Biology*, 69(4), 531–537.
- Toole, J. J., Knopf, J. L., Wozney, J. M., Sultzman, L. A., Buecker, J. L., Pittman, D. D., Kaufman, R. J., Brown, E., Shoemaker, C., Orr, E. C., Amphlett, G. W., Foster, W. B., Coe, M. Lou, Knutson, G. J., Fass, D. N., & Hewick, R. M. (1984). Molecular cloning of a cDNA encoding human antihemophilic factor. *Nature*, 312(5992), 342–347. <https://doi.org/10.1038/312342a0>.
- Tsien, R. Y. (1998). The Green Fluorescent Protein. *Annu. Rev. Biochem.*, 67, 509–544. <https://doi.org/10.1146/annurev.biochem.67.1.509>.
- Urlaub, G., & Chasin, L. A. (1980). Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proceedings of the National Academy of Sciences of the United States of America*, 77(7), 4216–4220. <https://doi.org/10.1073/pnas.77.7.4216>.
- Verrijzer, C. P., Chen, J. L., Yokomori, K., & Tjian, R. (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell*, 81(7), 1115–1125. [https://doi.org/10.1016/S0092-8674\(05\)80016-9](https://doi.org/10.1016/S0092-8674(05)80016-9).
- Von Heijne, G. (1990). The signal peptide. *The Journal of Membrane Biology*, 115(3), 195–201. <https://doi.org/10.1007/BF01868635>.
- Walsh, G. (2018). Biopharmaceutical benchmarks 2018. *Nature Biotechnology*, 36(12), 1136–

1145. <https://doi.org/10.1038/nbt.4305>.

Wen, B., Deng, Y., Guan, J., Yan, W., Wang, Y., Tan, W., & Gao, J. (2011). Signal peptide replacements enhance expression and secretion of hepatitis C virus envelope glycoproteins. *Acta Biochimica et Biophysica Sinica*, 43(2), 96–102. <https://doi.org/10.1093/abbs/gmq117>.

Wood, W. I., Capon, D. J., Simonsen, C. C., Eaton, D. L., Gitschier, J., Keyt, B., Seeburg, P. H., Smith, D. H., Hollingshead, P., Wion, K. L., Delwart, E., Tuddenham, E. G. D., Vchar, G. A., & Lawn, R. M. (1984). Expression of active human factor VIII from recombinant DNA clones. *Nature*, 312(5992), 330–337. <https://doi.org/10.1038/312330a0>.

Wurm, F. M. (2014). *Production of recombinant protein therapeutics in cultivated mammalian cells*. June. <https://doi.org/10.1038/nbt1026>.

Wurm, F., & Wurm, M. (2017). Cloning of CHO Cells, Productivity and Genetic Stability—A Discussion. *Processes*, 5(4), 20. <https://doi.org/10.3390/pr5020020>.

Yonemura, H., Sugawara, K., Nakashima, K., Nakahara, Y., Hamamoto, T., Mimaki, I., Yokomizo, K., Tajima, Y., Masuda, K., Imaizumi, A., et al. (1993) Efficient production of recombinant human factor VIII by co-expression of the heavy and light chains. *Journal of Biological Chemistry*, 268(12), 8485–8490.

Chusainow, J., Yang, Y. S., Yeo, J. H. M., Ton, P. C., Asvadi, P., Wong, N. S. C., & Yap, M. G. S. (2009). A study of monoclonal antibody-producing CHO cell lines: What makes a stable high producer? *Biotechnology and Bioengineering*, 102(4), 1182–1196. <https://doi.org/10.1002/bit.22158>

Young, R & Rance, J. (2008). Mammalian expression vector with a highly efficient secretory signal sequence. (U.S. Patent No:WO 2008/148519 A2). U.S. Patent and Trademark Office.

Zhang, L., Leng, Q., & Mixson, A. J. (2005). Alteration in the IL-2 signal peptide affects secretion of proteins in vitro and in vivo. *The Journal of Gene Medicine*, 7(3), 354–365. <https://doi.org/https://doi.org/10.1002/jgm.677>.

Zheng, N., & Gierasch, L. M. (1996). Signal sequences: The same yet different. *Cell*, 86(6), 849–852. [https://doi.org/10.1016/S0092-8674\(00\)80159-2](https://doi.org/10.1016/S0092-8674(00)80159-2).

IX. ANEXOS

9.1. Anexos A

Figura 14

Alineamiento del ensamblaje de la secuencia del clon bacteriano CRI con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST

Sequences producing significant alignments		Download	Select columns	Show	100			
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> None provided		3840	3840	100%	0.0	98.93%	2146	Query_5698991

Sequence ID: **Query_5698991** Length: **2146** Number of Matches: **1**

Range 1: 1 to 2146 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
3840 bits(2079)	0.0	2127/2150(99%)	4/2150(0%)	Plus/Plus

```

Query 1      CCCGGGATCCACCGGTCGCCACCATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCC 60
Sbjct 1      CCCGGGATCCACCGGTCGCCACCATGCAAATAGAGCTCTCCACCTGCTACTTTCTGTGCC 60

Query 61     TTTTGCATTCTGCTTTAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGC 120
Sbjct 61     TTTTGCATTCTGCTTTAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGC 120

Query 121    CCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTTCAAGCTGTCGGCGAGG 180
Sbjct 121    CCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTTCAAGCTGTCGGCGAGG 180

Query 181    GCGAGGGCGATGCCACCTACGGCAAGCTGACCTTGAAGTTTCAATCTGCACCACGGCAAGC 240
Sbjct 181    GCGAGGGCGATGCCACCTACGGCAAGCTGACCTTGAAGTTTCAATCTGCACCACGGCAAGC 240

Query 241    TGCCCGTGCCCTGGCCCAACCTCTGTCACCACTTTCGGCTACGGCTGCAAGTGTTCGCCC 300
Sbjct 241    TGCCCGTGCCCTGGCCCAACCTCTGTCACCACTTTCGGCTACGGCTGCAAGTGTTCGCCC 300

Query 301    GGTACCCCGACACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG 360
Sbjct 301    GGTACCCCGACACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG 360

Query 361    TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA 420
Sbjct 361    TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTAGAAGACCCGCGCCGAGGTGA 420

Query 421    AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG 480
Sbjct 421    AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG 480

Query 481    ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAACGCTATATCA 540
Sbjct 481    ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAACGCTATATCA 540

Query 541    TGGCCGACAAGCAGAAGAAGCGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGG 600
Sbjct 541    TGGCCGACAAGCAGAAGAAGCGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGG 600

Query 601    ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACCACCCATCGGGCAGCGCCCGC 660
Sbjct 601    ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACCACCCATCGGGCAGCGCCCGC 660

Query 661    TGCTGCTGCCCGACAACCACTACCTGAGTACCAGTCCGCCCTGAGCAAGACCCCAACG 720
Sbjct 661    TGCTGCTGCCCGACAACCACTACCTGAGTACCAGTCCGCCCTGAGCAAGACCCCAACG 720

Query 721    AGAAGCGCGATACATGGTCTGTCTGGAGTTCTGTGACCGCCGCGGGGATCACTCTCGGCA 780
Sbjct 721    AGAAGCGCGATACATGGTCTGTCTGGAGTTCTGTGACCGCCGCGGGGATCACTCTCGGCA 776

Query 781    TGGACGAGCTGTACAAGTAAAGCGCCGCTCGAGATCGATATCTGCGGCTAGCTAGCG 840
Sbjct 777    TGGACGAGCTGTACAAGTAAAGCGCCGCTCGAGATCGATATCTGCGGCTAGCTAGCG 836

Query 841    TTTAAACGCATCATCTAGGGCGGCAATTCCGCCCCCTCTCCCCCCCCCTCTCCCTC 900
Sbjct 837    TTTAAACGCATCATCTAGGGCGGCAATTCCGCCCCCTCTCCCCCCCCCTCTCCCTC 896

Query 901    CCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGAATAAAGGCCGGTGTGCGTTTGTCTA 960
Sbjct 897    CCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGAATAAAGGCCGGTGTGCGTTTGTCTA 956

```

```

Query 961  TATGTTATTTTCCACCATATTGCCGCTTTTGGCAATGTGAGGGCCCGGAAACCTGGGCC 1020
Sbjct 957  TATGTTATTTTCCACCATATTGCCGCTTTTGGCAATGTGAGGGCCCGGAAACCTGGGCC 1016

Query 1021 TGTCTTCTTGACGAGCATTCTAGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTCT 1080
Sbjct 1017 TGTCTTCTTGACGAGCATTCTAGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTCT 1076

Query 1081 GTTGAATGTCGTGAAGGAAGCAGTTCCTCGGAAGTCTTGAAGACAAACACGCTCTGT 1140
Sbjct 1077 GTTGAATGTCGTGAAGGAAGCAGTTCCTCGGAAGTCTTGAAGACAAACACGCTCTGT 1136

Query 1141 AGCGACCTTTGCAGGCAGCGGAACCCCACTGGCGACAGGTGCTCTGCGGCCAAAA 1200
Sbjct 1137 AGCGACCTTTGCAGGCAGCGGAACCCCACTGGCGACAGGTGCTCTGCGGCCAAAA 1196

Query 1201 GCCACGTGTATAAGATACACCTGCAAAAGCGGCACAACCCAGTGCACGTTGTGAGTTG 1260
Sbjct 1197 GCCACGTGTATAAGATACACCTGCAAAAGCGGCACAACCCAGTGCACGTTGTGAGTTG 1256

Query 1261 GATAGTTGTGAAAGAGTCAAATGGCTCTCTCAAGCGTATTCAACAAGGGCTGAAGGA 1320
Sbjct 1257 GATAGTTGTGAAAGAGTCAAATGGCTCTCTCAAGCGTATTCAACAAGGGCTGAAGGA 1316

Query 1321 TGCCCAAGAGTACCCATTGTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTAC 1380
Sbjct 1317 TGCCCAAGAGTACCCATTGTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTAC 1376

Query 1381 ATGTGTTAGTCGAGGTTAAAAAACGCTAGGCCCCCGAACACGCGGGACGTGGTTTT 1440
Sbjct 1377 ATGTGTTAGTCGAGGTTAAAAAACGCTAGGCCCCCGAACACGCGGGACGTGGTTTT 1436

Query 1441 CCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTGA 1500
Sbjct 1437 CCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTGA 1496

Query 1501 ACTGCATCGTCGCCGTGTCCAAAAATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1560
Sbjct 1497 ACTGCATCGTCGCCGTGTCCAAAAATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1556

Query 1561 CTCGGCTCAGGAACGAGTCAAGTACTTCCAAAGAATGACCACAACCTCTTCAGTGGGAG 1620
Sbjct 1557 CTCGGCTCAGGAACGAGTCAAGTACTTCCAAAGAATGACCACAACCTCTTCAGTGGGAG 1616

Query 1621 GTAACAGAAATCTGGTGATTATGGGTAGGAAAACTGGTCTCCATTCTGAGAAGAATC 1680
Sbjct 1617 GTAACAGAAATCTGGTGATTATGGGTAGGAAAACTGGTCTCCATTCTGAGAAGAATC 1676

Query 1681 GACCTTTAAAGGACAGAATTAATATAGTCTCAGTAGAGAACTCAAAGAACCAACGAG 1740
Sbjct 1677 GACCTTTAAAGGACAGAATTAATATAGTCTCAGTAGAGAACTCAAAGAACCAACGAG 1736

Query 1741 GAGCTCATTCTTTCGCAAAAGTTGGATGATGCTTAAGACTTATTGAACAACCGGAAT 1800
Sbjct 1737 GAGCTCATTCTTTCGCAAAAGTTGGATGATGCTTAAGACTTATTGAACAACCGGAAT 1796

Query 1801 TGGCAAGTAAAGTAGACATGGTTGGATAGTCGGAGGCAGTCTGTTTACCAGGAAGCA 1860
Sbjct 1797 TGGCAAGTAAAGTAGACATGGTTGGATAGTCGGAGGCAGTCTGTTTACCAGGAAGCA 1856

Query 1861 TGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1920
Sbjct 1857 TGAATCATCCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1916

Query 1921 ACACGTTTTTCCAGAAATTGATTTGGGAAATATAAACTTCTCCAGAATACCCAGGCG 1980
Sbjct 1917 ACACGTTTTTCCAGAAATTGATTTGGGAAATATAAACTTCTCCAGAATACCCAGGCG 1976

Query 1981 TCCCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAG 2040
Sbjct 1977 TCCCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAG 2036

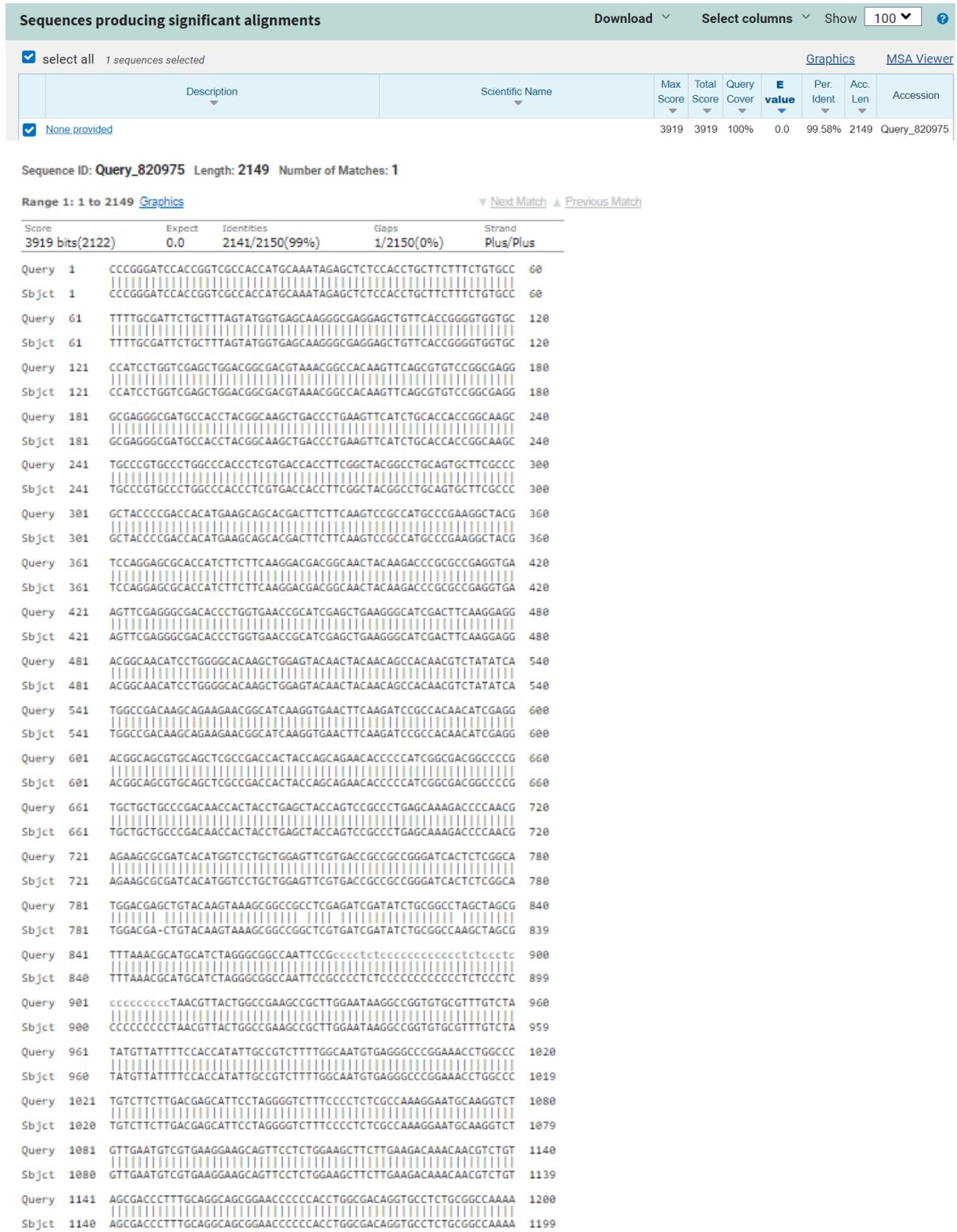
Query 2041 AAGACTAACAGGAAGATGCTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2100
Sbjct 2037 AAGACTAACAGGAAGATGCTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2096

Query 2101 ACTCCTCAGGTGACGGCTGCCTATCAGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2150
Sbjct 2097 ACTCCTCAGGTGACGGCTGCCTATCAGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2146

```

Figura 15

Alineamiento del ensamblaje de la secuencia del clon bacteriano CR5 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST



```

Query 1201  GCCACGTGTATAAGATACACCTGCAAAGGCCGCACAACCCAGTGCCACGTTGTGAGTTG 1260
Sbjct 1200  GCCACGTGTATAAGATACACCTGCAAAGGCCGCACAACCCAGTGCCACGTTGTGAGTTG 1259
Query 1261  GATAGTTGTGGAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGA 1320
Sbjct 1260  GATAGTTGTGGAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGA 1319
Query 1321  TGCCCAGAAGGTACCCCATTTGTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTAC 1380
Sbjct 1320  TGCCCAGAAGGTACCCCATTTGTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTAC 1379
Query 1381  ATGTGTTTAGTCGAGGTTAAAAAAGCTAGGCCCCCGAACACGGGGACGTGGTTTT 1440
Sbjct 1380  ATGTGTTTAGTCGAGGTTAAAAAAGCTAGGCCCCCGAACACGGGGACGTGGTTTT 1439
Query 1441  CCTTTGAAAAACACGATGATAAGCTTGCACAACCCGGGCCATCATGGTTCGACCATTTGA 1500
Sbjct 1440  CCTTTGAAAAACACGAGATTAGCTACCGACAACCCGGGCCATCATGGTTCGACCATTTGA 1499
Query 1501  ACTGCATCGTCGCGCTGTCCCAAAATATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1560
Sbjct 1500  ACTGCATCGTCGCGCTGTCCCAAAATATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1559
Query 1561  CTCGGCTCAGGAACGAGTTCAAGTACTTCCAAGAATGACCACAACCTCTTCAGTGGGAAG 1620
Sbjct 1560  CTCGGCTCAGGAACGAGTTCAAGTACTTCCAAGAATGACCACAACCTCTTCAGTGGGAAG 1619
Query 1621  GTAACAGAACTCTGGTGATTATGGGTAGGAAAACCTGGTTCTCCATTCTGAGAAGAATC 1680
Sbjct 1620  GTAACAGAACTCTGGTGATTATGGGTAGGAAAACCTGGTTCTCCATTCTGAGAAGAATC 1679
Query 1681  GACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAACTCAAAGAACCACACGAG 1740
Sbjct 1680  GACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAACTCAAAGAACCACACGAG 1739
Query 1741  GAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGACTTATTGAACAACCGGAAT 1800
Sbjct 1740  GAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGACTTATTGAACAACCGGAAT 1799
Query 1801  TGGCAAGTAAAGTAGACATGGTTGGATAGTCGGAGGCAGTTCTGTTTACCAGGAAGCCA 1860
Sbjct 1800  TGGCAAGTAAAGTAGACATGGTTGGATAGTCGGAGGCAGTTCTGTTTACCAGGAAGCCA 1859
Query 1861  TGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1920
Sbjct 1860  TGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1919
Query 1921  ACACGTTTTTCCAGAAAATTGATTTGGGGAATATAAACTTCTCCAGAAATACCCAGGCG 1980
Sbjct 1920  ACACGTTTTTCCAGAAAATTGATTTGGGGAATATAAACTTCTCCAGAAATACCCAGGCG 1979
Query 1981  TCCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGA 2040
Sbjct 1980  TCCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGA 2039
Query 2041  AAGACTAACAGGAAGATGCTTTCAAGTTCTCTGCTCCCCCTCAAAGCTATGCATCTAGC 2100
Sbjct 2040  AAGACTAACAGGAAGATGCTTTCAAGTTCTCTGCTCCCCCTCAAAGCTATGCATCTAGC 2099
Query 2101  ACTCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGGGTCGACGGG 2150
Sbjct 2100  ACTCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGGGTCGACGGG 2149

```

Figura 16

Alineamiento del ensamblaje de la secuencia del clon bacteriano CR20 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST

Sequences producing significant alignments		Download	Select columns	Show	100			
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> None provided		2741	2741	100%	0.0	90.05%	2102	Query_1552303

Sequence ID: **Query_1552303** Length: 2102 Number of Matches: 1Range 1: 1 to 2102 [Graphics](#)[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Gaps	Strand
2741 bits(1484)	0.0	1936/2150(90%)	48/2150(2%)	Plus/Plus
Query 1	CCCCGGATCCACCGGTCCGCCACCATGCAAATAGAGCTCCACCTGCTTCTTTCTGTGCC			60
Sbjct 1	CCCCGGATCCACCGGTCCGCCACCATGCAAATAGAGCTCCACCTGCTTCTTTCTGTGCC			60
Query 61	TTTTGCGATTCTGCTTAGTATGGTGAAGGGCGAGGAGCTGTTACCGGGTGGTGC			120
Sbjct 61	TTTTGCGATTCTGCTTAGTATGGTGAAGGGCGAGGAGCTGTTACCGGTGTGCTGC			120
Query 121	CCATCCTGGTCCGAGCTGGACGGCGACGTAACGGCCACAAGTTACGCGTGTCCGGCGAGG			180
Sbjct 121	CGATCCTAGTCCGAGCTGGACGGCGACGTATACGGCCACAAGTTCTGCGTGTCCGGCGATG			180
Query 181	GCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTTCATCTGCACCACCGCAAGC			240
Sbjct 181	GCGAGGGCGGTGCCACCTATGGCAAGCTGAGCTGAAGTTTCATCTGCACCACCGTCAAGC			240
Query 241	TGCCCGTGCCTGGCCACCTCTGTGACCACTTCGGCTACGGCTGCAGTGTTCGCC			300
Sbjct 241	TGCCCGTGCCTGGCCACCTCTGTGACCTTCGGCTACGGCTGCAGTGTTCGCC			300
Query 301	GCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGGCTACG			360
Sbjct 301	GCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGGCTAGG			360
Query 361	TCCAGGAGCGCACCATCTTCAAGGACGACGGCAACTACAAGACCGCGCCGAGGTGA			420
Sbjct 361	TCCAGGAGTGCACCATCTTCAAGGACGACGGCAACTACAAGACCGCGCCGAGGTGA			420
Query 421	AGTTCGAGGGCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG			480
Sbjct 421	AGTTCGAGGTCCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTGAAGGAGG			480
Query 481	ACGGCAACATCTGGGGCACAAGCTGGAGTACAACACAGCCACAACGCTATATCA			540
Sbjct 481	ACGGCAATATCTGGGGCACAATGCTGGAGTACAACACAGCCACAACGCTATATCA			540
Query 541	TGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTCAAGATCCGCCACAACATCGAGG			600
Sbjct 541	TGGGGCACAAGCAGTAGAACGGCATCTAGGTGAAGTCAAGATCCGCCACAACATCGTGG			600
Query 601	ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGGCAGCGCCCG			660
Sbjct 601	ACGGCAGCGTGCAGCTCGCCGAGCACTACCAGCAGAACTCCCCATCGGGCAGCGCCCG			660
Query 661	TGCTGCTGCCGACAACCACTACCTGAGTACCAGTCCGCCCTGAGCAAAGACCCCAACG			720
Sbjct 661	TGCTGTTGCCGACAACCTCTACGTGAGCTATCAGTCCGCCCTGAGCAAAGACCTCAACG			720
Query 721	AGAAGCGCATCACATGGTCTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCA			780
Sbjct 721	ACAAGCGCATCACTGGTCTCTGGAGTTCGT-----T-A-TC--GGCA			762
Query 781	TGGACGAGCTGTACAAGTAAAGCGCCCGCTCGAGATCGATATCTGCGGCTAGCTAGCG			840
Sbjct 763	-GG--GA-C-GTACAATAAGGGCCCGCTTGAGATCGCTATCTGGGGCCGAGCTAGGG			817
Query 841	TTTAAACGCATGCATCTAGGGCGGCCAATTCCGccccctctccccccccccccctctccctc			900
Sbjct 818	TTTAAACGCTTGCCTTAGCGCGGCCAATCCCGCCCCACTACACTCTCGCCCTT-CCTC			876
Query 901	ccccccccTAACGTTACTGGCCGAAGCGCTTGGGAATAAGGCGGTGTGCGTTTGTCTA			960
Sbjct 877	CTCGCCGCTAGCGTTACTGGCGGAAGTCGCTTGCAATAACGCGGAGTGCCTTTGTCTT			936
Query 961	TATGTTATTTTCCACCATATTGCCGCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCC			1020
Sbjct 937	TATGATAATTACGGCCTTAGGGCGTGTAGGGAATGAGACGGCCGGAAGCTGGCCC			996
Query 1021	TGTCTTCTGACGAGCATTCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTCT			1080
Sbjct 997	TGTCTACTTGACCAGCATTGCTAGGGGTCTTCCCTCTCCCAAGGATTGCAAGCTCT			1056

```

Query 1081 GTTGAATGCTGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACACGTCTGT 1140
Sbjct 1057 GTTGAATGCTGTGAACGAAGCTGTTCTTTGGAAGCATCTTGTAGACGAACAGCGTCTGT 1116

Query 1141 AGCGACCCTTTGCGAGGACGGGAACCCCACTGGCGACAGGTGCCTCTGCGGCCAAAA 1200
Sbjct 1117 AGCCACCCTTTGCTGGCAGCGCAACCCCGCAGCTGGGGACAGGAGCCTCTGGGGCCAATA 1176

Query 1201 GCCACGTGTATAAGATACACCTGCAAAAGCGGCACAACCCCAAGTGCACGTTGTGAGTTG 1260
Sbjct 1177 GCCAGGTGTAAAAGATACATCTGCAAAACGCGGGACAACCCCAAGAGCCACGCTGTGAGTTC 1236

Query 1261 GATAGTTGTGGAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGA 1320
Sbjct 1237 GATAGTTGAGGAAGAGTCATATGGCTCGCCTCAAGGGTATTCAACAGGGGCTGATGGA 1296

Query 1321 TGCCCAAGAGGTACCCCACTTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTAC 1380
Sbjct 1297 TGCCCAAGAGGTACCCCGTTGTATGGGAACTGATCTGGGGCCTCGGTGCACATGCTTTAC 1356

Query 1381 ATGTGTTTAGTCGAGGTT#####CGTCTAGGCCCCCGAACCACGGGGACGTGTTTT 1440
Sbjct 1357 ATGTGTTTGTGCGAGGTCAACAGAACGTCGAGGCCCGCGAAGCACCGGGACCTGGATTT 1416

Query 1441 CCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTGA 1500
Sbjct 1417 GCTTTGAAAAACACGATCATCAGCGTGGCTCAACCGGGCCATCATGGTTCGACCATTG- 1475

Query 1501 ACTGCATCGTCGCCGTGCCAAAATATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1560
Sbjct 1476 -----AATATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1512

Query 1561 CTCGGCTCAGGAACGAGTTCAGTACTTCCAAGAATGACCAACCTCTTCAGTGGGAAG 1620
Sbjct 1513 CTCGGCTCAGGAACGAGTTCAGTACTTCCAAGAATGACCAACCTCTTCAGTGGGAAG 1572

Query 1621 GTAAACAGAATCTGGTGATTATGGGTAGGAAAACTGGTTCTCCATTCTGAGAAGAATC 1680
Sbjct 1573 GTAAACAGAATCTGGTGATTATGGGTAGGAAAACTGGTTCTCCATTCTGAGAAGAATC 1632

Query 1681 GACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAACTCAAAGAACCACACGAG 1740
Sbjct 1633 GACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAACTCAAAGAACCACACGAG 1692

Query 1741 GAGCTCATTTTCTTGCCAAAAGTTGGATGATGCCCTAAGACTATTGAACAACCGGAAT 1800
Sbjct 1693 GAGCTCATTTTCTTGCCAAAAGTTGGATGATGCCCTAAGACTATTGAACAACCGGAAT 1752

Query 1801 TGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGTCTGTTTACCAGGAAGCCA 1860
Sbjct 1753 TGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGTCTGTTTACCAGGAAGCCA 1812

Query 1861 TGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1920
Sbjct 1813 TGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1872

Query 1921 ACACGTTTTTCCAGAAAATTGATTTGGGGAAATATAAACTTCTCCCAGAAATACCCAGGCG 1980
Sbjct 1873 ACACGTTTTTCCAGAAAATTGATTTGGGGAAATATAAACTTCTCCCAGAAATACCCAGGCG 1932

Query 1981 TCCTCTCTGAGGTCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGA 2040
Sbjct 1933 TCCTCTCTGCGGTCAGGTTGAGAACGGCACCAAGTATAAAGTTTGAAGTCTACGAGACGA 1992

Query 2041 AAGACTAACAGGAAGATGCTTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2100
Sbjct 1993 AAGACTAACAGGAAGATGCTTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2052

Query 2101 ACTCCTCAGGTGCAGGCTGCCTATCAGAAGTGGTGGCTGGGTCGACGGG 2150
Sbjct 2053 ACTCCACAGGTGCAGCCTGCCTATCAGAAGTCTGGCTGGGTCGACGGG 2102

```

Figura 17

Alineamiento del ensamblaje de la secuencia del clon bacteriano CR36 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST

Sequences producing significant alignments		Download	Select columns	Show	100			
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> None provided		2484	2484	100%	0.0	87.86%	2106	Query_8113857

Sequence ID: Query_8113857 Length: 2106 Number of Matches: 1

Range 1: 1 to 2106 [Graphics](#)[Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
2484 bits(1345)	0.0	1889/2150(88%)	44/2150(2%)	Plus/Plus
Query 1	CCCCGGATCCACCGGTCGCCACCATGCAAATAGAGCTCTCCACCTGCTTTCTGTGCC	60		
Sbjct 1	CCCCGGATCCACCGGTCGCCACCATGCAAATAGAGCTCTCCACCTGCTTTCTGTGCC	60		
Query 61	TTTTGCGATTCTGCTTTAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGC	120		
Sbjct 61	TTTTGCGATTCTGCTTTAGTATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGC	120		
Query 121	CCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAGG	180		
Sbjct 121	CCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAGG	180		
Query 181	GCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTTCATCTGCACCACCGGCAAGC	240		
Sbjct 181	GCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTTCATCTGCACCACCGGCAAGC	240		
Query 241	TGCCCGTGCCCTGGCCACCTCGTGACCACCTTCGGCTACGGCTGCAGTCTTCGCC	300		
Sbjct 241	TGCCCGTGCCCTGGCCACCTCGTGACCACCTTCGGCTACGGCTGCAGTCTTCGCC	300		
Query 301	GCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG	360		
Sbjct 301	GCTACCCCGACCATGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCGAAGGCTA--	358		
Query 361	TCCAGGAGCGACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA	420		
Sbjct 359	-----CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA	405		
Query 421	AGTTCGAGGGCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG	480		
Sbjct 406	AGTTCGAGGGCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG	465		
Query 481	ACGGCAACATCTCGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCA	540		
Sbjct 466	ACGGCAACATCTCGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCA	525		
Query 541	TGGCCGACAAGCAGAAGAAGCGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGG	600		
Sbjct 526	TGGCCGACAAGCAGAAGAAGCGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGG	585		
Query 601	ACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACCACCCATCGGGACGGCCCG	660		
Sbjct 586	ACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACCACCCATCGGGACGGCCCG	645		
Query 661	TGCTGTGCCCGACAACCACTACCTGAGTACCAGTCCGCCCTGAGCAAGACCCCAACG	720		
Sbjct 646	TGCTGTGCCCGACAACCACTACCTGAGTACCAGTCCGCCCTGAGCAAGACCCCAACG	705		
Query 721	AGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCA	780		
Sbjct 706	AGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCA	765		
Query 781	TGGACGAGTGTACAAGTAAAGCGCCGCTCGAGATCGATATCTCGGCCCTAGTAGCG	840		
Sbjct 766	TGGACGAGTGTCTAGTTAACCAGCCGGCTAGAGTTCGTTATGTGCTCCTGGTAGCG	825		
Query 841	TTTAAACGCATGCATCTAGGGCGGCAATTCCGccccctccccccccccccctctccctc	900		
Sbjct 826	AAAAAACCCATGCTTCAAGCGGGTCAATTGCGCGGTCACTCTCCACGCGTCAAGCTC	885		
Query 901	ccccccccTAAAGTTACTGGCCGAAGCCGCTTGAATAAGGCCGGTGTGCGTTTGTCTA	960		
Sbjct 886	CCCCCGCCAAAGGTGACTGCCCAAGTCGCATGGCATAACGCCCGTCTGCCCTTCTCAA	945		
Query 961	TATGTTATTTTCCACCATATTGCCGTCTTTGGCAATGTGAGGGCCCGGAAACCTGGCCC	1020		
Sbjct 946	TTTGTTTTTTCTCCATTTTGGGTCAAAAGGCTATGTGAGGGCCCGGTAACCTGCCCC	1005		
Query 1021	TGCTTCTTGACGAGCATTCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCT	1080		
Sbjct 1006	TGCTTCTAGACCAGCAATCTACGGGTCTTTCCCTCTGGCCAAATGGAATCAAGGTGT	1065		
Query 1081	GTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAACAACGTCTGT	1140		
Sbjct 1066	GTTGTATGTCGAGAAGGATGCAGTACCACTGGAAGGTTGTTCAAGTCAAAACAAGGTCTGT	1125		

```

Query 1141 AGCGACCCCTTTGCGGCGAGCGGAACCCCACTGGGACAGGTGCCTCGCGGCCAAAA 1200
          |||
Sbjct 1126 TGCGACCCCTTTGCGGCGAGCGGTACCCCGACGTGGGACACGTGCCTCAGCG-CC---- 1180

Query 1201 GCCACGTGATAAAGATACACCTGCAAAAGCGGCACAACCCAGTGCCACGTTGTGAGTTG 1260
          |||
Sbjct 1181 -C-----CACCTGCAAAAGCGGCACACCCAGTCCACAGATGTGACTTG 1224

Query 1261 GATAGTTGTGGAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGGA 1320
          |||
Sbjct 1225 GAAAGTTGTGGATAGAGACAAATGGCTGTCTCATGCGTATCCAACAAGCGGCTGAAGCA 1284

Query 1321 TGCCGAGAAGGTACCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTAC 1380
          |||
Sbjct 1285 TGCCGAGATGGTACCCGATTGTATGCGATCTGATCTCGGGCCTCGGTGCACATGCTTTAC 1344

Query 1381 ATGTGTTTAGTCGAGGTTaaaaaaCGCTAGGCCCCCCGAACCACGGGGACGTGGTTTT 1440
          |||
Sbjct 1345 ATGTGAAAAGTCGAGGTTATAACAACGTGAGGCCGCCCAACCACCGGGACGAGTTTT 1404

Query 1441 CCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTGA 1500
          |||
Sbjct 1405 CGTTTGTAAATCAGGAAGTTACGCATGGCAGAAAGCCGGGCAATCATGGTACGACATTGC 1464

Query 1501 ACTGCATCGTCGCCGTGTCCCAAAATATGGGGATTGGCAAGAACGGGAGACCTACCTGGC 1560
          |||
Sbjct 1465 ACTGCATGGTCGCCGTGTGCAATATATGGCGATTGGGAAGAGCGGTGACCTTCCCTGCC 1524

Query 1561 CTCGGCTCAGGAACGAGTTCAAGTACTTCCAAGAATGACCACAACCTCTTCAGTGGAA 1620
          |||
Sbjct 1525 CTCGGCTCACGAACGAGATCAAGAACTTCAAAAGATTGACCACTACCTCGTCAGTGGTAG 1584

Query 1621 GTAACACAGAATCTGGTGATTATGGGTAGGAAAACCTGGTTCTCCATTCTGAGAAGAATC 1680
          |||
Sbjct 1585 GGAAA-----GGTGATCATGGGTGGGAAATCTGCTTCTCTTCTGACAAGAATG 1636

Query 1681 GACCTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAAGTCAAGAACCACCCAGAG 1740
          |||
Sbjct 1637 GACCTGTAAGGTGAGAATTAGTATAGGTCAGGAGAGATCTCAGAGAACCCTCCACGGG 1696

Query 1741 GAGCTCATTTTCTTGCCAAAAGTTGGATGATGCTTAAGACTTATTGAACAACCCGAAT 1800
          |||
Sbjct 1697 GACCTCATTTCTTCCAATAGTTTGCATGTTGCTTAAAGAGTTATTCAACAACGGGAAG 1756

Query 1801 TGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGTTCGTTTACCAGGAAGCCA 1860
          |||
Sbjct 1757 TGGTAGTAAATGAGACTTGGTTGGGATAGTCGAGGCAGACTGTTTAAACCAGCAAGCGA 1816

Query 1861 TGAATCAACAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1920
          |||
Sbjct 1817 TGATTCAACTGGCCAGCTCACACTCTTTGAGACAAGCATCATGGAGGAATGTGAAAGTC 1876

Query 1921 ACACGTTTTTCCAGAAATTGATTTGGGAAATATAAACTTCTCCAGAAATACCAGGCG 1980
          |||
Sbjct 1877 ACACGCTCTTCCAGGAATTGATTTGGGAAATATAAACTTCTCCAGAAATACCAGGCG 1936

Query 1981 TCCTCTCTGAGGTCCAGGAGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGA 2040
          |||
Sbjct 1937 TCCTCTCTGAGGTCCAGGAGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGA 1996

Query 2041 AAGACTAACAGGAAGATGCTTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2100
          |||
Sbjct 1997 AAGACTAACAGGAAGATGCTTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2056

Query 2101 ACTCCTCAGGTGCAGGCTGCCTATCAGAAGTGGTGGCTGGGTGCACGGG 2150
          |||
Sbjct 2057 ACTCCTCAGGTGCAGGCTGCCTATCAGAAGTGGTGGCTGGGTGCACGGG 2106

```

Figura 18

Alineamiento del ensamblaje de la secuencia del clon bacteriano CR60 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST

Sequences producing significant alignments		Download	Select columns	Show	100			
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> None provided		3869	3869	100%	0.0	99.21%	2142	Query_2416421

Range 1: 1 to 2142 [Graphics](#)[Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
3869 bits(2095)	0.0	2133/2150(99%)	8/2150(0%)	Plus/Plus
Query 1	CCCCGGATCCACCGGTCGCCACCATGCAAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCC			60
Sbjct 1	CCCCGGATCCACCGGTCGCCACCATGCAAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCC			60
Query 61	TTTTGCGATTCTGCTTTAGTATGGTGAAGGGCCGAGGAGCTGTTACCGGGGTGGTGC			120
Sbjct 61	TTTTGCGATTCTGCTTTAGTATGGTGAAGGGCCGAGGAGCTGTTACCGGGGTGGTGC			120
Query 121	CCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAAGCTGTCCGGCGAGG			180
Sbjct 121	CCATCCTGGTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAAGCTGTCCGGCGAGG			180
Query 181	GCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACCGCAAGC			240
Sbjct 181	GCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCACCACCGCAAGC			240
Query 241	TGCCCCTGCTGCTGGCCACCTCTGTGACACCTTCGGCTACGGCTGCAGTCTCGCCC			300
Sbjct 241	TGCCCCTGCTGCTGGCCACCTCTGTGACACCTTCGGCTACGGCTGCAGTCTCGCCC			300
Query 301	GCTACCCCGACACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG			360
Sbjct 301	GCTACCCCGACACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG			360
Query 361	TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCACTACAAGACCCGCGCCGAGGTGA			420
Sbjct 361	TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCACTACAAGACCCGCGCCGAGGTGA			420
Query 421	AGTTCGAGGGCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG			480
Sbjct 421	AGTTCGAGGGCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG			480
Query 481	ACGGCAACATCTCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAACGTCTATATCA			540
Sbjct 481	ACGGCAACATCTCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAACGTCTATATCA			540
Query 541	TGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGG			600
Sbjct 541	TGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGG			600
Query 601	ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGGCAGCGCCCCG			660
Sbjct 601	ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGGCAGCGCCCCG			660
Query 661	TGCTGCTGCCCGACAACCACCTACCTGAGCTACCAGTCCGCCCTGAGCAAGACCCCAACG			720
Sbjct 661	TGCTGCTGCCCGACAACCACCTACCTGAGCTACCAGTCCGCCCTGAGC---GACCCCAACG			717
Query 721	AGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCA			780
Sbjct 718	AGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCA			777
Query 781	TGGACGAGCTGTACAAGTAAAGCGGCGGCTCGAGATCGATATCTGGGCCCTAGCTAGCG			840
Sbjct 778	TGGACGAGCTGTACAAGTAAAGCGGCGGCTCGATATCGATATCTCGGCCCTAGCTAGCT			837
Query 841	TTTAAACGCATGCATCTAGGGCGGCAATTCCGccccctctccccccccccccctctcccc			900
Sbjct 838	TTTAAACGCATGCATCTAGGGCGGCAATTCCCCCCTCTCCCTCCCTCCCTCTCCCTC			897
Query 901	ccccccccTAAAGTACTGGCCGAAGCCGCTTGAATAAGGCCGGTGTGCGTTTGTCTA			960
Sbjct 898	CCCCCCCCCTAAGCTTACTGGCCGAAGCCGCTTGAATAAGGCCGGTGTGCGTTTGTCTA			957
Query 961	TATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCC			1020
Sbjct 958	TATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCC			1017
Query 1021	TGTCTCTTGACGAGCATTCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTCT			1080
Sbjct 1018	TGTCTCTTGACGAGCATTCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTCT			1077

```

Query 1081 GTTGAATGTCGTGAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAAACACGCTGT 1140
Sbjct 1078 GTTGAATGTCGTGAAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAAACACGCTGT 1137

Query 1141 AGCGACCCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAA 1200
Sbjct 1138 AGCGACCCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAA 1197

Query 1201 GCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCAGTGCCACGTTGTGAGTTG 1260
Sbjct 1198 GCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCAGTGCCACGTTGTGAGTTG 1257

Query 1261 GATAGTTGTGGAAAGAGTCAAATGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGA 1320
Sbjct 1258 GATAGTTGTGGAAAGAGTCAAATGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGA 1317

Query 1321 TGCCCAAGAGGTACCCCATTTGATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTAC 1380
Sbjct 1318 TGCCCAAGAGGTACCCCATTTGATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTAC 1377

Query 1381 ATGTGTTTAGTCGAGGTTAAAAAAGTCGTAGGCCCCCGAACACCGGGGACGTGGTTTT 1440
Sbjct 1378 ATGTGTTTAGTCGAGGTTAAAAAAGTCGTAGGCCCCCGAACACCGGGGACGTGGTTTT 1437

Query 1441 CCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTTGA 1500
Sbjct 1438 CCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTTGA 1497

Query 1501 ACTGCATCGTCGCCGTGTCCCAAAATATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1560
Sbjct 1498 ACTGCATCGTCGCCGTGTCCCAAAATATGGGGATTGGCAAGAACGGAGACCTACCCCTGGC 1557

Query 1561 CTCCGCTCAGGAACGAGTTCAAGTACTTCAAAGAATGACCAACCTCTTCAGTGGAAAG 1620
Sbjct 1558 CTCCGCTCAGGAACGAGTTCAAGTACTTCAAAGAATGACCAACCTCTTCAGTGGAAAG 1617

Query 1621 GTAAACAGAACTGGTGATTATGGGTAGGAAAACCTGGTCTCCATTCTGAGAAGAATC 1680
Sbjct 1618 GTAAACAGAACTGGTGATTATGGGTAGGAAAACCTGGTCTCCATTCTGAGAAGAATC 1672

Query 1681 GACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAAGCTCAAAGAACCACACGAG 1740
Sbjct 1673 GACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAAGCTCAAAGAACCACACGAG 1732

Query 1741 GAGCTCATTTTCTTGCCAAAAGTTGGATGATGCCCTTAAGACTTATTGAACAACCGGAAT 1800
Sbjct 1733 GAGCTCATTTTCTTGCCAAAAGTTGGATGATGCCCTTAAGACTTATTGAACAACCGGAAT 1792

Query 1801 TGGCAAGTAAAGTAGACATGGTTGGATAGTCGGAGGCAGTTCTGTTTACCAGGAAGCCA 1860
Sbjct 1793 TGGCAAGTAAAGTAGACATGGTTGGATAGTCGGAGGCAGTTCTGTTTACCAGGAAGCCA 1852

Query 1861 TGAATCAACAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1920
Sbjct 1853 TGAATCAACAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGTG 1912

Query 1921 ACACGTTTTTCCAGAAATGATTTGGGAAATATAAACTTCTCCAGAAATACCCAGGCG 1980
Sbjct 1913 ACACGTTTTTCCAGAAATGATTTGGGAAATATAAACTTCTCCAGAAATACCCAGGCG 1972

Query 1981 TCCTCTCTGAGGTCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGA 2040
Sbjct 1973 TCCTCTCTGAGGTCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAGA 2032

Query 2041 AAGACTAACAGGAAGATGCTTTCAAGTTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2100
Sbjct 2033 AAGACTAACAGGAAGATGCTTTCAAGTTCTCTGCTCCCTCCTAAAGCTATGCATCTAGC 2092

Query 2101 ACTCCTCAGGTGCAGGCTGCCTATCAGAAGTGGTGGCTGGTGCACGGG 2150
Sbjct 2093 ACTCCTCAGGTGCAGGCTGCCTATCAGAAGTGGTGGCTGGTGCACGGG 2142

```

Figura 19

Alineamiento del ensamblaje de la secuencia del clon bacteriano CR82 con la secuencia teórica del vector pLV/spYFP_DHFR con la herramienta BLAST

Sequences producing significant alignments		Download	Select columns	Show	100			
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> None provided		2695	2695	100%	0.0	89.54%	2118	Query_661669

Sequence ID: **Query_661669** Length: **2118** Number of Matches: **1**Range 1: 1 to 2118 [Graphics](#)[Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
2695 bits(1459)	0.0	1926/2151(90%)	34/2151(1%)	Plus/Plus
Query 1	CCCCGGATCCACCGTCCGCCACCATGCAAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCC	60		
Sbjct 1	CCCCGGATCCACCGTCCGCCACCATGCAAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCC	60		
Query 61	TTTTGCG-ATTCTGCTTTAGTATGGTGAGCAAGGGCGAGGAGCTTCCACGGGGTGGTG	119		
Sbjct 61	TTATCCGTACTCTTCTATAGTATGGTGAGCAAGGGCGAGGAGCTTCCACGGGGTGGTG	120		
Query 120	CCCATCTGGTCTGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAG	179		
Sbjct 121	CCGATGCTGCTCGAGTTGGACCCCGACGGAAACGTCCACTAGTTCGCGTGCCTGCCGAC	180		
Query 180	GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACGGCAAG	239		
Sbjct 181	GGCGTGGGGTCCACCTTCGGCTAGCTGACTCTGAACCTCATGTGCAGCACGGCAAC	240		
Query 240	CTGCCGTGCTCCGGCCACCTCGTGACCACCTTCGGCTACGGCTGCAGTGTCTCGCC	299		
Sbjct 241	CTGCCGGGGCGTCCCTCGTGTCCACGTTCCGCTTCGGCGTGCAGTGTCTCGCG	300		
Query 300	CGTACCCCGACCATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTAC	359		
Sbjct 301	CGTACCCCGTCCACATGAAGCAGCACGAGTTCATCAAGACCTCGATGCGCGAAGGGTAC	360		
Query 360	GTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG	419		
Sbjct 361	GACCAAGAGGGCAGCATCATCTTCAACGTCGACGGCTACTTCAAGTCCCGCCGACGCTG	420		
Query 420	AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAG	479		
Sbjct 421	TAGATCCAGGGCGACACCCTGGTGTATCGTATGGACCTCAACCCCATCTACTTTAAGGAT	480		
Query 480	GACGGCAACATCTGGGGCACAAGCTGGAGTACAACATAACAGCCACAACGCTATATC	539		
Sbjct 481	GACGGCTACATCTGGGGCACAAGTGTAGTACAACATAACAGCCACAACGCTATAAC	540		
Query 540	ATGGCCGACAAGCAGAAGAAGCGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAG	599		
Sbjct 541	ATGGCGACTAGCACAAGATCGGCTTCAACGTCAGTGAAGATCCGCCACAACATCGAG	600		
Query 600	GACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGGCAGGGCCCC	659		
Sbjct 601	CACGGCTGCGAGCAGCGCGCCGAGCACTAGCAGCAGTACATCCCATTTGGCCACGGCGCG	660		
Query 660	GTGCTGCTGCCGACAACCACTACCTGAGCTACCAGTCCGCCCTGAGCAAAAGACCCCAAC	719		
Sbjct 661	GTGCTCCTGCCCTACAACCACTACCTGAGTACC--T--GTCC---G---GATCCCAAC	709		
Query 720	GAGAAGCGCATCACATGGTCTGCTGGAGTTCGTGACCGCCGGGATCACTCTCGGC	779		
Sbjct 710	GAGTAGCGCATCACATGGTCTGCTGGAGTTCGTGACCTCCGGCGGATCACTCTCGGC	769		
Query 780	ATGGACGAGCTGTACAAGTAAAGCGGCCCTCGAGATCGATATCTGGGCCCTAGCTAGC	839		
Sbjct 770	ATGGTCGAGCTGTACAAGTAAAGCGGGCGGCTCGATATCGTTATCTCGGGCTAGCGAGC	829		
Query 840	GTTTAAACGCATGCATCTAGGGCGGCAATTCCGccccctccccccccccccctccccct	899		
Sbjct 830	ATTTAATCGGATCCATTTAGCGCGGCAATTTCCCTCTCTCTGCTGCCGTCCTCACCGT	889		
Query 900	ccccccccccTAACGTTACTGGCCGAAGCCGCTTGGAAATAAGGCCGGTGTGCGTTGTCT	959		
Sbjct 890	CCCCCCCCCTTAACCTTACTGGCGGAACCCGGTTGGTATATGGCGGGTGTGCGTGTCT	949		
Query 960	ATATGTTATTTCCACCATATTGCCGCTTTTGGCAATGTGAGGGCCGGAAACCTGGCC	1019		
Sbjct 950	AGATGTTAATTTCCACCATATTGCCGCTTTTGGCAATGGGAGGGCCGGAAACCTCGCC	1009		
Query 1020	CTGTCTTTCTGACGAGCATCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGGTC	1079		
Sbjct 1010	CTGTCTTATTGACGATCATCTAGGGGTCTTCCGCTCTGGCCAAATGGAAACGCAAGCTC	1069		
Query 1080	TGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAACAACGCTCG	1139		
Sbjct 1070	TGTTCAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAACAACGCTCG	1129		
Query 1140	TAGCGACCTTTGACGGCAGCGGAACCCCACTGGCGACAGGTGCTCTGCGGCCAAA	1199		
Sbjct 1130	TAGCGACCTTTGACGGCAGCGGAACCCCACTGGCGACAGGTGCTCTGCGGCCAAA	1189		

```

Query 1200 AGCCACGTGTATAAGATACACCTGCAAAAGCGGCACAAACCCAGTGCCACGTTGTGAGTT 1259
Sbjct 1190 AGCCACGTGTATAAGATACACCTGCAAAAGCGGCACAAACCCAGTGCCACGTTGTGAGTT 1249
Query 1260 GGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGG 1319
Sbjct 1250 GGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTGAAGG 1309
Query 1320 ATGCCCAGAAGGTACCCCATTTGATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTA 1379
Sbjct 1310 ATGCCCAGAAGGTACCCCATTTGATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTA 1369
Query 1380 CATGTGTTTAGTCGAGGTTAAAAAAGCTCTAGGCCCCCGAACCCAGGGGACGTGGTTT 1439
Sbjct 1370 CATGTGTTTAGTCGAGGTTAAAAAAGCTCTAGGCCCCCGAACCCAGGGGACGTGGTTT 1429
Query 1440 TCCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTG 1499
Sbjct 1430 TCCTTTGAAAAACACGATGATAAGCTTGCCACAACCCGGGCCATCATGGTTCGACCATTG 1489
Query 1500 AACTGCATCGTCGCCGTGTCCAAAATATGGGGATTTGGCAAGAACGGAGACCTACCCCTGG 1559
Sbjct 1490 AACTGCATCGTCGCCGTGTCCAAAATATGGGGATTTGGCAAGAACGGAGACCTACCCCTGG 1549
Query 1560 CCTCCGCTCAGGAACGAGTTCAGTACTTCAAAGAATGACCACAACCTCTTCAGTGGAA 1619
Sbjct 1550 CCTCCGCTCAGGAACGAGTTCAGTACTTCAAAGAATGACCACAACCTCTTCAGTGGAA 1609
Query 1620 GGTAAACAGAATCTGGTGATTATGGGTAGGAAAACCTGGTTCCTCATTCTGAGAAGAAT 1679
Sbjct 1610 GGTAAACAGAATCTGGTGATTATGGGTAGGAAAACCTGGT----- 1658
Query 1680 CGACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAACTCAAAGAACCACCACGA 1739
Sbjct 1651 ---CCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAACTCAAAGAACCACCACGA 1707
Query 1740 GGAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGACTTATTGAACAACCGGAA 1799
Sbjct 1708 GGAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGACTTATTGAACAACCGGAA 1767
Query 1800 TTGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGTTCGTGTTACCAGGAAGCC 1859
Sbjct 1768 TTGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGTTCGTGTTACCAGGAAGCC 1827
Query 1860 ATGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGT 1919
Sbjct 1828 ATGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATCATGCAGGAATTTGAAAGT 1887
Query 1920 GACACGTTTTTCCAGAAAATGATTTGGGGAAAATATAAACTTCTCCAGAAATACCCAGGC 1979
Sbjct 1888 GACACGTTTTTCCAGAAAATGATTTGGGGAAAATATAAACTTCTCCAGAAATACCCAGGC 1947
Query 1980 GTCCTCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAG 2039
Sbjct 1948 GTCCTCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAGTTTGAAGTCTACGAGAAG 2007
Query 2040 AAAGACTAACAGGAAGATGCTTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAG 2099
Sbjct 2008 AAAGACTAACAGGAAGATGCTTTCAAGTCTCTGCTCCCTCCTAAAGCTATGCATCTAG 2067
Query 2100 CACTCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGGGTGCACGGG 2150
Sbjct 2068 CACTCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGGGTGCACGGG 2118

```

9.2. anexos B

Tabla 1

Evaluación de los niveles de secreción de la proteína EYFP en las poblaciones celulares mixtas generadas por transfección transitoria en las células HEK293T

Repetición/población celular	Recuento de células finales (10 ⁴ Células /ml)			Proporción de células finales/células iniciales *			Absorbancia			Valores normalizados		
	CR5	IL	AB	CR5	IL	AB	CR5	IL	AB	CR5	IL	AB
1	32	39.6	39.3855	3.2	3.96	3.93855	2.70E+11	2.40E+11	1.95E+11	8.64E+11	9.50E+11	7.68E+11
2	31.612	37.95	39.6	3.1612	3.795	3.96	2.10E+11	1.50E+11	1.38E+11	6.64E+11	5.69E+11	5.46E+11
3	32.998	34.1	42.56	3.2998	3.41	4.256	2.80E+11	2.50E+11	1.64E+11	9.24E+11	8.52E+11	6.98E+11

*la cantidad inicial de células fue de 10 x 10⁴ células/ml

Ordinary one-way ANOVA						
ANOVA results						
1	Table Analyzed	Data 1				
2	Data sets analyzed	A-C				
3						
4	ANOVA summary					
5	F	0.7772				
6	P value	0.5010				
7	P value summary	ns				
8	Significant diff. among means (P < 0.05)?	No				
9	R square	0.2058				
10						
11	Brown-Forsythe test					
12	F (DFn, DFd)	0.1865 (2, 6)				
13	P value	0.8345				
14	P value summary	ns				
15	Are SDs significantly different (P < 0.05)?	No				
16						
17	Bartlett's test					
18	Bartlett's statistic (corrected)					
19	P value					
20	P value summary					
21	Are SDs significantly different (P < 0.05)?					
22						
23	ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
24	Treatment (between columns)	3.656e+022	2	1.828e+022	F (2, 6) = 0.7772	P=0.5010
25	Residual (within columns)	1.411e+023	6	2.352e+022		
26	Total	1.777e+023	8			
27						
28	Data summary					
29	Number of treatments (columns)	3				
30	Number of values (total)	9				

Ordinary one-way ANOVA									
Multiple comparisons									
1	Number of families	1							
2	Number of comparisons per family	3							
3	Alpha	0.05							
4									
5	Tukey's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adjusted P Value			
6	AB vs. IL	27000000000	-357200577933 to 411200577933	No	ns	0.9748	A-B		
7	AB vs. CR5	146666666667	-237533911266 to 530867244599	No	ns	0.5103	A-C		
8	IL vs. CR5	119666666667	-264533911266 to 503867244599	No	ns	0.6284	B-C		
9									
10	Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	n1	n2	q	DF
11	AB vs. IL	817333333333	790333333333	27000000000	125217144726	3	3	0.3049	6
12	AB vs. CR5	817333333333	670666666667	146666666667	125217144726	3	3	1.656	6
13	IL vs. CR5	790333333333	670666666667	119666666667	125217144726	3	3	1.352	6

Tabla 2:

Evaluación de los niveles de secreción de la proteína EYFP en las poblaciones celulares mixtas generadas por transfección estable en las células CHODG44

Repetición/población celular	Recuento de células finales (10 ⁴ Células /ml)			Proporción de células finales/células iniciales *			Absorbancia			Valores normalizados		
	CR5	AB	IL	CR5	AB	IL	CR5	AB	IL	CR5	AB	IL
1	27.37	30.42	33.729	2.737	3.042	3.3729	7.09E+08	5.89E+08	3.89E+08	1.94E+09	1.79E+09	1.31E+09
2	26.23	32.282	35.703	2.623	3.2282	3.5703	6.60E+08	6.60E+08	4.70E+08	1.73E+09	2.13E+09	1.68E+09
3	27.9	35	34.2	2.79	3.5	3.42	7.49E+08	5.79E+08	4.39E+08	2.09E+09	2.03E+09	1.50E+09

*la cantidad inicial de células fue de 10 x 10⁴ células/ml

Ordinary one-way ANOVA						
ANOVA results						
1	Table Analyzed	teste CHO				
2	Data sets analyzed	A-C				
3						
4	ANOVA summary					
5	F	6.573				
6	P value	0.0308				
7	P value summary	*				
8	Significant diff. among means (P < 0.05)?	Yes				
9	R square	0.6866				
10						
11	Brown-Forsythe test					
12	F (DFn, DFd)	0.005945 (2, 6)				
13	P value	0.9941				
14	P value summary	ns				
15	Are SDs significantly different (P < 0.05)?	No				
16						
17	Bartlett's test					
18	Bartlett's statistic (corrected)					
19	P value					
20	P value summary					
21	Are SDs significantly different (P < 0.05)?					
22						
23	ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
24	Treatment (between columns)	4.190e+017	2	2.095e+017	F (2, 6) = 6.573	P=0.0308
25	Residual (within columns)	1.912e+017	6	3.187e+016		
26	Total	6.103e+017	8			
27						
28	Data summary					
29	Number of treatments (columns)	3				
30	Number of values (total)	9				

Ordinary one-way ANOVA									
Multiple comparisons									
1	Number of families	1							
2	Number of comparisons per family	3							
3	Alpha	0.05							
4									
5	Tukey's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adjusted P Value			
6	AB vs. IL	485785543	38522676 to 933048409	Yes	*	0.0363	A-B		
7	AB vs. CR5	62565046	-384697820 to 509827912	No	ns	0.9050	A-C		
8	IL vs. CR5	-423220497	-870483363 to 24042370	No	ns	0.0613	B-C		
9									
10	Test details	Mean 1	Mean 2	Mean Diff.	SE of diff.	n1	n2	q	DF
11	AB vs. IL	1983606333	1497820790	485785543	145770158	3	3	4.713	6
12	AB vs. CR5	1983606333	1921041287	62565046	145770158	3	3	0.6070	6
13	IL vs. CR5	1497820790	1921041287	-423220497	145770158	3	3	4.106	6