



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

REDUCCIÓN DEL PROTEOMA DE *ESCHERICHIA COLI* MEDIANTE  
SILENCIAMIENTO DE LA EXPRESIÓN DE GENES USANDO CRISPR Y SU  
APLICACIÓN EN BIOLOGÍA SINTÉTICA.

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:  
LIC. MIGUEL ANGEL BELLO GONZÁLEZ

TUTOR PRINCIPAL  
Dr. José Utrilla Carreri  
(Centro de Ciencias Genómicas, UNAM)

MIEMBROS DEL COMITÉ TUTOR  
Dr. Miguel Cevallos  
(Centro de Ciencias Genómicas, UNAM)  
Dr. Adelfo Escalante  
(Instituto de Biotecnología)

Cuernavaca, Morelos, junio, 2019



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## Índice

I. Introducción:	4
II. Antecedentes:	5
<b>2.1 Carga metabólica de construcciones sintéticas:</b>	5
<b>2.2 Factores de Transcripción</b>	9
<b>2.3 CRISPR/Cas9 (Clustered Regularly Interspaced Palindromic Repeats):</b>	14
<b>2.4 CRISPRi:</b>	16
<b>2.5 Sistema de medición de recursos celulares:</b>	18
<b>2.6 Antecedentes experimentales:</b>	21
III. Metodología	23
<b>3.1 pCRISPathBrick:</b>	23
<b>3.2 Diseño de los gARN:</b>	24
<b>3.3 Detección de la liberación de recursos celulares:</b>	25
<b>3.4 Selección de los factores de transcripción:</b>	26
<b>3.5 Experimentos de liberación de recursos celulares:</b>	27
IV Resultados y Discusión	27
<b>4.1 Generación del sistema CRISPRi</b>	27
<b>4.2 Prueba de represión <i>in vivo</i>:</b>	29
<b>4.3 Efecto de la expresión de dCas9 en la disponibilidad de recursos celulares</b>	34
<b>4.4 Experimento cinético de liberación de recursos celulares:</b>	36
<b>4.5 Efecto de los gARN's en la fluorescencia máxima, densidad óptica máxima y velocidad de crecimiento máxima.</b>	47
V. Conclusiones	53
VI. Perspectivas	54
VII. Bibliografía	55

## Índice de Figuras

FOTOGRAFÍA 4. 1.....	28
FOTOGRAFÍA 4. 2.....	29
FOTOGRAFÍA 4. 3.....	30
FOTOGRAFÍA 4. 4.....	32
FIGURA 4.1.....	31
FIGURA 4. 2.....	33
FIGURA 4. 3.....	35
FIGURA 4. 4.....	37
FIGURA 4. 5.....	39
FIGURA 4. 6.....	41
FIGURA 4. 7.....	43
FIGURA 4. 8.....	44
FIGURA 4. 9.....	46
FIGURA 4. 10.....	47
FIGURA 4. 11.....	49
FIGURA 4. 12.....	51

## **I. Introducción:**

La biología sintética combina principios de la biología molecular, ingeniería y computación para el diseño racional y construcción de circuitos genéticos con comportamientos predecibles en el contexto celular. Estas construcciones genéticas permiten crear sistemas biológicos con características superiores a las encontradas en el medio ambiente u otras completamente nuevas (Borkowski *et al.*, 2016). A pesar de los grandes avances y aplicaciones en este campo de la ciencia en los últimos años, se enfrentan desafíos considerables en su desarrollo. Esto es debido, fundamentalmente, a la extraordinaria complejidad de los sistemas biológicos, a las interacciones muchas veces impredecibles entre los sistemas heterólogos en el contexto celular y a la influencia del medio ambiente. Estos factores atentan contra el desempeño de los sistemas sintéticos para cubrir las altas expectativas de la industria biotecnológica (Cardinale and Arkin., 2012).

En aras de sobrepasar tales dificultades, se están haciendo grandes esfuerzos, como son la utilización de sistemas de expresión ortogonales, es decir que no interfieran con la maquinaria natural del hospedero (Glasscock *et al.*, 2016), o bien la construcción de sistemas biológicos con genomas mínimos (Hutchison *et al.*, 2016) o el uso de sistemas sintéticos libres de células (Yuan., 2017). En la naturaleza los organismos necesitan asignar recursos celulares de manera dinámica a varias funciones imprescindibles, para garantizar su crecimiento y supervivencia frente a las condiciones cambiantes de su medio ambiente. Para lo cual, los microorganismos deben mantener la expresión de gran cantidad de proteínas con diversas funciones para contrarrestar estos efectos negativos del ambiente. Estas funciones se denominan funciones de cobertura, ya que preparan al microorganismo a las nuevas condiciones adversas. Están descritos mecanismos que dentro de la población microbiana favorecen la aparición de fenotipos que permitan de manera general lidiar con fluctuaciones del ambiente que puedan ser perjudiciales para la población en sentido general. Esto permite que las células que presentan estos fenotipos puedan garantizar la supervivencia de toda la población o al menos de una parte, en caso de la aparición de un evento negativo repentino (Acar *et al.*, 2008).

Estas funciones si bien cumplen un papel sumamente importante en las condiciones naturales de los microorganismos, no son imprescindibles en las

condiciones de la industria, donde se recrean las condiciones óptimas (pH, temperatura, concentración de sales, ausencia de inhibidores, concentración de sustrato, entre otras), para garantizar la mayor productividad del proceso biotecnológico. En el caso particular de la bacteria modelo *Escherichia coli* se han realizado varios estudios donde se ha podido determinar que fracciones de los recursos celulares (proteoma) son asignados o asociados a funciones biológicas globales (crecimiento, funciones de cobertura, sector ribosomal) dentro de la célula.

Estos estudios se han basado en el modelado a escala genómica (GSM) haciendo posible calcular los recursos necesarios para mantener la tasa de crecimiento y permitir la expresión de una construcción sintética en una condición determinada (O'Brien *et al.*, 2016). Numerosos estudios plantean que el proteoma es finito, así como lo es la capacidad de la célula para expresar una cierta cantidad de proteínas. Cuando se sobreexplota esta capacidad de la célula para expresar la cantidad máxima de proteínas que biológicamente es posible, ocurre el fenómeno de "carga proteómica" o "burden", lo que genera que los niveles de expresión del sistema sintético sean muy inferiores a lo esperado (Ceroni *et al.*, 2015).

Se ha observado que la reducción de la expresión del 2 - 5% de los genes con funciones de cobertura puede redirigir alrededor de un tercio de la energía de estas funciones hacia crecimiento; llevando a un aumento del 22% en la velocidad de crecimiento en un entorno constante (Utrilla *et al.*, 2016). Estos estudios soportan la idea que reduciendo la expresión de este segmento del proteoma dedicado a las funciones de cobertura, se podría incrementar la reproducibilidad y el rendimiento de las construcciones sintéticas en la célula.

## **II. Antecedentes:**

### **2.1 Carga metabólica de construcciones sintéticas:**

Una de las metas más deseadas de la biología sintética es el empleo de circuitos sintéticos ortogonales dentro de un organismo chasis, ya que estos circuitos ortogonales no deberían tener interacción con la expresión genética y elementos regulatorios presentes de manera natural en el organismo chasis. En aras de

lograr la predictibilidad de la función de estos circuitos, considerando que la expresión de estos circuitos requiere compartir los recursos celulares limitados de la célula, dígase, nucleótidos, ATP, maquinaria replicativa, transcripcional y traduccional, lo que genera un efecto de carga metabólica. En teoría, un sistema ortogonal debería estar lo más alejado posible evolutivamente del organismo chasis, donde se pretende implantar, para así evitar al mínimo posible cualquier interferencia con la maquinaria nativa de la célula; ya que esto permitiría conseguir que dicho circuito se comporte de manera estable en diferentes circunstancias y estados fisiológicos de la célula hospedero.

Un aspecto importante del empleo de estos circuitos sintéticos es el número de copias presentes el vector, así como el número de copias que puede presentar este dentro de la célula. Estos aspectos fueron determinados por Liu *et al.*, 2018, donde comprobaron que la influencia del número de copias del vector y el número de copias del circuito sintético, en este caso una puerta lógica, la cual requiere de la presencia de dos señales al unísono para que el circuito pueda activarse (AND), se comprobó que sí afectan de manera considerable el transcriptoma y el metabolismo en general del organismo chasis. En este estudio, los autores demuestran que el número de copias del vector es capaz de influir en el transcriptoma de la célula. El número de copias los plásmidos afectan la velocidad de crecimiento celular, siendo los plásmidos de medio número de copias los de mayor influencia con respecto a los plásmidos con bajo número de copias. Esto se debe fundamentalmente a que en un medio rico y apto para el crecimiento acelerado de *E. coli*, las cepas con un plásmido de mayor número de copias no compiten favorablemente por los recursos celulares limitados con una cepa con menor número de copias.

Estos autores también demostraron que uno de los factores más importantes en incrementar la carga metabólica es el número de copias del circuito, debido al costo metabólico, lo cual disminuye la expresión de los genes asociados al circuito ortogonal. En este estudio también se comprobó que los circuitos tienen poco efecto negativo en los niveles de expresión de la maquinaria biosintética y de replicación celular. En contraste con este hecho, los autores anteriormente referenciados, comprobaron que el efecto negativo predominante sobre la velocidad de crecimiento, es la expresión genética de los circuitos sintéticos y la carga metabólica asociada a dicha expresión, siendo esta la causa de mayor

impacto en el incremento de la carga proteómica celular. También comprobaron que los niveles de expresión nativos más afectados, estaban relacionados con rutas metabólicas asociadas con la biosíntesis de aminoácidos, varios transportadores del tipo ABC y en especial transportadores especializados en internalizar aminoácidos.

Estos resultados reafirman fuertemente la hipótesis que la carga metabólica impuesta por los elementos ortogonales está directamente asociada con la síntesis de proteínas en el hospedero en el sentido general; y que esta pudiera ser responsable en gran medida de la aparición de fenotipos de crecimiento lento asociado a la expresión de dichos elementos ortogonales. Estos resultados son de gran importancia ya que brindan una idea clara de cómo el hospedero “sufre” al ser sometido a la expresión de un vector y, más importante, aun nos guía sobre qué estrategias se pudieran seguir para evitar estos efectos negativos en las construcciones sintéticas de interés (Liu *et al.*, 2018).

Por otra parte, varios autores plantean la necesidad de comprender de una mejor manera qué genes son activados o reprimidos ante la presencia del fenómeno de “burden”. Para realizar este análisis, se han realizado experimentos de ARNseq para semicuantificar el mARN total de *E coli* en un momento dado: esta técnica también puede ser utilizada de manera diferencial, ya que se puede cuantificar la variación de los niveles de expresión génica de un mismo gen en diferentes condiciones o a diferentes tiempos. Esta técnica fue aplicada de conjunto con un sistema previamente diseñado (Ceroni *et al.*, 2017), llamado monitor de capacidad, que permite conocer el *burden* a la que es sometida la célula ante una construcción genética determinada. Esto permite conocer, entre otras cosas, qué diseños sintéticos presentan menor carga metabólica a la célula y por tanto pueden tener un comportamiento más predecible y más robusto en el tiempo.

Estos mismos autores (Ceroni *et al.*, 2017), desarrollaron un circuito que permite la expresión de un ARNg (ARN guía) fusionado transcripcionalmente a un promotor de choque térmico, el cual tiene la capacidad de activarse ante la presencia de sobrecarga metabólica: con este promotor fueron capaces de construir un sistema de control por retroalimentación negativa, es capaz de controlar la expresión de un circuito sintético determinado, reduciendo el *burden*



al que es sometida la célula en presencia de una construcción sintética, a través de la represión del mismo utilizando el sistema CRISPR/dCas9.

Como parte de este estudio (Ceroni *et al.*, 2017), los autores emplearon dos cepas de *E. coli* y cuatro construcciones inducibles: expresión de reportero inducible, la sobreexpresión de una proteína heteróloga no funcional de gran tamaño, la expresión de un operón de una ruta metabólica. Comprobaron que estas construcciones impusieran una carga metabólica al hospedero antes de realizar los estudios de transcripción, a través del empleo del monitor de capacidad anteriormente mencionado. Estos autores determinaron que el promotor que presentó un mejor comportamiento de apagado/encendido en presencia de los inductores de las cuatro construcciones evaluadas fue *htpG1*. De esta manera, decidieron emplear dicho promotor para diseñar un biosensor que permite detectar los efectos de la carga metabólica y la expresión de una molécula efectora que permita reprimir la expresión de los genes que están causando dicha sobrecarga en el hospedero. Los autores de este trabajo implementaron un sistema que comprendía la expresión del ARN guía, bajo el control de dicho promotor. Además utilizaron la expresión de dCas9 de manera constitutiva, lo que permite que los efectos de la represión fuesen más inmediatos; sólo dependiendo de que el promotor *htpG1* permita la transcripción del gARN.

Posteriormente, co-expresaron este sistema de control del *burden* y una de las construcciones que presento mayor carga metabólica para la célula (pSB1C3-H2) en una cepa que presentaba el monitor de capacidad, la construcción empleada en este estudio cuenta con un sitio de unión a ribosomas muy fuerte y una proteína heteróloga relativamente grande (VioB), fusionada con la proteína mCherry, bajo el control de un promotor inducible por arabinosa. El gARN fue diseñado para unirse al promotor de VioB-mCherry. Los niveles de expresión de VioB-mcherry fueron bajos a pesar de estar en presencia de los mayores niveles del inductor y también la carga metabólica era lo bastante baja como para no afectar la demasiado la velocidad de crecimiento. Como conclusiones de este estudio, los autores plantean que las construcciones sintéticas solo podrían utilizar en el mejor de los casos un cuarto de los recursos transcripcionales de la célula. Y concluyen que el impacto sobre los recursos traduccionales celulares

sería mucho mayor, sobre todo teniendo en cuenta que la mayoría de las construcciones sintéticas poseen o son expresamente diseñadas con sitios de unión del ribosoma muy fuertes, con lo cual estos secuestrarían gran cantidad de ribosomas que son necesarios para mantener los niveles de expresión de las proteínas nativas del hospedero y que son necesarias para el adecuado crecimiento del organismo.

Es válido aclarar que estos experimentos de modulación *in vivo* de la expresión de genes, han tenido un gran avance gracias al desarrollo de la técnica de CRISPR. Un aspecto a tener en cuenta es que el uso de múltiples gARN para la represión o activación de múltiples genes diana simultáneamente, es una de las características más promisorias del sistema CRISPRi y CRISPRa respectivamente, sobre todo se está explorando su aplicación para controlar cascadas de señalización y redes de regulación debido a su extraordinaria capacidad modular. Sin embargo, el uso de múltiples gARN implica que estos deben competir en un momento dado por una cantidad menor de moléculas dCas9 generando interacciones no deseadas, se pensó que el incremento de los niveles de dCas9 podrían ayudar a mitigar este efecto pero elevados niveles de esta proteínas tienen efectos tóxicos en la célula y afectan directamente la velocidad de crecimiento de las bacterias (Chen and Del Vecchio, 2018).

## **2.2 Factores de Transcripción**

### ***marA***

El locus *mar* es quizás uno de los sistemas cromosomales mejor descrito en la literatura. Es responsable de proveer resistencia a bajos niveles de varios antibióticos, tanto en *E. coli* como en *Salmonella sp.* Este locus se transcribe en forma de operón, es inducido por *marA*, el cual a su vez esta codificado dentro de este mismo operón (Randall and Woodward, 2002). En estudios anteriores se ha comprobado que este locus es responsable de permitir que algunas bacterias presenten resistencia a antibióticos como cloranfenicol y tetraciclina fundamentalmente, aunque también se ha encontrado que brinda resistencia a antibióticos como rifampicina, puromicina, ampicilina, ácido nalidíxico y penicilina G entre algunos otros. El locus *mar* consta de dos unidades transcripcionales en sentido opuesto, en cuyo centro se encuentra el operador *marO*. Una de estas unidades transcripcionales es *marC*, la cual se cree que codifica para una

proteína integral de membrana. La otra unidad transcripcional comprende *marRAB* la cual codifica para una proteína represora *marR*, una proteína activadora llamada *marA* y una proteína con función desconocida llamada *marB*. El represor *mar* normalmente se encuentra unido a su sitio operador *marO*, impidiendo la transcripción de la unidad transcripcional *marRAB* solo cuando aparece su inductor este se despegar y permite la transcripción de *marRAB*. Se ha demostrado que existen diferentes sustancias que pueden actuar como inductores en *E. coli*, estos pueden ser cloranfenicol, tetraciclina, benzoato de sodio, acetaminofén, entre otros. Una vez que *marRAB* se puede transcribir, *marA* permite la auto sobreexpresión del propio operón, además es responsable de la sobreexpresión o disminución de la expresión de alrededor de sesenta genes, muchos de los cuales están involucrados en la disminución del influjo de antibióticos o en el incremento del eflujo de estos.

Un ejemplo claro de este efecto es la sobreexpresión del sistema de eflujo inducido por estrés *acrAB-TolC*, *acrAB* es una proteína de la membrana encargada de sacar de la célula sustancias tóxicas para la misma y posiblemente *TolC* sea otra proteína de membrana que esté involucrada en este mismo proceso, otra forma sería la inducción de la sobreexpresión de un ARN antisentido que favorece la degradación del transcrito de *ompF*, el cual está relacionado con la internalización a la célula de sustancias acuosas y quizás esté relacionado con la degradación de otros transcritos que codifiquen para otras proteínas integrales de membrana. Vale destacar que conjuntamente con *marA*, existen otros dos sistemas que regulan el mismo conjunto de genes confiriendo resistencia a bajos niveles de antibióticos y solventes orgánicos, estos son *soxS* y *robA*.

Se ha investigado la posibilidad de que la inducción del locus *mar* sea responsable en gran medida de la resistencia a antibióticos que presenta de manera natural los biofilms. En este estudio, se evaluaron diferentes métodos de cultivo (lote o continuo) así como la formación de biofilms en filtros Sorbaroud, además se evaluó la influencia de dos medios de cultivo, uno rico (LB) y otro definido, para comprobar cuáles condiciones favorecerían o no la expresión del locus *mar*, también se fusionó *marO* con el gen de la beta-galactosidasa (*lacZ*) para poder monitorear los niveles de expresión del locus *mar*. En este estudio se comprobó que en quimiostato existía una dependencia entre el crecimiento lento

y el incremento de la actividad de la beta-lactosidasa, sin embargo en condiciones lote no presentaban el mismo comportamiento. Se logró determinar que la sobreexpresión de *mar* depende del estadio fisiológico de la célula, y que existía una mayor expresión en medio mínimo que en medio rico, lo que indica que si bien la sobreexpresión del locus *mar* no es el responsable absoluto de la resistencia a antibióticos que presentan los biofilms, influye considerablemente. Se plantea que en células que se encuentran formando biofilm y por tanto de son de lento crecimiento, la sobreexpresión del locus *mar* pudiera ser un factor a considerar en aras de mitigar el efecto nocivo de los biofilm en la industria y en la medicina. Sobre todo teniendo en cuenta la posibilidad de brindar resistencia a múltiples antibióticos no relacionados estructuralmente, así como resistencia a solventes orgánicos, sobre todo en ecosistemas donde exista limitaciones de nutrientes (Maira-Litran *et al.*, 2000)

### ***csgD***

También llamado regulador maestro de la producción de biofilm, es un factor de transcripción que está involucrado directamente en la regulación de genes encargados de la síntesis del biofilm y en la producción de las fimbrias. Favoreciendo la adhesión, anclaje y protección de la célula contra la presencia de agentes tóxicos, desecación, carencia de nutrientes y presencia de antibióticos, en organismos como *E. coli* y *Salmonella* sp. El factor transcripcional *csgD* regula directamente el operón *csgABC* y el operón *csgDEFG* los cuales son responsable de la síntesis de las fimbrias, estructuras celulares implicadas en la adhesión celular, específicamente en interacciones célula-célula. En el caso del operón *csgABC*, *csgD* se une directamente al promotor de dicho operón y facilita el reclutamiento de la ARN polimerasa permitiendo la transcripción. Este operón codifica para los componentes principales de las fimbrias y el segundo operón (*csgDEFG*) codifica para un conjunto de proteínas que actúan como factores de ensamblaje y polimerización de las fimbrias. El operón *csgABC* es controlado también por el factor sigma *rpoS* en la fase estacionaria del crecimiento de *E. coli*. (Perni *et al.*, 2016). La producción de biofilm está sometida a una regulación fuerte debido a que si bien es capaz de brindarle a las células que los producen una ventaja adaptativa para sobrevivir en condiciones de estrés, también es cierto que es un proceso bastante caro desde un punto de

vista metabólico y energético para la célula. Siendo *csgD* el encargado de reconocer varias señales de estrés e inducir la expresión de los genes anteriormente mencionados, también induce la expresión de varios genes como *adrA*, el cual codifica para la enzima que sintetiza el di-GMP cíclico, el cual es un mensajero secundario bacteriano que favorece la síntesis de celulosa, componente fundamental de la matriz de exopolisacáridos que constituyen el biofilm. El di-GMP cíclico actúa a través de la unión directa de este compuesto con el dominio PliZ de la enzima celulosa sintasa, induciendo un cambio conformacional en la enzima lo que provoca que un incremento en la actividad enzimática de la misma (Ogasawara *et al.*, 2011).

Se ha encontrado que si bien la sobreexpresión del factor transcripcional *csgD* es el principal responsable de la producción de la matriz de exopolisacárido que forma el biofilm, se ha podido observar que dicha sobreexpresión no es homogénea en la comunidad bacteriana. Teniendo un carácter bi-estable, donde una subpoblación posee mayores niveles de expresión de este factor de transcripción mientras que otra subpoblación presenta menores niveles. De esta manera la subpoblación con mayores niveles de expresión es la responsable de la producción de la matriz de exopolisacárido y la formación de la estructura del biofilm, el cual encapsula a la subpoblación con menores niveles de expresión de *csgD*. De esta manera al ser un proceso tan costoso metabólicamente, la comunidad bacteriana en general es capaz de mantener la ventaja en condiciones adversas, manteniendo también el potencial de la comunidad para crecer (Grantcharova *et al.*, 2010). Un mecanismo interesante de regulación del biofilm que involucra a *csgD* y *rpoS* fue encontrado en *Salmonella* en relación con cambios en temperatura. Este factor ambiental afecta la producción de biofilm en este organismo, la región intergénica presente en *Salmonella* entre los operones *csgABC* y *csgDEFG* es de aproximadamente 582 pb. En dicha región se ha demostrado que se pueden unir varios factores de transcripción que pueden regular la expresión de dichos operones. Durante la fase estacionaria de este organismo, el factor sigma *rpoS* recluta la ARN polimerasa formando la holoenzima, lo que genera la unión a promotores que solo son inducidos en la fase estacionaria del crecimiento y se inicia la transcripción del ARN mensajero. Durante la fase estacionaria también se empieza a acumular la proteína accesoria Crl, la cual es más estable a temperaturas que rondan los 28 grados

Celsius, siendo mucho más inestable a 37 grados Celsius, lo curioso de este mecanismo es que *rpoS* favorece una disminución de la actividad polimerizadora de la ARN polimerasa, hasta la unión de la proteína Crl, teniendo en cuenta que al menos en *Salmonella* la producción de biofilm ocurre a 28 grados Celsius se piensa que esta proteína sirve como sensor de temperatura favoreciendo la formación del biofilm ante este estrés térmico. (Liu *et al.*, 2014).

### ***rpoS***

*rpoS* es un factor sigma que actúa sobre un gran conjunto de genes durante la fase estacionaria del crecimiento de varias bacterias incluyendo *E. coli* y *Salmonella sp.* La ausencia de nutrientes constituye uno de los principales detonadores para aumentar la síntesis de *rpoS*, factores como la ausencia de nitrógeno, carbono o fosfato son señales claras del inicio de la fase estacionaria. La célula percibe este estrés y dispara la síntesis de *rpoS* que a su vez va a modular los niveles de expresión de múltiples genes. En estudios anteriores se ha determinado que los niveles de *rpoS* se empiezan a incrementar no solo al inicio de la fase estacionaria ante señales de disminución de nutrientes, sino también a partir que el crecimiento microbiano empieza a disminuir, mucho antes que este se detenga (Griffith *et al.*, 2004).

La regulación de *rpoS* ocurre a tres niveles (transcripcional, traduccional y postraduccional). La secuencia codificante de *rpoS* se encuentra adyacente y en el mismo sentido de la transcripción que el gen *nlpD*, se plantea que la mayor parte de los niveles de *rpoS* detectados en la fase exponencial del crecimiento se deben a precisamente a dos promotores débiles que se encuentran situados corriente arriba de *nlpD* que inician la transcripción. Al inicio de la fase estacionaria se activan otros dos promotores que son responsables de la mayoría de los niveles de expresión durante la fase estacionaria.

También se ha comprobado la presencia de dos dominios potenciales implicados en la unión de CRP corriente arriba de estos promotores, sugiriendo que elevados niveles de cAMP reducirán la expresión de este factor de transcripción. Posiblemente el nivel de mayor regulación para *rpoS* es el nivel postranscripcional, donde un ARN anti-sentido ubicado corriente arriba, este una vez transcrito se aparea con la región correspondiente al sitio de unión del ribosoma del transcrito de *rpoS*, inhibiendo de esta manera la traducción del mismo. Estudios avalan la presencia de un ARN anti-anti-sentido llamado *dsrA*,

el cual es capaz de aparearse con dicho ARN anti-sentido y de esta manera favorecer la expresión de *rpoS* (Hitsch and Elliot, 2002).

Se plantea que estos promotores poseen una secuencia -10 muy similar a las que poseen la mayoría de los promotores dependientes de sigma 70. Sin embargo se ha comprobado que en los promotores que están regulados por *rpoS* no presentan la típica secuencia -35. También se ha sugerido que de conjunto con la secuencia -10 esté implicado en el reconocimiento de la región promotora una topología específica que permite que *rpoS* se una y promueva la transcripción de determinados genes en la transición a la fase estacionaria. Debido a la superposición de muchos de los genes que son regulados por ambos factores de transcripción, se sugiere que solo aquellos promotores con las secuencias -10 y -35 serán reconocidos por sigma 70, y que aquellos que posean la secuencia -10 apropiada en la vecindad de una curvatura intrínseca serán reconocidos por *rpoS*.

### **2.3 CRISPR/Cas9 (Clustered Regularly Interspaced Palindromic Repeats):**

La capacidad para perturbar de manera controlada la expresión genética, es una tecnología necesaria e imprescindible en los campos de la biología sintética. En aras de comprender como funcionan estos sistemas complejos, así como mejorar la capacidad de diseño e implementación de los circuitos genéticos diseñados por los investigadores. Sin embargo existen pocas herramientas que sean capaces de llevar a cabo la regulación de la transcripción de manera específica, sencilla y dinámica. Una de estas herramientas es el sistema CRISPR, está presente en más del 40% de las bacterias y en hasta el 90% de las archaeas, brindándoles a estos organismos resistencia ante la presencia de ADN foráneo en su interior. Este sistema presenta dos tipos de secuencias, secuencias espaciadoras que están flanqueadas por secuencias palindrómicas repetidas idénticas, estas secuencias espaciadoras se derivan de ADN plásmidico o viral y son el vestigio de infecciones anteriores a las que fue sometido el microorganismo. Este arreglo recibió el nombre de CRISPR (*Clustered Regularly Interspaced Palindromic Repeats*). Asociados a estas secuencias características se encuentran una serie de genes denominados genes *cas* (CRISPR associated), los cuales codifican para proteínas con

actividad nucleasa y helicasa, los cuales se encargan del procesamiento del transcrito correspondiente al locus donde se localizan las secuencias espaciadoras y repetidas. Además permiten que fragmentos de ADN foráneo sea incorporado a este locus por lo que se ha propuesto que este sistema CRISPR/Cas es un mecanismo de inmunidad adaptativa de muchos microorganismos. Este mecanismo de manera natural se lleva a cabo a través de tres fases: primero ocurre la inserción de un pequeño fragmento de ADN foráneo como una secuencia espaciadora dentro del arreglo CRISPR, segundo ocurre la transcripción de un crARN precursor, el cual es sometido a un proceso de maduración generándose crARN(CRISPR ARN) individuales, los cuales están compuestos por una secuencia repetida y una secuencia espaciadora proveniente de ADN foráneo, tercero el crARN sirve de guía para el corte de ADN foráneo por la proteína Cas9 a través de la complementariedad de bases existente entre el crARN y el ADN foráneo (Doudna and Charpentier., 2014). Hasta el momento se han descrito tres sistemas CRISPR (I, II, III), que emplean distintos mecanismos para ejecutar este proceso de reconocimiento y corte del ADN diana, siendo el más aplicado en la actualidad por su relativa sencillez y eficacia es el sistema II encontrado en *Streptococcus pyogenes*, ya que necesita de una sola proteína(Cas 9) para efectuar este proceso. Por supuesto con la ayuda del crARN, a diferencia de los tipos I y III, que requieren de grandes complejos proteicos Cas para efectuar el corte del ADN. Un requerimiento sumamente importante en el sistema II es la presencia de una secuencia PAM (*protospacer adjacent motif*) la cual debe estar adyacente a la secuencia diana del crARN, para que se efectúe el silenciamiento del ADN. La maduración del crARN requiere del tracrARN (un ARN complementario que actúa en *trans*) y la ARNasa III, sin embargo en la práctica en el laboratorio se emplea un ARN guía pequeño (sgARN) de 20 nucleótidos sintético diseñado para poseer una estructura en forma de gancho que mimetiza el complejo tracrARN-crARN. La complementariedad de bases entre el sgARN y su secuencia diana en el ADN provoca un corte de doble cadena debido a la acción de Cas9. (Qi *et al.*, 2013); (Vigouroux *et al.*, 2017).

Este sistema ha sido ampliamente aplicado no solo en bacterias sino también en varios modelos eucariontes. Uno de estos ejemplos es la creación de una librería a escala genómica de sgARN para determinar la esencialidad de genes en la



viabilidad de células cancerosas y células madre pluripotentes, a través de un screening genético de 18,080 genes que fueron editados con este sistema para la selección tanto positiva como negativa de mutantes. Muchos de estos estudios se habían realizado previamente utilizando iARN para la disminución de la expresión de dichos genes sin embargo utilizando esta herramienta se presentaban muchas interferencias cruzadas y además no posibilitaba el *knockout* de genes (Shalem *et al.*, 2013)

En plantas también han empezado a aparecer numerosas aplicaciones, las cuales se enfocan en la utilización del sistema CRISPR/Cas9 para la producción de NHEJ (*Non Homologous End Joining*), para la creación de mutaciones heredables a la progenie de numerosas especies de plantas. A través de la formación de cortes en doble cadena en marcos abiertos de lectura de los genes de interés, destacando el *knockout* de seis alelos de los tres genes *mlo* presentes en el trigo, lo que le confirió resistencia ante la enfermedad del oídio, esta resistencia fue heredada a la descendencia. También resulta muy interesante que en algunos países de la Unión Europea, estas modificaciones en los cultivos no se consideran como organismos genéticamente modificados (GMO), lo que impulsa todavía más la investigación en este sentido al eliminarse este impedimento legal (Puchta., 2016).

#### **2.4 CRISPRi:**

El sistema CRISPRi es una variante del sistema original CRISPR, donde se emplea una variante mutagenizada de la proteína original (Cas9), para que esta pierda su capacidad endonucleasa (dCas9). Esta variante de la nucleasa catalíticamente muerta (dCas9) se acopla con el gARN y se unen a la secuencia diana, este complejo impide estéricamente que la ARN polimerasa transcriba un gen determinado. Como la represión del sistema CRISPRi se deriva de la complementariedad de bases entre el gARN y la secuencia del gen diana, se puede virtualmente reprimir cualquier gen rediseñando el gARN, esta característica junto con la capacidad de reprimir reversiblemente (ya que no hay modificación o corte de la secuencia de ADN) y de manera simultánea múltiples genes hacen de este sistema la herramienta ideal para este trabajo. (Hawkins *et al.*, 2015). Este sistema ya ha sido muy empleado en varias aplicaciones para la

represión de genes específicos en aras de incrementar la productividad de fármacos y otras sustancias de gran interés económico (Kim *et al.*, 2017). Este sistema brinda la posibilidad de ser utilizado como herramienta para la represión de múltiples genes dianas sin alterar directamente la secuencia de ADN. Recientemente ha sido utilizado como herramienta para la represión múltiple de genes que participan en varias rutas metabólicas, que compiten entre sí, por precursores comunes. Enriqueciendo la disponibilidad de acetil-CoA a través de la represión de varios genes endógenos responsables de la formación de subproductos como el acetato, etanol, lactato y succinato, para redireccionar el flujo metabólico a la producción de n-butanol en *E. coli*. Para ello diseñaron un plásmido pSECRi-PFLA derivado de las series pSEVA el cual posee un promotor inducible por ramnosa, la secuencia dCas9, sgARN y un marcador de selección a cloranfenicol. Estos autores seleccionaron los genes *pta*, *frdA*, *ldhA*, and *adhE*, los cuales codifican para las enzimas fosfato acetil transferasa, fumarato reductasa, lactato deshidrogenasa y alcohol/aldehído deshidrogenasa respectivamente. Como resultado obtuvieron que la represión múltiple de estos genes permitió un incremento en la producción de n-butanol en 2.1 veces con respecto al control. Para comprobar que efectivamente el incremento obtenido se debía a la represión producida por su sistema emplearon concentraciones variables del inductor lo cual arrojó que efectivamente en ausencia de este los niveles de n-butanol disminuían. También comprobaron que la presencia de la enzima CAT (cloranfenicol acetiltransferasa) codificada por el gen *cat*, la cual es la responsable de conferir la resistencia a cloranfenicol también era la responsable de la formación de acetato de butilo y butil ester, con lo cual reemplazaron dicho gen con el gen *b/a*, el cual codifica por la enzima beta-lactamasa (nuevo marcador de resistencia) y se logró un incremento 1.5 veces más con respecto a las cepas que poseían resistencia a cloranfenicol (Kim *et al.*, 2017).

Otra aplicación es la implementada por los autores Wu *et al* (2015) con el objetivo de incrementar la producción intracelular de malonil-CoA. El cuales un precursor muy importante en la síntesis de ácidos grasos, flavonoides y policétidos, utilizando la producción heteróloga de naringina como sistema modelo. Logrando reprimir la expresión de genes endógenos que forman parte de rutas del metabolismo central. Para ello emplearon dos plásmidos, uno codificante para la

proteína dCas9 bajo la influencia de un promotor T7 y otro para los sgARN específicos para unirse a las secuencias dianas, con un promotor Trc. Se tuvieron en cuenta el efecto inhibitorio del crecimiento celular que pudiesen tener tanto las perturbaciones de genes involucrados en el metabolismo central como las concentraciones de malonil-CoA. Estos autores determinaron la mejor combinación de genes (*fabF/fumC/fabB/sucC/adhE*) a reprimir que al mismo tiempo no afectaran considerablemente el crecimiento celular (10% de reducción) logrando aumentar la producción de naringina a una concentración de 421.6 mg/L.

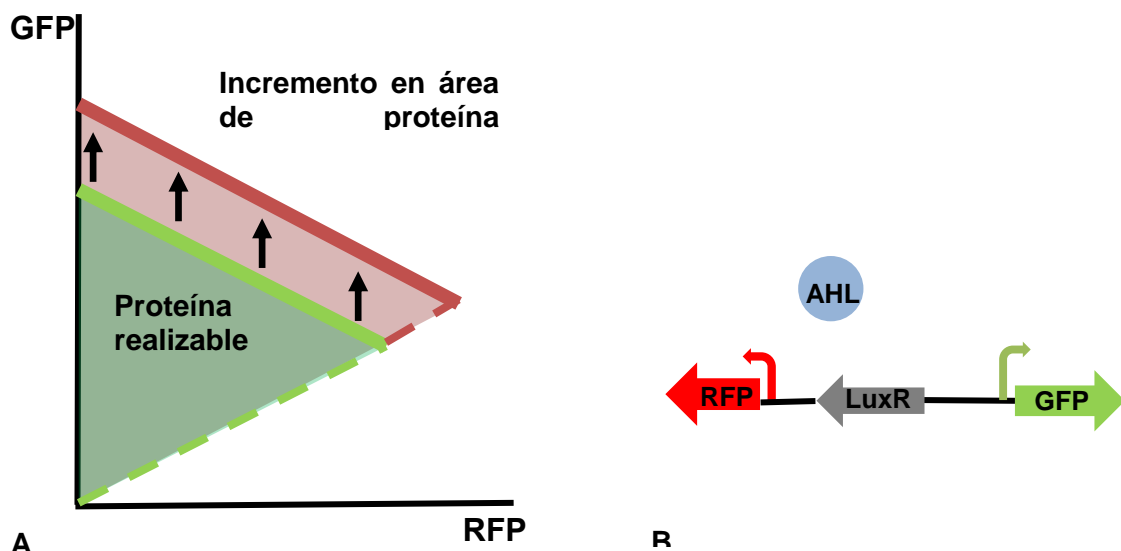
También se han realizado estudios para determinar la cinética de represión que puede tener un solo complejo sgARN-dCas9. En este estudio se fusionó la proteína fluorescente YPet a la dCas9 para permitir el rastreo celular del complejo una vez reconociera la secuencia de *lacO1* en la región operadora (secuencia diana). Estos autores determinaron que como promedio una sola molécula de gARN-dCas9 se demora en encontrar su objetivo unas 6 horas como máximo, determinándose que dCas9 tiene que explorar todas las secuencias PAM posible hasta hallar la correcta tomándose aproximadamente 30 ms para el desenrollado e inspección de cada una de estas secuencias. También comprobaron cuanto tiempo permanecía unido este complejo al ADN a través de la suplementación a con IPTG al inicio del cultivo y posteriormente eliminado a t=0 y la disociación fue medida a través de la disminución de puntos fluorescentes en la célula. Demostrando que la unión de este complejo con una determinada secuencia es muy fuerte y solo es reversible cuando ocurre la replicación del ADN. Se determinó que, si bien la cinética de unión de dCas9 es un poco lenta en comparación con otros mecanismos y debido fundamentalmente a la flexibilidad del mecanismo de unión, esto puede ser superado a través del empleo de un mayor número de copias (Jones *et al.*, 2017).

## **2.5 Sistema de medición de recursos celulares:**

Cuando nuevos genes son introducidos dentro de la célula los recursos celulares finitos que esta posee, deben ser compartidos y redirigidos a la expresión de los nuevos genes introducidos. Lo que establece una afectación a la robustez de la célula, modificando parámetros como la densidad óptica y la velocidad de crecimiento, pero además afecta la predictibilidad de los circuitos sintéticos

introducidos en el hospedero. Esto sugiere que la economía que rige la distribución de dichos recursos celulares puede jugar un papel preponderante en la predictibilidad y estabilidad de los circuitos genéticos. Parte de estos recursos celulares de gran importancia con los que cuenta la célula son, la disponibilidad de ribosomas y de ARN polimerasa. Siendo la polimerasa quizás el recurso celular limitante más importante de la transcripción, ya que es la enzima encargada de llevar a cabo este proceso; y al igual que esta, la disponibilidad de ribosomas es el recurso limitante quizás más importante en el proceso de traducción.

Debido a esta problemática Gyorgy et al, (2015), diseñaron un sistema de dos proteínas reporteras acopladas en un mismo plásmido, llamado pSEVA63-Dual. Este plásmido cuenta con un origen de replicación PBBR1, como marcador de selección un gen de resistencia a gentamicina. Presenta dos proteínas fluorescentes, la proteína verde fluorescente (GFP), la cual está bajo el control de un promotor constitutivo y la proteína roja fluorescente (RFP), la cual está bajo el control de un promotor inducible por AHL. Ambas proteínas son capaces de transcribirse de manera independiente como unidades transcripcionales separadas, por lo que la regulación de la expresión de una



**Figura 2. 1.** Representación esquemática del sistema de cuantificación de la fluorescencia empleado en este estudio. En la parte A, se representa el aumento esperado en la emisión de fluorescencia total, producto del aumento en los recursos celulares disponibles. En la parte B, se puede observar una representación esquemática del circuito existente en el plásmido pSEVA63-Dual.

proteína no afecta la regulación de la expresión de la otra. En este trabajo los autores se interesaron también en la influencia del número de copias de los plásmidos y la fortaleza de del sitio de unión al ribosoma (RBS) podían afectar o no la distribución de recursos celulares. Los autores fueron capaces de establecer una relación lineal e inversamente proporcional entre los niveles de expresión de ambas proteínas fluorescentes no relacionadas regulatoriamente, es decir comprobaron que a mayores niveles del inductor AHL menores niveles de fluorescencia eran obtenidos con la GFP. Lo que permitió que establecieran un paralelismo con un concepto de la microeconomía denominado líneas de isocosto, el cual básicamente indica como con un presupuesto limitado se puede comprar dos productos, en este caso los dos productos a “comprar” serían las proteínas fluorescentes y el “presupuesto” con el que contamos, los recursos celulares.

Estos autores fueron capaces de comprobar que a pesar de ser unidades transcripcionales independientes, cuando están localizadas en el mismo vector ocurre un fenómeno de acoplamiento en los niveles de expresión de ambas proteínas. Debido a la limitación de ARN polimerasa y ribosomas, dependiendo entre otros factores del número de copias del vector, oscilando entre un 60%-29% cuando el vector posee dichos circuitos en un vector con un número de copias medio y un vector con muy bajo número de copias respectivamente.

Otra observación muy interesante de este trabajo es en cuanto a la fortaleza de los RBS, los autores comprobaron que los niveles de fluorescencia de GFP no resultaban afectados al máximo nivel de inducción con AHL, cuando la proteína RFP poseía un RBS débil, indicando que la causa principal de afectación de los niveles de expresión de la proteína GFP pudieran deberse a la baja disponibilidad de ribosomas. En cuanto a la demanda de ARN polimerasa comprobaron que a mayor fortaleza relativa de los RBS se requieren menores niveles de inducción de RFP para alcanzar elevados niveles de concentración de proteína debido a que son necesarios menores niveles de ARN polimerasa, lo que permite que otros genes mantengan sus niveles de expresión relativamente óptimos ya que existe baja demanda de la ARN polimerasa.

## 2.6 Antecedentes experimentales:

En trabajos anteriores en el grupo del Dr. Utrilla del Laboratorio de Biología de Sistemas y Biología Sintética, se han determinado hasta 42 factores de transcripción (TF's) que regulan positivamente al menos un gen no esencial en *E.coli*. Dichos genes codifican para proteínas con funciones de cobertura. Esta selección fue realizada a través del análisis de datos obtenidos por medios experimentales basados en la disrupción de genes y datos bioinformáticos generados a partir de modelos *in silico* del metabolismo y de expresión de genes a escala global.

También se tuvieron en cuenta datos de proteómica cuantitativa disponibles en la literatura especializada (Schmidt *et al.*, 2015). El análisis de la red de regulación transcripcional fue realizado utilizando la última versión de la base de datos RegulonDB, donde se determinó la esencialidad y no esencialidad de cada factor de transcripción, de acuerdo a estas mismas características de los genes que están sujetos a su regulación, escogiéndose para este análisis solo aquellos TF's no esenciales. También se tuvo en cuenta la subred de interacciones entre los distintos TF's. Determinándose el grado de la salida de cada TF ( $TF_{out}$ ) que a su vez se subdividió en positiva ( $TF_{out+}$ ) y negativa ( $TF_{out-}$ ) y también el grado de entrada de los genes ( $GEN_{in}$ ) regulados por estos TF's, tanto positivamente ( $GEN_{in+}$ ) y negativamente ( $GEN_{out-}$ ).

Con estos y otros datos se determinó cuales genes eran candidatos reales a silenciar, estos TF's debían regular positivamente al menos un gen no esencial y que dicho gen no poseyera interacciones negativas, con lo cual al silenciar su TF se garantiza la no expresión del mismo. También se emplearon un conjunto de datos proteómicos cuantitativos para calcular la carga proteómica (PL) en fg/célula de cada gen utilizando la expresión:

$$P_L = MW_P * C_{cell} * N_A$$

Donde se tuvo en cuenta la masa molecular de la proteína, el número de copias en la célula y el número de Avogadro. Se determinó el balance proteómico de cada TF ( $P_B$ ) el cual se calcula como la sumatoria de la carga proteómica de los genes reprimidos menos la sumatoria de los genes inducidos.

$$P_B = \sum \text{Silenced}P_L - \sum \text{Induced}P_L$$

Estos planteamientos matemáticos, y otros desarrollados en nuestro laboratorio fueron utilizados para el desarrollo de una herramienta computacional en perfeccionamiento, con la cual se han logrado calcular las mejores combinaciones de TF's a eliminar, que resulten en la mayor cantidad de carga proteómica liberada. Esta herramienta prioriza aquellos TF's que tengan bajo su influencia genes que se expresen con un nivel expresión elevado sobre aquellos TF's que regulen un número mayor de genes pero que su nivel de expresión sea bajo.

Debido a la complejidad y cantidad de genes que son potenciales dianas en este trabajo hemos decidido utilizar el sistema CRISPR/dCas9, ya que permite utilizar múltiples combinaciones de TF's blanco. Logrando reprimir la expresión de múltiples objetivos de manera simultánea y controlable en el tiempo. En este proyecto no se utilizará la eliminación por *knock-out* de genes, ya que la posibilidad de realizar múltiples combinaciones de silenciamientos es limitada, solo ofrece soluciones definitivas (edición del genoma), es decir se lograría la represión de los genes permanentemente.

Además empleando CRISPRi, se pudieran efectuar la represión de genes en cepas que ya tengan genes *knock-out* e interrogar las redes de regulación más eficientemente (Tesis de Doctorado en proceso autor: Gustavo Lastri). En este trabajo utilizaremos **CRISPRi para reprimir combinaciones de TF's**.

Nuestra diana es la red regulatoria de *E. coli*, inhibiendo la expresión de factores de transcripción específicos que controlan la expresión del segmento del proteoma encargado de las funciones de cobertura para incrementar las disponibilidades de recursos celulares para el crecimiento y expresión de construcciones sintéticas. La efectividad del acercamiento se evaluará a través de las características mejoradas de producción de proteínas fluorescentes como modelo de producción.

### **Hipótesis:**

Al modificar la red de regulación transcripcional (TRN) de *E. coli*, se logrará una liberación de recursos en el proteoma, dada por la reducción de la expresión de

proteínas no asociadas al crecimiento, lo cual aumentará la disponibilidad de recursos celulares para la adecuada integración y expresión de funciones sintéticas.

### **Objetivo General:**

Reducir la expresión de sectores del proteoma no asociado al crecimiento de *E. coli* para mejorar la integración de funciones sintéticas.

### **Objetivos Específicos:**

- a) Establecer una metodología de silenciamiento simultáneo de genes por CRISRPi.
- b) Probar el sistema en las combinaciones de 42 factores de transcripción.
- c) Evaluar el efecto del silenciamiento de sectores del proteoma en la capacidad de producir proteínas fluorescentes y su aplicación en biología sintética.

## **III. Metodología**

### **3.1 pCRISPathBrick:**

Se empleó el plásmido pCRISPathBrick, del repositorio Addgene (#65006) (Cress *et al.*, 2015), el cual tiene un tamaño de 9.3 kb. Contiene un origen de replicación p15A, como marcador de resistencia cloranfenicol, la secuencia del tracrARN, crARN leader, y una región que mantiene la arquitectura de un arreglo CRISPR. También cuenta con un non-targeting gARN diseñado para no tener ningún off-target con el genoma de *E.coli*. Este fue modificado con el único sitio de corte para la enzima de restricción BsaI, dicho sitio es donde se ensamblarán los nuevos gran, previa digestión. Estos gARN's se encontrarán bajo el control de un mismo promotor constitutivo J23119. Este plásmido también posee la secuencia de la proteína dCas9, bajo el control del promotor constitutivo. Esto permitió la represión de manera constitutiva de los factores de transcripción seleccionados, de manera individual y simultánea en el caso que se logre clonar más de un gARN en dicho vector.

Sin embargo, debido a dificultades usando el sistema de restricción-ligamiento convencional, se decidió utilizar el método de Gibson assembly para integrar en el vector los nuevos gARN. Las nuevas construcciones (secuencia espaciadora-



secuencia repetida) son incorporadas como oligonucleótidos de simple cadena de 96 nt, que poseerán extremos complementarios para facilitar el ligamiento con el esqueleto del vector, de acuerdo con los principios del Gibson assembly. En el ligamiento del extremo orientado corriente abajo de la nueva secuencia espaciadora - secuencia repetida se regenera el sitio de corte original de la enzima BsaI, el cual puede ser redigerido y regenerado cíclicamente en cada incorporación de una nueva construcción. (Cress *et al.*, 2015). Se empleó una estrategia de clonaje combinatoria que permite el ensamblaje modular de múltiples regiones espaciadoras en dicho vector, estas regiones espaciadoras pueden ser diseñadas específicamente para cada uno de los objetivos transcripcionales.

### **3.2 Diseño de los gARN:**

Se empleó la herramienta CHOPCHOP (<http://chopchop.cbu.uib.no/>), para diseñar las secuencias de 20 pb complementarias a los genes a reprimir. Esta herramienta realiza un puntaje teniendo en cuenta: el porcentaje de GC de la región a la que une el gARN, regiones de auto-complementariedad, número de posibles off-target's, localización del sitio de unión de gARN (más cercano al extremo 5' mejor clasificación, más cercano al extremo 3' menor clasificación). Los gARN's serán sintetizados en forma de oligonucleótidos de cadena sencilla, con un tamaño de 96 nt, una vez seleccionada la secuencia de mejor puntaje

Para garantizar una represión más efectiva se deben seguir una serie de principios que deben regir el diseño de los gARN:

- El gARN debe de reconocer la cadena codificante del gen para garantizar una mayor eficiencia en la represión. Se priorizo aquellos gARN que reconocieran la región cercana al extremo 5' del gen seleccionado, ya que esta región favorece mayores niveles de represión. Teniendo en cuenta que debe existir la secuencia PAM (NGG o NAG) adyacente a dicha secuencia para garantizar el efecto deseado. También se debe tener en cuenta que no tiene que existir una complementariedad de bases estricta entre los 20 nucleótidos que participan en el reconocimiento de la secuencia diana, siendo los más importante aquellos nucleótidos que se

unirán a la secuencia semilla del gen de interés (nucleótidos adyacentes a la secuencia PAM).

- La especificidad genómica de los gARN es otra variable muy importante a tener en cuenta. Debido a que estos se podrían unir a cualquier secuencia de ADN con una homología muy similar, sobre todo si presentan secuencias PAM adyacentes. Para evitar esta interferencia se debe correr un BLAST para cada secuencia diana de los gARN contra todo el genoma del organismo, esto permitirá arrojar posibles secuencias no deseadas a las que se podrían unir los gran. Estas secuencias deberán ser analizadas para determinar su grado de homología y poder introducir mismatches (uno o dos son suficientes) en la región del gARN que se unirá a la región semilla del gen, lo que permitirá la reducción de la eficiencia de unión en un orden de magnitud como mínimo, lo que se traduce en una reducción casi total de la posibilidad de apareamiento indebido de los gARN's con secuencias que no son las deseadas.
- La calidad del plegamiento en el diseño de los gARN, es un factor crítico a considerar para ello una vez que se haya seleccionado la secuencia de 20 nucleótidos. Se debe analizar que no se afecte el plegamiento adecuado del gARN, a través de la comparación con gARN cuya estructura este validada, ya que si el asa por la que une dCas9 es modificada, el gARN no será capaz de unirse a esta proteína y la eficiencia de la represión disminuirá considerablemente (Hawkins *et al.*, 2017).

### **3.3 Detección de la liberación de recursos celulares:**

Como método de cuantificación de la liberación de recursos celulares en nuestro organismo modelo se utilizó un vector llamado pSEVA63- Dual con origen de replicación pBBR1, cuenta como marcador de selección un cassette de resistencia a gentamicina (Gyorgy *et al.*, 2015). Este plásmido fue ensamblado como un sistema que cuenta con dos proteínas reporteras fluorescentes, específicamente proteína verde fluorescente (GFP) y la proteína roja fluorescente (RFP). En este plásmido la GFP se encuentra bajo el control de un promotor constitutivo, por lo que siempre está emitiendo fluorescencia, la RFP está sujeta a la inducción de AHL, por lo que solo emitirá fluorescencia en presencia del inductor.

Este sistema permitió establecer un sistema donde no existe ninguna relación regulatoria entre ambos productos génicos en un mismo plásmido, en donde se podía observar una relación lineal entre los niveles de expresión de ambas proteínas cuando se sometía a inducción la RFP.

Por lo que un aumento creciente de la expresión de RFP, se producía una disminución en los niveles de expresión de GFP. Con este comportamiento se estableció una comparación con un concepto empleado en análisis de microeconomía, utilizado para describir el empleo de un presupuesto limitado, denominado líneas de isocosto, en este caso entiéndase como presupuesto limitado a los recursos celulares los cuales son también limitantes para la expresión de dicho circuito (**Figura 2.1**).

Se empleó el lector de placas Synergy H1, de la marca Biotek, con un protocolo de mediciones de la densidad óptica (OD) y mediciones de fluorescencia a la longitud de onda de excitación y emisión de las proteínas fluorescentes GFP y RFP cada 20 minutos, con una temperatura de incubación de 37 grados Celsius y agitación lineal, durante 24 horas.

### **3.4 Selección de los factores de transcripción:**

Solo se escogieron aquellos factores de transcripción (TF's) que modulen positivamente al menos un solo gen que codifique para una función no esencial en el metabolismo de *E. coli*. Teniendo en cuenta que dichos genes solo sean modulados por un único factor de transcripción y no presenten interacciones negativas y no podrán ser inducidos al remover otro factor de transcripción. Estos genes se considerarán como "gen real a silenciar".

También se tuvo en cuenta la carga proteómica que se liberaría con la represión de dichos factores para priorizar el silenciamiento de aquellos TF's, lo que garantiza que se libere una mayor carga con el silenciamiento de un factor de transcripción que regule al menos un gen que se expresa en cantidades mucho mayores que varios TF's que regulen genes que expresen diversas proteínas pero en baja concentración (Tesis de doctorado, Gustavo Lastiri).

### **3.5 Experimentos de liberación de recursos celulares:**

Se co-transformó la cepa silvestre BW25113 (WT), de dichas cepas. Esta cepa fue previamente transformada con el plásmido pSEVA63-Dual, anteriormente descrito. Esto permite cuantificar la liberación de carga proteómica a través de la emisión de GFP, a diferentes concentraciones de inductor (AHL) que influye en la expresión de la proteína RFP la cual necesita la presencia del activador LuxR. Posteriormente esta cepa WT- pSEVA63-DUAL, se co-transformó con las cuatro variantes de gARN's clonadas en el plásmido CRISPathbrick (CPB), denominadas como CPB-marA, CPB-csgD, CPB-rpoS, donde cada variante del plásmido CPB es llamado además con el nombre del factor de transcripción al cual reprimen. Por último en el caso de la combinación doble gARN's clonados en el vector se denominó como CPB-marA\_csgD, ya que contiene ambos gARN's.

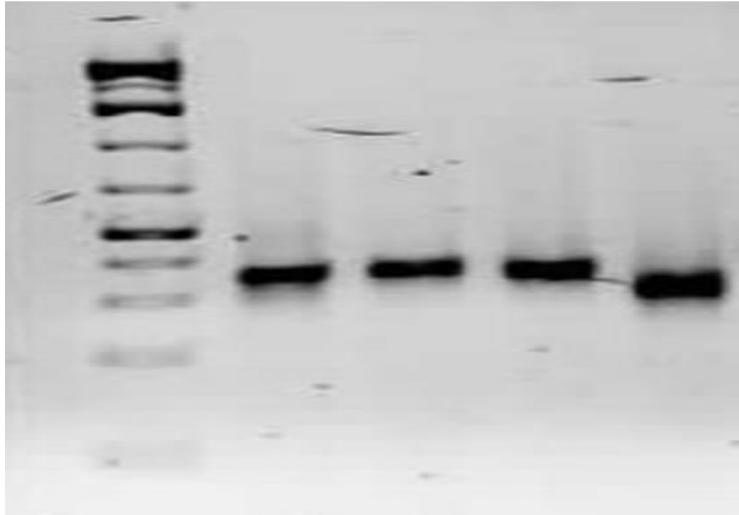
## **IV Resultados y Discusión**

### **4.1 Generación del sistema CRISPRi**

En el proceso de clonación de los gARN's diseñados para reprimir los diferentes reguladores en este estudio se utilizaron gARN's diseñados para unirse al extremo 5' de los marcos abiertos de lectura del factor sigma y los dos factores de transcripción: *rpoS*, *marA* y *csgD* respectivamente. El diseño fue realizado utilizando la herramienta CHOPCHOP teniendo en cuenta los parámetros principales de diseño.

Los objetivos genéticos a reprimir fueron seleccionados basándonos en la carga proteómica teórica que debe liberarse una vez se encuentren reprimidos cada uno de estos. Los factores de transcripción *marA* y *csgD* fueron seleccionados debido a ser la combinación de TF's predicha que mayor carga proteómica pueden liberar según análisis computacional de nuestro laboratorio (Gustavo Lastiri, Tesis de Doctorado en proceso).

Una vez sintetizados los oligonucleótidos de cadena sencilla, se prosiguió con el protocolo de ensamble descrito en la metodología de este trabajo. Posterior a la transformación por electroporación, se realizaron reacciones de PCR de colonia para comprobar la presencia de los insertos en nuestro plásmido por incremento de tamaño del amplicón.

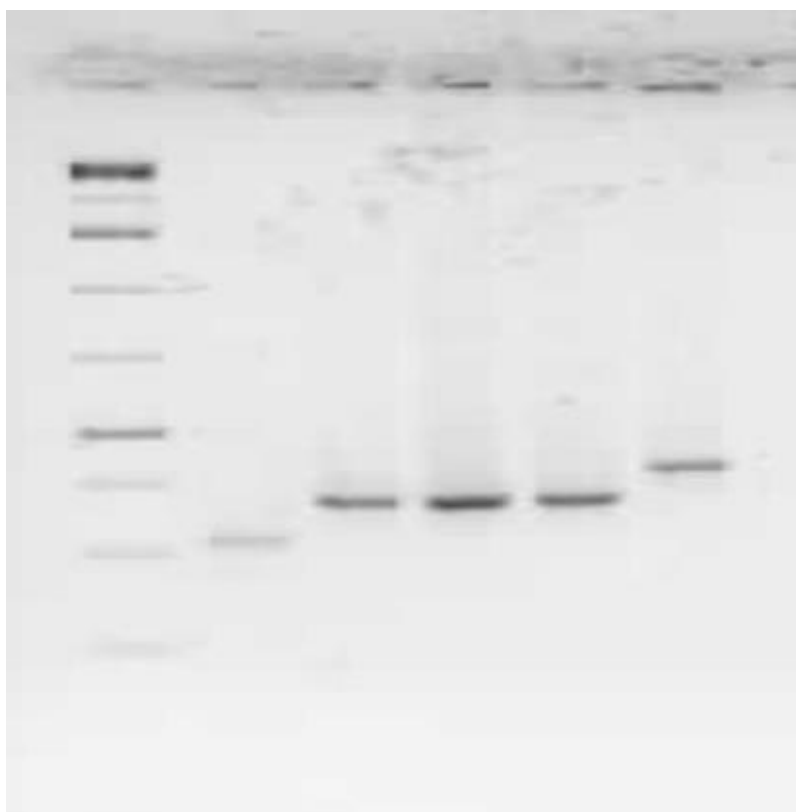


**Fotografía 4. 1.** Electroforesis en Gel de poliacrilamida 2% para comprobar mediante PCR la inserción de los gARN para silenciar *rpoS*, *marA* y *csgD*, Carril 1: marcador de peso molecular (GeneRuler 1kb), carril 2: CPB-*rpoS*, Carril 3: CPB-*marA*, Carril4: CPB-*csgD* Carril 5: CPB vacío (control negativo).

Como se puede observar en la **fotografía 4.1** donde se obtuvieron colonias donde el ensamble fue exitoso para los tres gARN's. Una vez identificadas las colonias positivas a la transformación con los gARN's correspondientes, estos plásmidos fueron extraídos, purificados y enviados a secuenciación, donde se comprobó la correcta presencia de la secuencia que se había diseñado previamente. Una vez comprobada la presencia de los insertos, se prosiguió a intentar clonar un segundo gARN en el vector, es decir se intentó realizar un segundo ensamble consecutivo en una versión del CRISPathBrick que ya contuviese un primer gARN insertado. Se seleccionó la variante CPB-*csgD* (CRISPathBrick-*csgD*), para este experimento de clonación de múltiples gARN's, y el gARN diseñado para reprimir el factor de transcripción *marA*, debido computacional que pueden liberar una mayor carga proteómica en *Escherichia coli*.

Se comprobó por PCR de colonias y por secuenciación la presencia del nuevo gARN clonado y la integridad de la secuencia, formando la variante con el doble gARN denominada como CPB-*csgD\_marA* (**fotografía 4.2**). Los oligonucleótidos diseñados para realizar la evaluación por PCR producen un amplicón de 315 nucleótidos, con la incorporación de cada nuevo gARN, se produce un aumento de 56 nucleótidos. a ser una de las combinaciones que están predichas por

nuestra herramienta De esta manera se puede comprobar la presencia o no de los insertos al compararse contra un control.



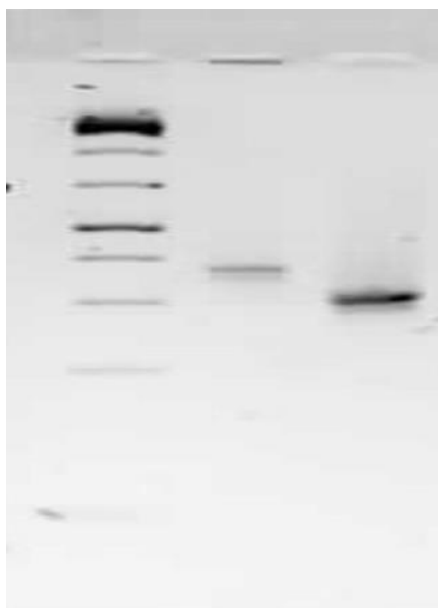
**Fotografía 4. 2** Gel al 2% de agarosa. En los pocillos 2, 3, 4 y 5 se ubican los productos de PCR del CPB, CPB-marA, CPB-csgD, CPB-rpoS y CPB-csgD\_marA. Obsérvese el incremento aproximado de 56 nucleótidos observado entre el plásmido

#### **4.2 Prueba de represión *in vivo*:**

En este experimento nos propusimos comprobar la capacidad real del sistema CRISPathBrick para inhibir la expresión de un gen, en nuestro caso se empleó un gen reportero que codifica para una proteína fluorescente (mCherry) como prueba de concepto. El gARN fue diseñado para unirse a la región promotora de la proteína mCherry la cual se expresa constitutivamente. Este gARN fue clonado en CPB utilizando el kit NEBuilder Hifi, siguiendo un protocolo que permite economizar el uso de este kit. Posteriormente se comprobó por PCR la inserción del gARN en el plásmido CRISPathBrick.

Para este amplicón se esperaba un tamaño de 468pb que son 56pb más que el plásmido sin el inserto. En la **fotografía 4.3** se puede observar la electroforesis en gel de poliacrilamida 2% donde se muestra el aumento del tamaño del

amplicón esperado en alrededor de 56 nt con respecto al control negativo (plásmido sin el inserto).

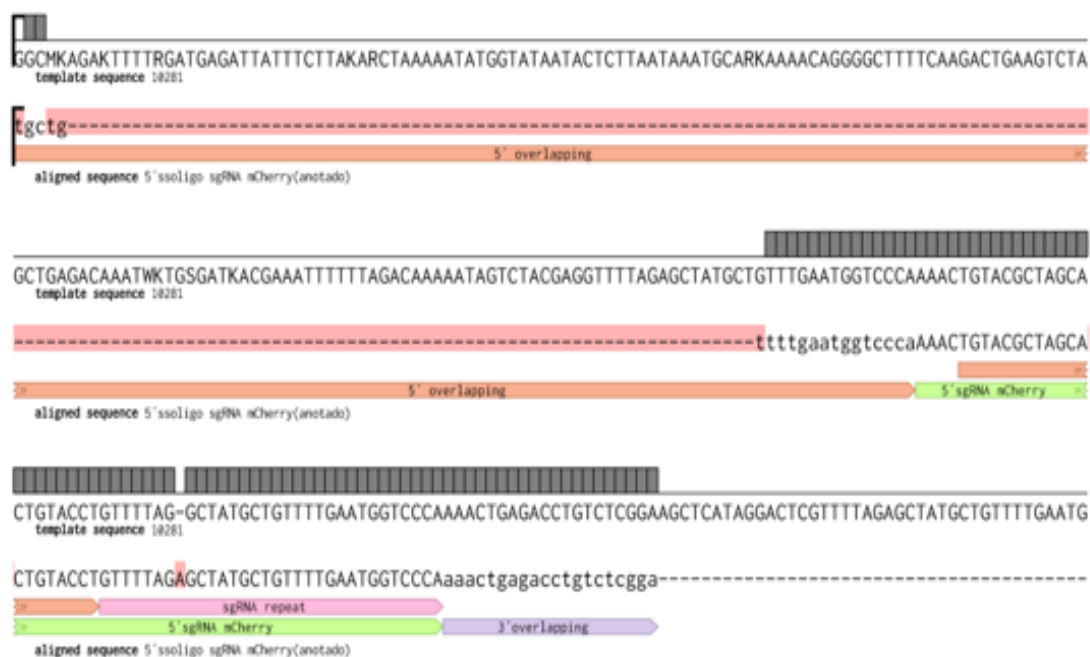


**Fotografía 4. 3.** Electroforesis en Gel de poliacrilamida 2% para comprobar mediante PCR la inserción del gARN que tiene como diana el promotor de la proteína fluorescente mCherry. Carril 1: marcador de peso molecular (GeneRuler 1kb), carril 2: plásmido CRISPathBrick con el gARN insertado. Carril 3: plásmido CRISPathBrick.

El vector CPB con el gARN fue secuenciado para comprobar la existencia y calidad de la secuencia. Los resultados de la secuenciación fueron exitosos, como se puede observar en la **figura 4.1**, por lo que posteriormente fue utilizado para transformar cepas de *E. coli* DH5-alpha, para comprobar su efecto biológico.

Como se puede observar en la **figura 4.1** no se encontró ninguna mutación importante en el inserto correspondiente al gARN diseñado para reprimir la proteína fluorescente mCherry. Es importante destacar que a pesar de encontrar una mutación en el inserto esta se encuentra fuera de la región semilla (seed region ó región I), clasificada así por Qi *et al.*, 2013. En este estudio, los autores indujeron mutaciones puntuales en cada uno de los 20 nucleótidos del gARN para comprobar la robustez del sistema. Los autores fueron capaces de discernir tres regiones fundamentales correlacionadas con la efectividad de represión y las mutaciones generadas.

## Secuencia del gARN mCherry clonado en CPB



**Figura 4.1.** Alineamiento de secuencias entre el diseño realizado *in silico* y la secuenciación de CPB con el sgARN diseñado para silenciar mCherry.

Determinando que aquellas mutaciones encontradas en los primeros siete nucleótidos próximos al extremo 5' del transcrito ocasionan una afectación considerable a la capacidad de represión del sistema dCas9/gARN. Como se puede observar en la **figura 4.1** la mutación ocurre en el nucleótido 8 del extremo 5' del inserto por lo que al momento de transcribirse dicha mutación queda ubicada bastante fuera de las regiones críticas que garantizan niveles de represión elevados.

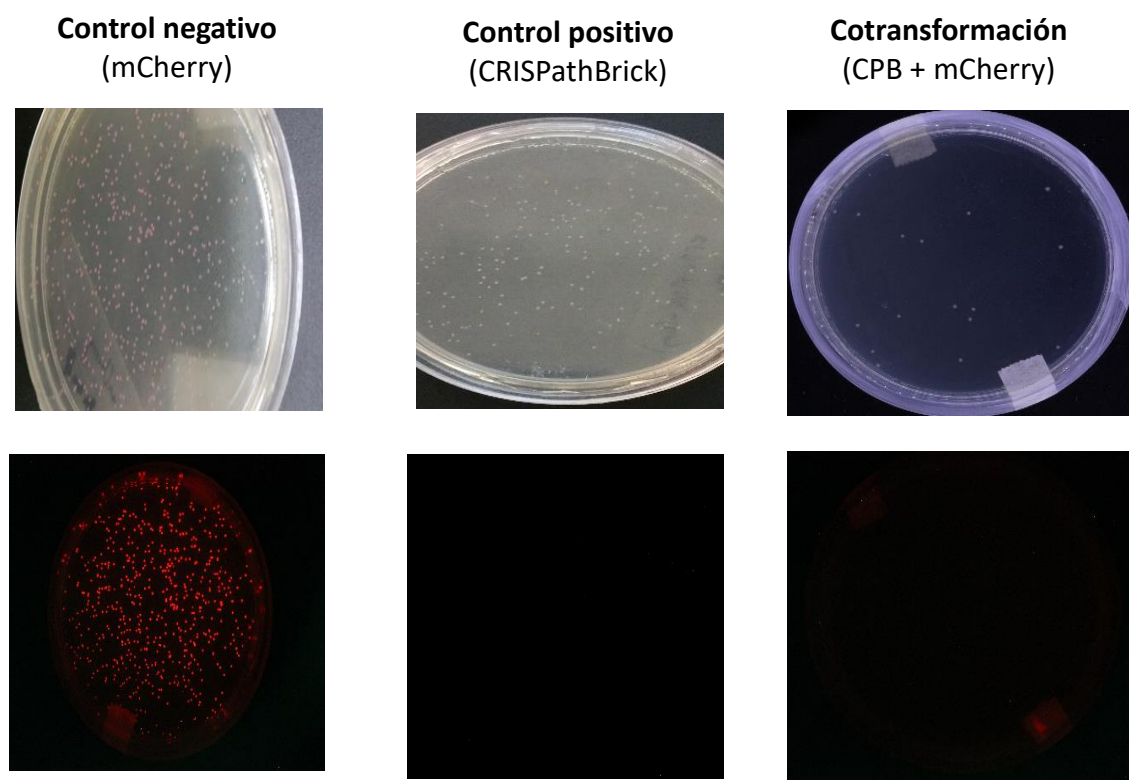
Posteriormente se efectuó la cotransformación de los plásmidos CPB-mCherry y mCherry/pTU 1-A-lacZ. Se comprobó el cambio fenotípico de las células contransformadas con ambos plásmidos, se empleó como control negativo *E. coli* DH5-alpha transformada únicamente con el plásmido mCherry/pTU 1-A-lacZ y como control positivo la cepa de *E. coli* DH5-alpha transformada únicamente con CRISPathBrick vacío. Para observar los resultados se tomaron fotografías con luz blanca y posteriormente las cajas Petri fueron irradiadas con luz a una longitud de onda capaz de excitar la mCherry y detectar la presencia o no de colonias fluorescentes, este tratamiento se realizó por igual a las cepas control como a la muestra. Luego de irradiar las cepas se observó que el control negativo



emitía fluorescencia, lo cual era de esperar pues estaban transformadas con el plásmido que porta la proteína fluorescente, demostrando que la proteína mCherry se expresa correctamente.

El control positivo no mostró fluorescencia, resultado también esperado pues no contiene la proteína fluorescente. Sin embargo, la cepa cotransformada mostró un cambio fenotípico entre las cepas control y la muestra, como se puede apreciar en la **fotografía 4.4**, donde se observa que al parecer esta cepa no expresa la proteína fluorescente lo que indica que el gARN es capaz de unirse al promotor de la proteína e inhibir su expresión.

Posteriormente se realizó un experimento dinámico con el objetivo de comprobar que efectivamente nuestro sistema era el responsable por la ausencia de fluorescencia (**figura 4.2**).

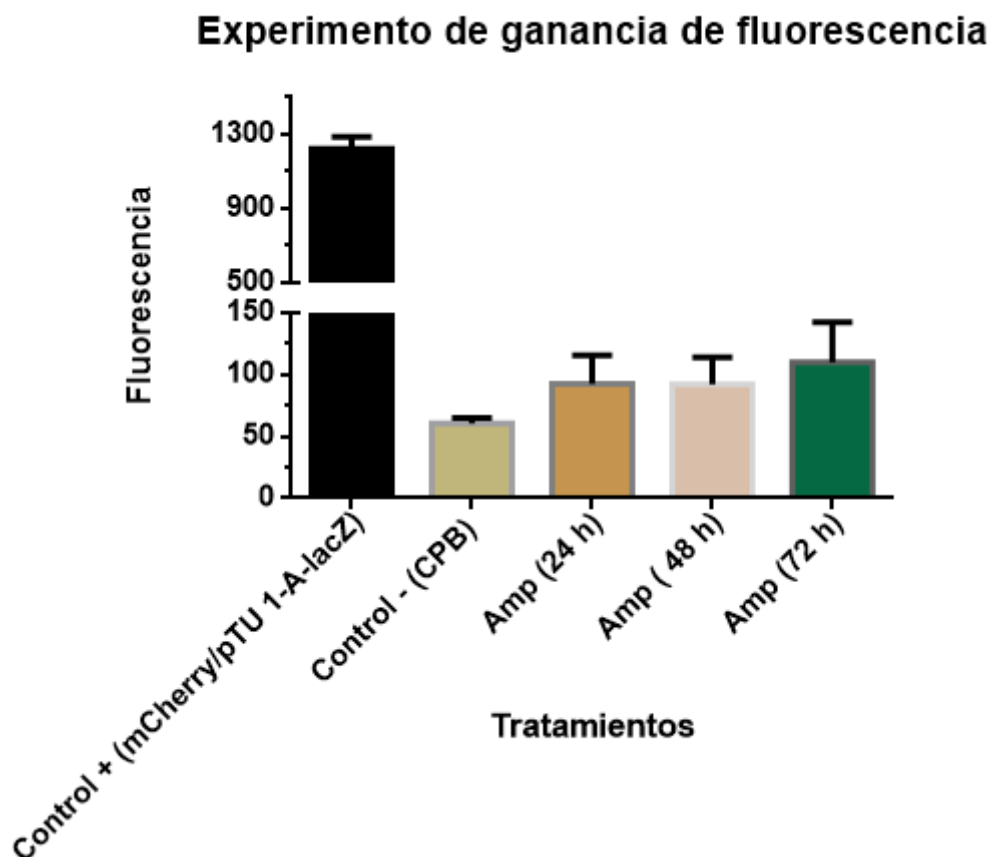


**Fotografía 4. 4.** Experimento de fluorescencia de doble transformación con CRISPathBrick y mCherry/pTU 1-A-lacZ (mCherry constitutiva). Se puede apreciar la ausencia de fluorescencia por parte del control (+) y de la muestra.

En este experimento se efectuó de dos formas en paralelo, las cepas fueron inoculadas en medio LB en placas de 96 pocillos diseñadas para poder cuantificar fluorescencia. Se inoculó la muestra cotransformada con dos

antibióticos: cloranfenicol y ampicilina, los cuales son marcadores de selección de CPB y mCherry/pTU 1-A-lacZ respectivamente. Simultáneamente se inoculó esta cepa en medio LB sin cloranfenicol, pero con ampicilina para observar el efecto de la ausencia de presión selectiva (antibiótico cloranfenicol) en la cepa cotransformada.

También se emplearon los controles pertinentes: una cepa con mCherry/pTU 1-A-lacZ como control positivo y otra cepa con CPB como control negativo. Se realizaron subcultivos de la misma muestra cada 24 horas durante tres días y se midió tanto la fluorescencia como la densidad óptica. Como se puede observar a partir del tercer día de subcultivo se puede observar un aumento en los niveles de fluorescencia correspondientes. Lo que sugiere que con los pases consecutivos de esta cepa en medio LB sin cloranfenicol (marcador de resistencia del CPB) pero con ampicilina (marcador de resistencia del plásmido que codifica para la proteína mCherry) como



**Figura 4. 2.** Cambio en la fluorescencia de las colonias cotransformadas con los plásmidos CPB y mCherry/pTU 1-A-lacZ. Los subcultivos se realizaron cada 24 horas durante 3 días en medio LB suplementado solamente con ampicilina, como presión de selección.

mecanismo para eliminar la presión de selección que favorece la permanencia en la célula del plásmido CPB. Favoreciéndose el crecimiento de las cepas que perdieron dicho plásmido pero mantienen el plásmido que expresa mCherry al estar en presencia de ampicilina en el medio. La recuperación de la fluorescencia una vez que se pierde el CPB indica que se perdió parcialmente el sistema gARN-dCas9 que es capaz de actuar sobre el promotor de la mCherry. Debido a la carga metabólica que ejercen ambos plásmidos sobre las cepas cotransformadas, estas no pueden crecer a la misma velocidad que las cepas que pierden el CPB.

Por lo que la población que pierde el CPB al no poseer esta presión de selección es capaz de destinar más recursos celulares a crecer más rápidamente. Debido a un aumento en la disponibilidad de dichos recursos provocado por la ausencia del plásmido CPB. Comprobándose la efectividad del sistema la capacidad de represión del plásmido CRISPathBrick y validándose la metodología de diseño de los gARN's *in vivo* durante el experimento dinámico de pérdida del CPB.

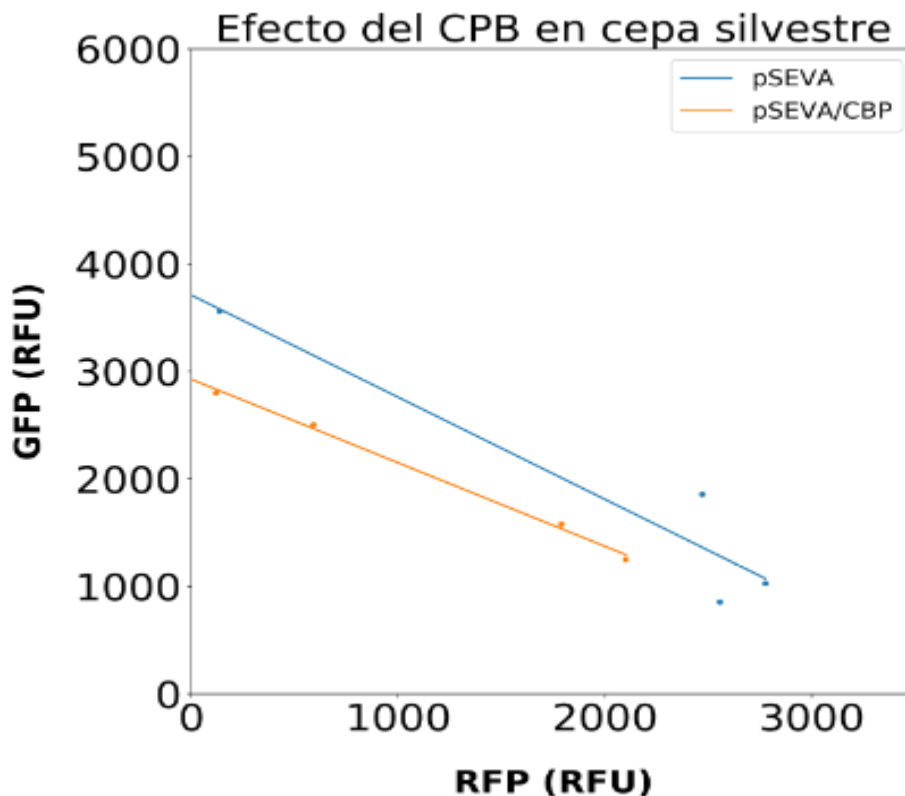
#### **4.3 Efecto de la expresión de dCas9 en la disponibilidad de recursos celulares**

En este experimento tuvo como propósito comprobar la carga metabólica que provoca el plásmido CPB vacío (sin gARN clonado) en la célula, así como el efecto del plásmido pSEVA63-Dual. Midiéndose el efecto por separado y durante la cotransformación con ambos plásmidos.

Como se puede observar en la **figura 4.3**, la cepa cotransformada (pSEVA/CPB) presenta una disminución significativa en la producción de proteína fluorescente en comparación con la cepa transformada solamente con el pSEVA63-Dual (pSEVA). Esto pudiera estar relacionado con el efecto tóxico que puede presentar la dCas9 en *E. coli* reportado recientemente (Cui *et al.*, 2018; Sho *et al.*, 2018). Una de las causas atribuidas a esta toxicidad está relacionada con el hecho que la dCas9 es capaz de unirse al ADN de doble cadena sin la presencia del gARN correspondiente lo que conlleva a múltiples actos de unión al ADN y rastreo de la secuencia PAM, aunque esta unión es temporal, a diferencia de la unión al ADN mediada por un gARN (mucho más fuerte), este evento puede

ocasionar la interrupción de eventos moleculares de transcripción, que aunque sean momentáneos, pueden ocurrir en múltiples regiones del genoma generando el efecto de toxicidad observado en otros estudios.

Otra posible causa de este efecto en la fluorescencia se debe a que la cepa transformada con el plásmido CRISPathBrick vacío, el cual no libera ningún recurso celular, solo es capaz de generar más carga sobre la célula.



**Figura 4. 3.** Se grafican las líneas de isocosto de las cepas transformadas con el plásmido pSEVA63-Dual y la cepa cotransformada con dicho plásmido y el CPB simultáneamente.

Factores como el número de copias de este plásmido (5 – 10 copias), pero sobre todo la demanda de recursos celulares que conlleva el mantenimiento del mismo, teniendo en cuenta que la proteína dCas9 es una proteína de 160 kDa con más de 1000 aminoácidos en su estructura, la cual se expresa de manera constitutiva, además del costo que conlleva expresar las proteínas propias del vector que le confieren resistencia a cloranfenicol.

Otra razón que puede explicar esta diferencia tan marcada en la emisión de fluorescencia se debe a que la cepa transformada con el plásmido pSEVA63-Dual solamente solo está sometida a la carga proteómica de un solo plásmido,

por tanto presentar un mejor comportamiento en las mediciones de fluorescencia total, debido a que los recursos que posee pueden ser utilizados de manera más eficiente para el crecimiento y para la producción de proteína fluorescente.

#### **4.4 Experimento cinético de liberación de recursos celulares:**

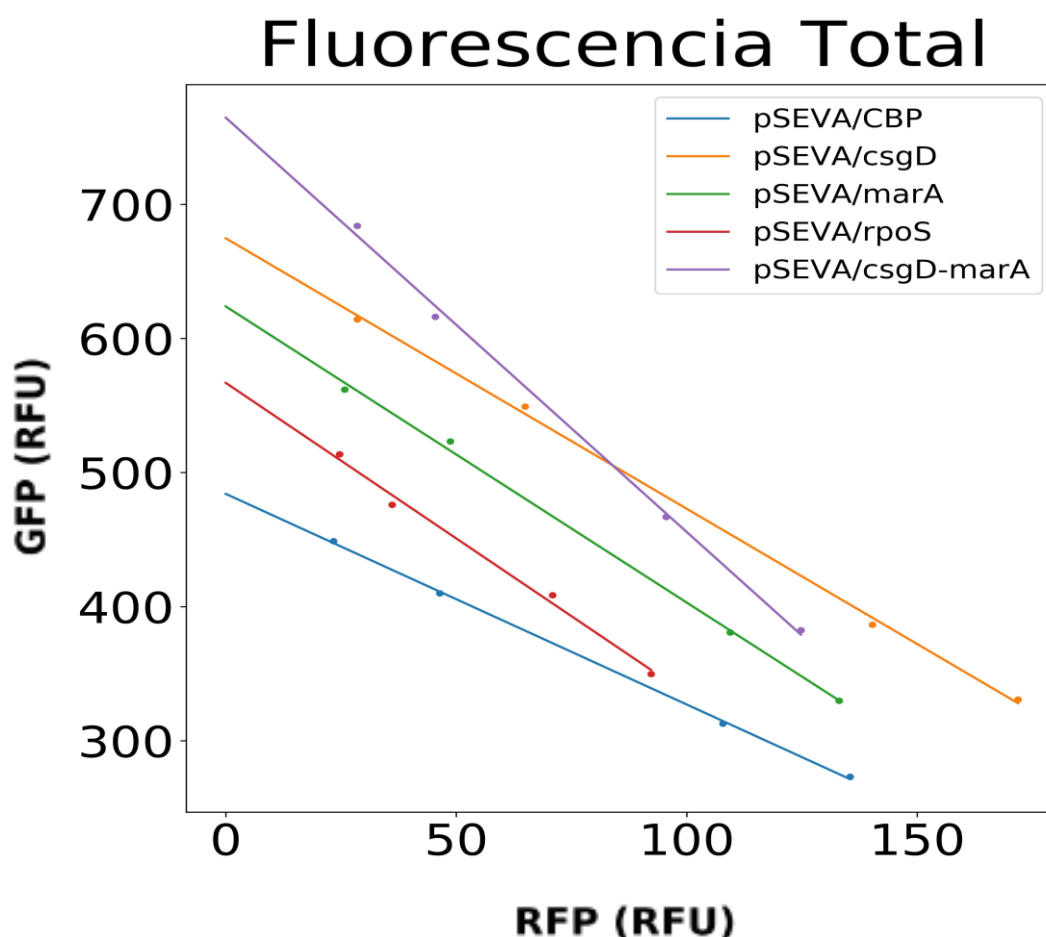
Una vez confirmada la presencia del doble gARN clonado en el plásmido CRISPathBrick y la secuencia en correcto estado (**Fotografías 4.1 y 4.2**), se prosiguió a la elaboración y ejecución de un experimento cinético para comprobar si el silenciamiento de la expresión de los TF's con el sistema CRISPRi es capaz de liberar recursos celulares.

El efecto de los gARN's que fuesen capaces de apagar los genes reguladores seleccionados y liberar recursos celulares, fueron estudiados mediante la elaboración de líneas de isocosto que tuviesen en cuenta la fluorescencia de ambas proteínas heterólogas (GFP y RFP). Se cotransformaron la cepa WT, con el plásmido pSEVA63-Dual (origen de replicación pBBR1), el cual posee un origen de replicación compatible con el origen de replicación del CRISPathBrick (origen de replicación p15A), con lo que se elimina la posibilidad de pérdida de alguno de los plásmidos siempre y cuando se mantenga la presión de selección que en el caso del pSEVA63-Dual presenta un casete de resistencia a gentamicina y el CRISPathBrick, presenta un casete diferente que le proporciona resistencia a cloranfenicol.

En el diseño del experimento se tuvo en cuenta las concentraciones de inductor AHL que son necesarias para alcanzar niveles de producción de la proteína roja fluorescente (RFP) que fuesen perceptibles por el lector de placas pero al mismo tiempo produjesen una variación en los niveles de expresión de RFP lo suficientemente diferentes como para efectuar un gradiente adecuado. Se seleccionaron tres concentraciones de inductor y una condición con la ausencia del inductor AHL (0 nM, 2.5 nM, 10 nM, 20 nM), donde cada uno de los puntos de inducción representa una de estas concentración de AHL. En este experimento se realizaron tres replicas biológicas y dos replicas estadísticas para cada una de las condiciones evaluadas en medio M9 suplementado con glucosa como única fuente de carbono. En la **figura 4.4**, se puede observar el comportamiento de los diferentes gARN's, en el denominado *estado estable*,

donde la producción de proteína es constante en el tiempo con respecto al aumento de la biomasa y su tasa es máxima, en el caso de nuestro sistema de estudio con *E. coli* este estado se alcanza a partir de las cinco horas de incubación en este medio aproximadamente.

Se cuantificó la emisión de fluorescencia (GFP y RFP) y con estos datos se elaboraron las líneas de isocosto correspondientes, tomando las emisiones de fluorescencia totales de las diferentes variantes de gARN's y se puede observar que en todos los casos evaluados se pudo producir más fluorescencia total que la cepa pSEVA/CPB que usamos como comparación pues lleva el plasmido CPB sin gARN clonado. La variante CPB/rpoS fue la variante de menor capacidad liberadora de recursos celulares en este ensayo después de la variante con el plásmido vacío (pSEVA/CPB), esto puede estar relacionado con el hecho que en



**Figura 4. 4.** Se grafican las líneas de isocosto de la fluorescencia total (medida en unidades relativas de fluorescencia), para cada uno de los gARN's en la cepa evaluada (WT) a las cinco horas de cultivo (*estado estable*).

este momento fisiológico de la célula (crecimiento exponencial), los diferentes genes cuya expresión es controlada por *rpoS* no se encuentran inducidos.

Debido a que *rpoS* es un factor sigma o *master regulator* que actúa sobre una cantidad masiva de genes relacionados con la contención de diferentes situaciones de estrés, cuya expresión es necesaria cuando empieza a aparecer la escases de nutrientes en el ambiente, iniciándose la fase estacionaria del crecimiento (Griffith *et al.*, 2004), tiene mucho sentido que esta sea la variante que menos recursos celulares libere en este estadio celular.

Una de las variantes de mayor liberación de recursos celulares fue pSEVA/csgD, esto se justifica por el hecho que este factor de transcripción está fuertemente relacionado con la expresión de genes involucrados en la síntesis del biofilm. El proceso de síntesis de los componentes del biofilm es muy costoso para la célula, pero también la expresión del conjunto de genes que están involucrados en la síntesis del mismo tiene un alto costo celular (Ogasawara *et al.*, 2011).

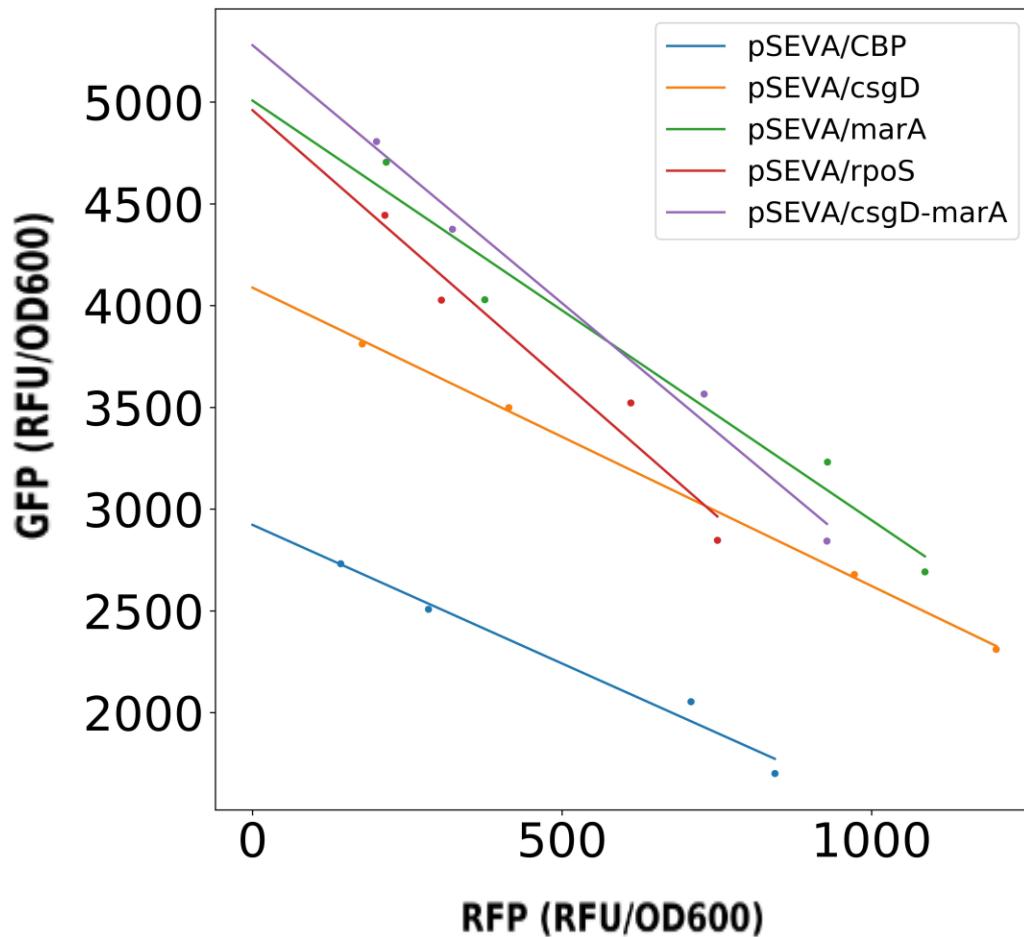
La variante más significativa en cuanto a la liberación de recursos celulares fue la combinación pSEVA/csgD\_marA, debido a que estos dos factores de transcripción, según análisis efectuados en nuestro laboratorio son de los factores de transcripción que mayor carga proteómica pueden liberar del total de factores de transcripción analizados, esto se debe al gran número de genes asociados a dichos factores de transcripción y a los niveles de expresión elevados reportados para algunos de dichos genes. La represión de ambos factores de transcripción al parecer libera una mayor cantidad de recursos celulares que actuando por separado, lo que se traduce en un incremento de la fluorescencia total.

Al normalizar (**figura 4.5**) los datos de fluorescencia divididos por la densidad óptica de cada una de las cepas evaluadas, podemos observar que en todos los casos de los factores de transcripción reprimidos por nuestros gARN's liberan una mayor cantidad de fluorescencia, incluso que la propia cepa control pSEVA/CPB de manera muy significativa. Esto sugiere que la producción de proteína heteróloga en nuestras cepas con genes reguladores reprimidos es mayor que la cepa control (pSEVA/CPB).

En este caso se puede observar que la cepa transformada con el CRISPathBrick vacío solamente mantiene el mismo comportamiento observado en la **figura 4.4**, siendo la variante con menor producción de fluorescencia, debido a la gran carga

proteómica a la que se sometida la célula y a la ausencia del efecto de liberación de recursos celulares a través de la represión de factores de transcripción presente en las demás cepas cotransformadas. Se puede observar claramente el efecto de rescate de la fluorescencia producida por la acción de los gARN's.

## Fluorescencia Normalizada



**Figura 4. 5.** Se grafican las líneas de isocosto de la fluorescencia normalizada (medida en unidades relativas de fluorescencia), para cada uno de los gARN's en la cepa evaluada (WT) a las cinco horas de cultivo (estado estable).

La combinación pSEVA/csgD\_marA también se mantuvo como la variante de mayor liberación de recursos celulares, seguida de la cepa transformada con el gARN que reprime al factor de transcripción *marA*, este factor de transcripción está involucrado en la resistencia basal a antibióticos fundamentalmente, por lo que al estar reprimido no induce la expresión de un conjunto de proteínas involucradas en la degradación de antibióticos a bajas concentraciones, que son excretados en el ambiente como un mecanismo para inhibir el crecimiento de otros grupos bacterianos y eliminar así la competencia potencial por nutrientes



como ocurre en la naturaleza. De esta manera estos recursos celulares pueden ser aprovechados/desviados para incrementar la producción de proteína heteróloga, en nuestro caso GFP y RFP.

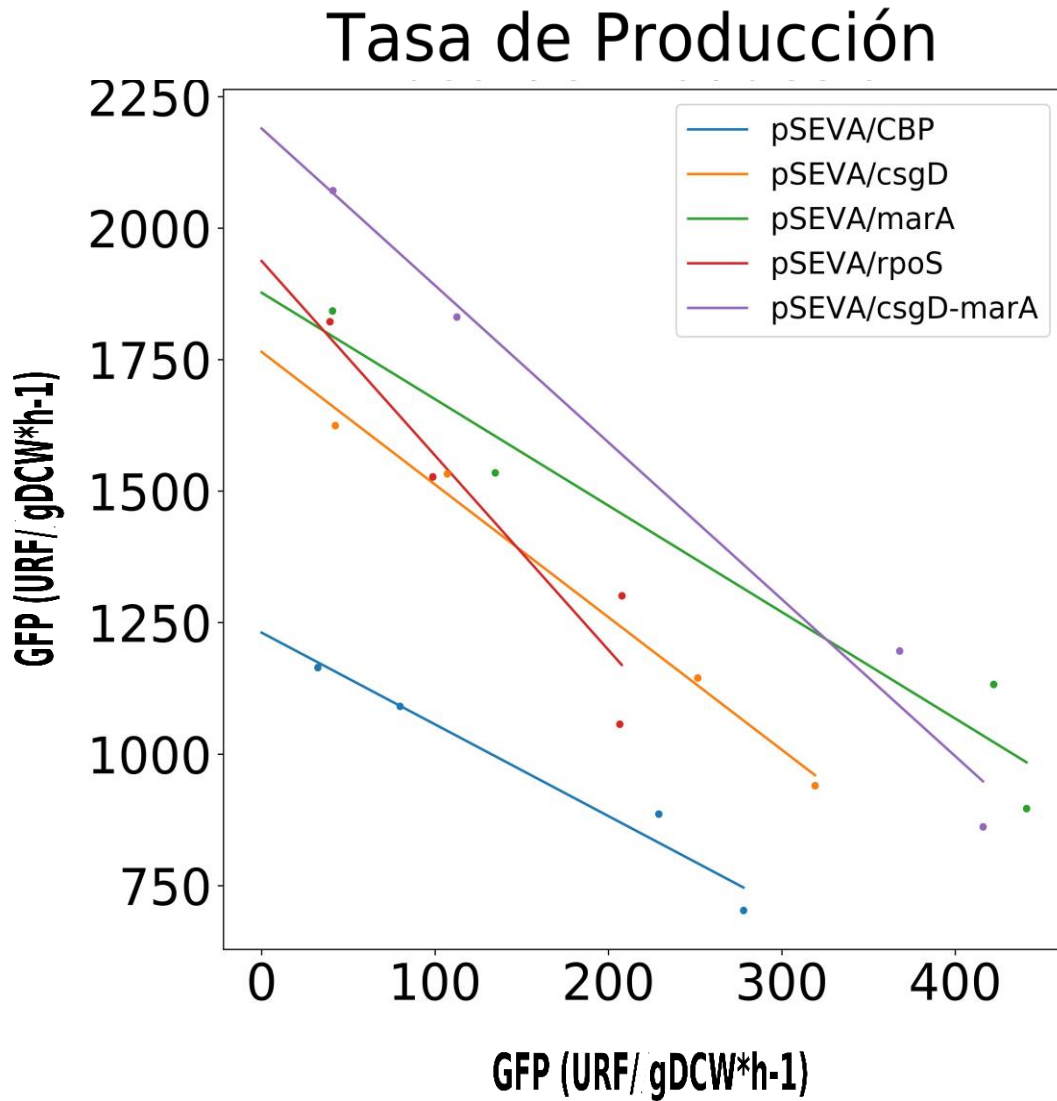
El factor de transcripción *csgD*, como se mencionó anteriormente está involucrado en la producción del biofilm, este mecanismo es necesario para garantizar la supervivencia de la colonia en condiciones de estrés por presencia de antibióticos y por ausencia de nutrientes, entre otros factores. Sin embargo esta función no es necesaria en las condiciones de cultivo empleadas en la industria biotecnología en sentido general por lo que la represión de este factor de transcripción el cual regula una cantidad importante de genes con niveles de expresión considerables, permite que estos recursos celulares puedan ser desviados a la expresión de proteína fluorescente.

La cepa transformada con la combinación pSEVA/*rpoS*, presentó un comportamiento mejor que en el gráfico de fluorescencia total “produciendo” mayores niveles de fluorescencia que la variante pSEVA/*csgD*.

En este análisis pudimos observar que la variante pSEVA/*rpoS* fue de las variantes que mayor fluorescencia produjo, a diferencia de la **figura 4.5** donde fue de las combinaciones que menores recursos liberó, esto nos sugiere que la variante pSEVA/*csgD*, es capaz de producir mucha biomasa y al efectuar el análisis de fluorescencia normalizada por la densidad óptica se puede comprobar que realmente es capaz de liberar menos recursos celulares de los pensado inicialmente. Este análisis está en consonancia con el hecho de que si bien *rpoS* tiene sus máximos niveles de inducción en la fase estacionaria del crecimiento, desde la propia fase exponencial pueden estarse expresando, quizás como un mecanismo de heterogeneidad poblacional anteriormente descrito. La disponibilidad recursos celulares no está limitado solamente a recursos como ATP, aminoácidos y poder reductor, otros factores también tienen un gran peso en la economía celular.

También se evaluó como nuestros gARN's afectaban la tasa de producción de proteína fluorescente, como se puede observar en la **figura 4.6** En este caso se puede observar un comportamiento muy similar al observado en la **figura 4.5**, donde una vez más los gARN's evaluados son capaces de incrementar la tasa de producción de proteína heteróloga en la célula. Siendo esta producción más rápida que la propia cepa control pSEVA/CPB, a pesar de que esta cepa como

ya se ha aclarado anteriormente se encuentra sometida a la carga metabólica de un solo plásmido (pSEVA63-Dual) y por tanto su carga proteómica debiese ser menor que las cepas cotransformadas.



**Figura 4. 6.** Se grafican las líneas de isocosto de la fluorescencia teniendo en cuenta la tasa de producción de proteína, para cada uno de los gARN's en la cepa evaluada a las cinco horas de cultivo (estado estable).

La variante pSEVA/CPB mantiene su comportamiento de ser la variante de menor tasa de producción de fluorescencia debido a la ausencia de gARN's que liberen recursos celulares. Podemos observar que la variante pSEVA/marA, fue de las mejores variantes al inducir una tasa de producción de proteína fluorescente de las más elevadas. La combinación doble de nuestro experimento (pSEVA/csgD\_marA), volvió a mantener su comportamiento como la variante de mayor tasa de producción de proteína heteróloga, lo que confirma el hecho de

que los recursos celulares que son liberados por esta variante son superiores al resto de combinaciones en las que se emplean un solo gARN reprimiendo a un solo gen regulador. Esto sugiere que el efecto en la capacidad de producción de proteína heteróloga provocada por la acción de nuestros gARN's es lo suficientemente grande como para superar la propia carga proteómica que provoca el pCRISPathBrick por sí solo. Liberando los recursos celulares suficientes, ya que al menos en los casos evaluados se puede observar una mayor producción de proteínas fluorescentes que la propia cepa control (pSEVA/CPB).

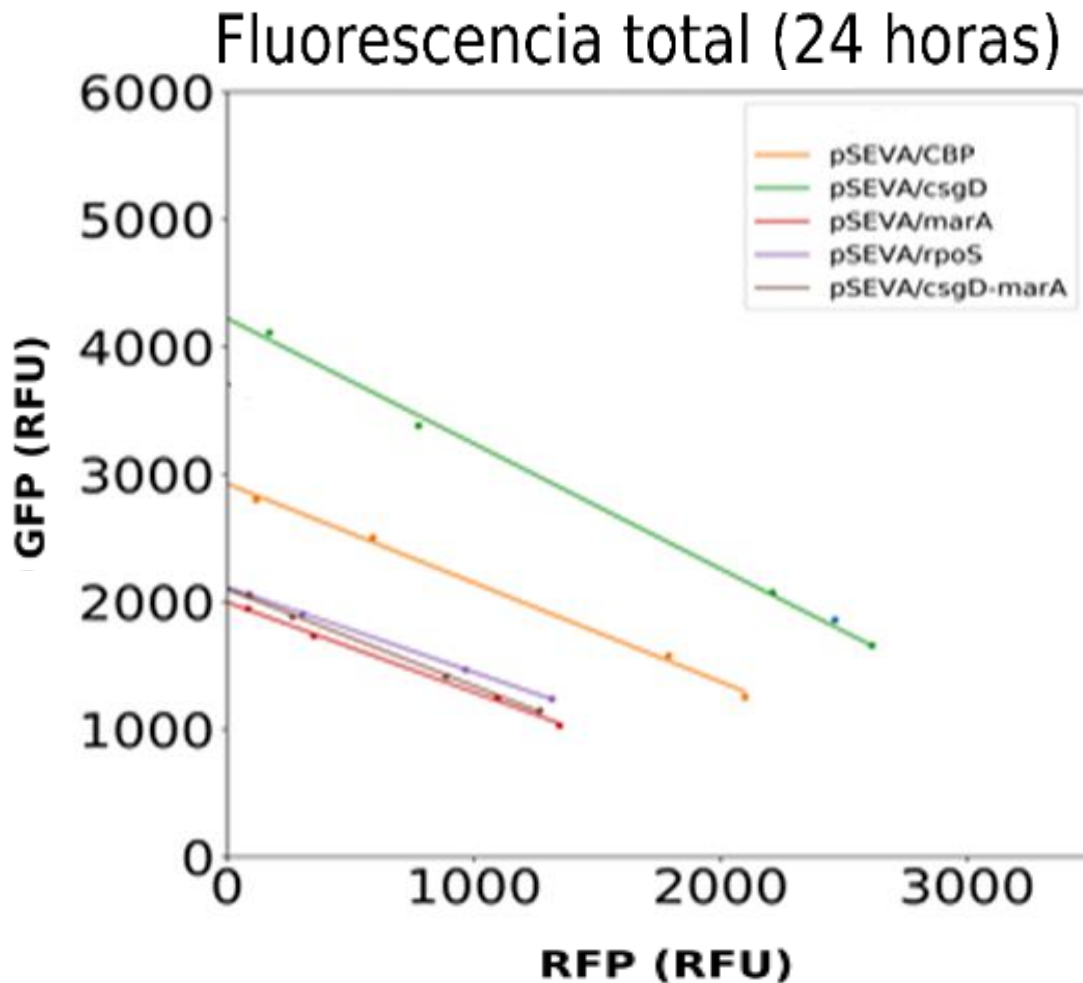
Resultó interesante que el comportamiento de los gARN's utilizados en este estudio fue un poco diferente en determinados momentos de la cinética, como se puede apreciar en la **figura 4.7**, donde se grafica la fluorescencia a las 24 horas de cinética. Notamos que la variante del plásmido con mejor comportamiento fue pSEVA/csgD, el cual fue que alcanzo los mayores niveles de emisión de fluorescencia al final de la cinética. Un resultado un poco desconcertante es el caso donde la cepa transformada con el plásmido vacío (pSEVA/CPB), fue capaz de producir mayor fluorescencia total que casi todos los gARN's evaluados, incluyendo la cepa transformada con la doble combinación (csgD-marA).

Una razón que pudiera justificar este tipo de comportamiento es que al no liberar ningún recurso celular la célula percibe esta carga proteómica extra a la que se encuentra sometida por el costo de mantenimiento y expresión de ambos plásmidos, sin embargo se sigue priorizando el crecimiento por sobre la producción de proteínas no asociadas con el crecimiento directamente. Provocando el efecto de producir una mayor fluorescencia, cuando realmente lo que ocurre es un aumento en la biomasa. Es importante recalcar que cuando se plantea la liberación de recursos celulares también estamos incluyendo la disponibilidad de maquinaria transcripcional como la ARN polimerasa, y maquinaria traduccional como son los ribosomas. Los cuales son constituyentes de la célula extremadamente importantes además los otros recursos celulares más conocidos como aminoácidos, ATP y poder reductor para realizar las reacciones de síntesis.

En el caso de la **figura 4.7**, pensamos que quizás en este estadio fisiológico de la célula (fase estacionaria), la represión de este factor de transcripción libere

una cantidad considerable de recursos celulares que son redirigidos a las funciones de crecimiento.

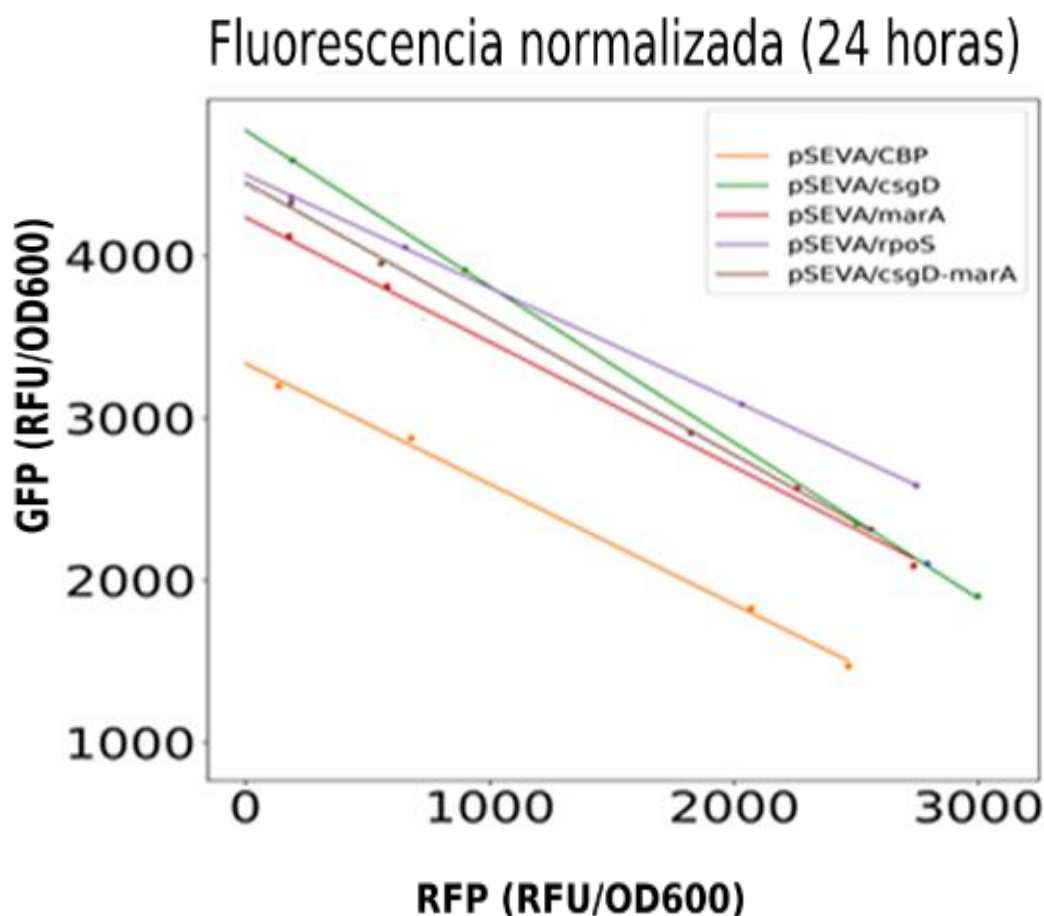
Esta hipótesis se pudo comprobar cuando se realizó el análisis con los datos de fluorescencia normalizados, como se puede observar en la **figura 4.8**. En esta grafica se puede observar el comportamiento esperado teniendo en cuenta solamente la producción de proteína fluorescente neta en cada una de las cepas evaluadas.



**Figura 4. 7.** Se grafican las líneas de isocosto de la fluorescencia total para cada uno de los gARN's evaluados al final de la cinética (**24 horas**).

La cepa transformada con el plásmido vacío (pSEVA/CPB), presentó las menores mediciones de fluorescencia normalizada, lo que sugiere que al igual que en el estado fisiológico anteriormente discutido, esta cepa es sometida a la mayor carga proteómica y su producción de GFP/RFP es la menor observada. Por otra parte, obtuvimos que todos los genes reguladores reprimidos con nuestros gARN's son capaces de inducir una mayor producción de fluorescencia

debido a su acción represora sobre nuestras dianas genéticas y la consecuente liberación de recursos celulares obtenida como resultado.



**Figura 4. 8.** Se grafican las líneas de isocosto de la fluorescencia normalizada, para cada uno de los gARN's en la cepa evaluada, al final de la cinética (24 horas).

En este análisis la variante pSEVA/csgD se mantuvo como uno de las de mayor capacidad liberadora de recursos celulares, lo cual posee una lógica biológica detrás bastante contundente a nuestro entender. Numerosos estudios han descrito la importancia de la producción de biofilm para las comunidades bacterianas en condiciones de estrés, a nuestro entender estas condiciones de estrés son alcanzadas en este momento de la cinética debido a la sobrepoblación y la escase de nutrientes. Esta situación que no se podía observar en las mediciones realizadas en el *estado estable* de la bacteria, donde los nutrientes abundan.

También se conoce de la literatura que el factor de transcripción *csgD* no solo regula los genes involucrados en la síntesis del biofilm, también regula positivamente genes involucrados en la síntesis de las fimbrias, las cuales son

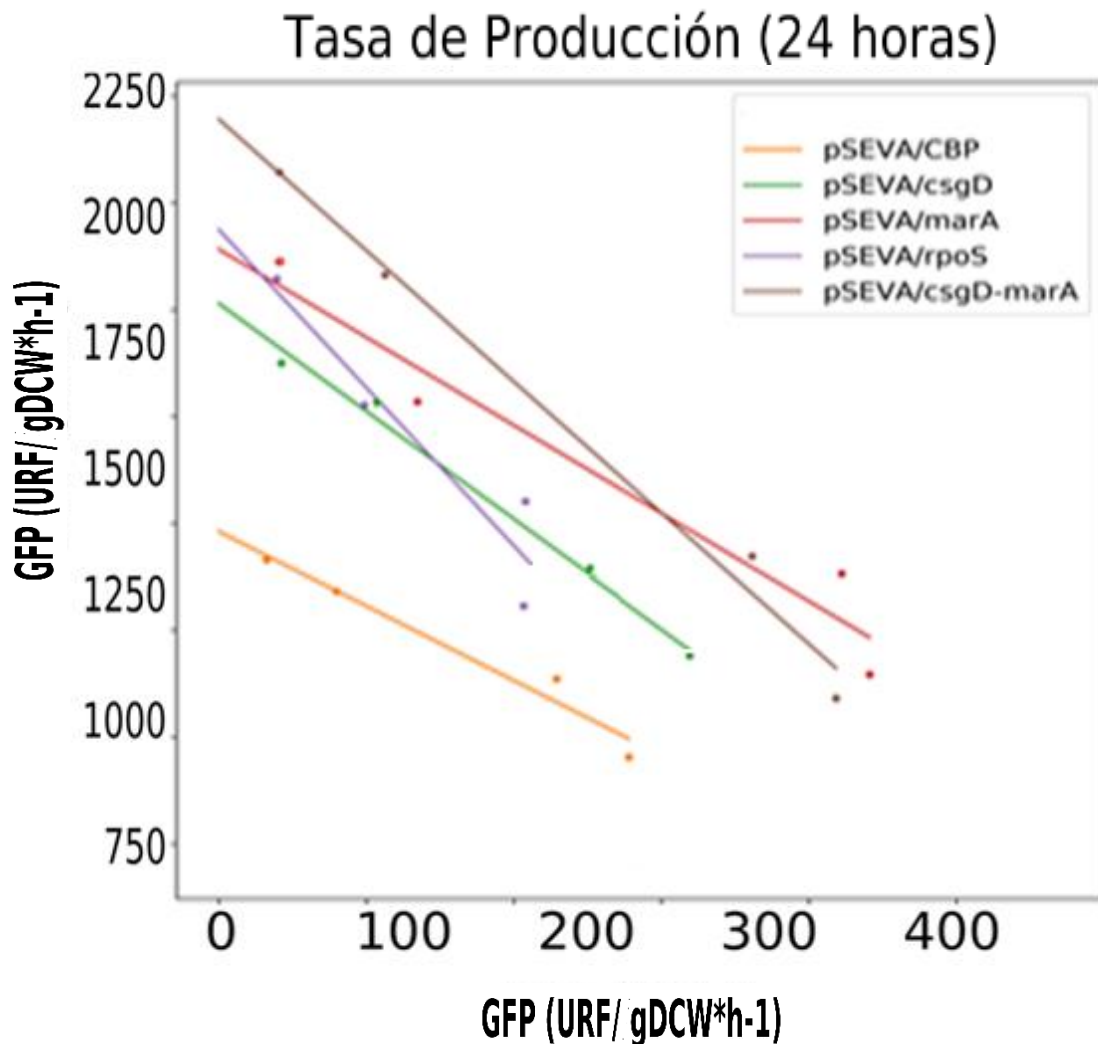
estructuras macromoleculares involucradas en favorecer la adhesión a superficies y también la adhesión célula-célula, lo cual se pudiera interpretar con un aumento en la disponibilidad de recursos celulares considerable en este estadio fisiológico en particular.

En el análisis de estos resultados no pudimos dejar de observar que la doble combinación de gARN's, pSEVA/csgD\_marA, la cual, fue la variante con mejor comportamiento en los análisis anteriores, no presentó el mismo comportamiento. Si bien es cierto que se puede ubicar entre las variantes que producen una elevada fluorescencia no fue la mejor, sugiriendo que al menos en este caso no significa necesariamente un mejor comportamiento.

Este análisis nos sugiere que el estadio fisiológico es un factor muy importante a la hora del diseño de candidatos a reprimir siguiendo la metodología expuesta en este trabajo. Debido a los propios cambios en la red de regulación transcripcional de *E. coli*, puede generar un efecto similar al observado cuando se emplean cepas con diferente fondo genético, es de suponer que el comportamiento sea también diferente. En el caso del factor sigma *rpoS*, hasta este momento había presentado un comportamiento relativamente modesto, sin embargo, al final de la cinética este comportamiento empieza a cambiar produciendo niveles de fluorescencia significativos incluso a las mayores concentraciones de inductor. Esto se debe a la propia naturaleza de la función biológica que desempeña, ya que está involucrado en la regulación de un número elevado de genes que a su vez están involucrados en garantizar la supervivencia de la célula en periodos de carencia de nutrientes y disímiles condiciones de estrés.

Resulta lógico suponer que si este factor sigma *rpoS* se encuentra reprimido en un momento donde los nutrientes son limitantes no se debe activar toda la respuesta celular para contender con este estrés, con lo que no se desvían recursos celulares a la expresión de un número significativo de genes involucrados en esta función y estos recursos celulares quedan disponibles para la producción de proteínas fluorescentes. Quedando también disponibles los recursos celulares que debiesen ser utilizados por los productos génicos de estos genes bajo el control de *rpoS* para realizar su función biológica.

En el caso de la doble combinación de gARN's (CPB-csgD\_marA), si bien no fue capaz de producir los mayores niveles de fluorescencia, sin embargo fue capaz de tener la mayor tasa de producción como se puede observar en la **figura 4.9**. Lo que sugiere que si bien en este momento particular de la cinética, no es la variante que favorece la mayor emisión de fluorescencia, si es la variante que produce proteína heteróloga más rápidamente.



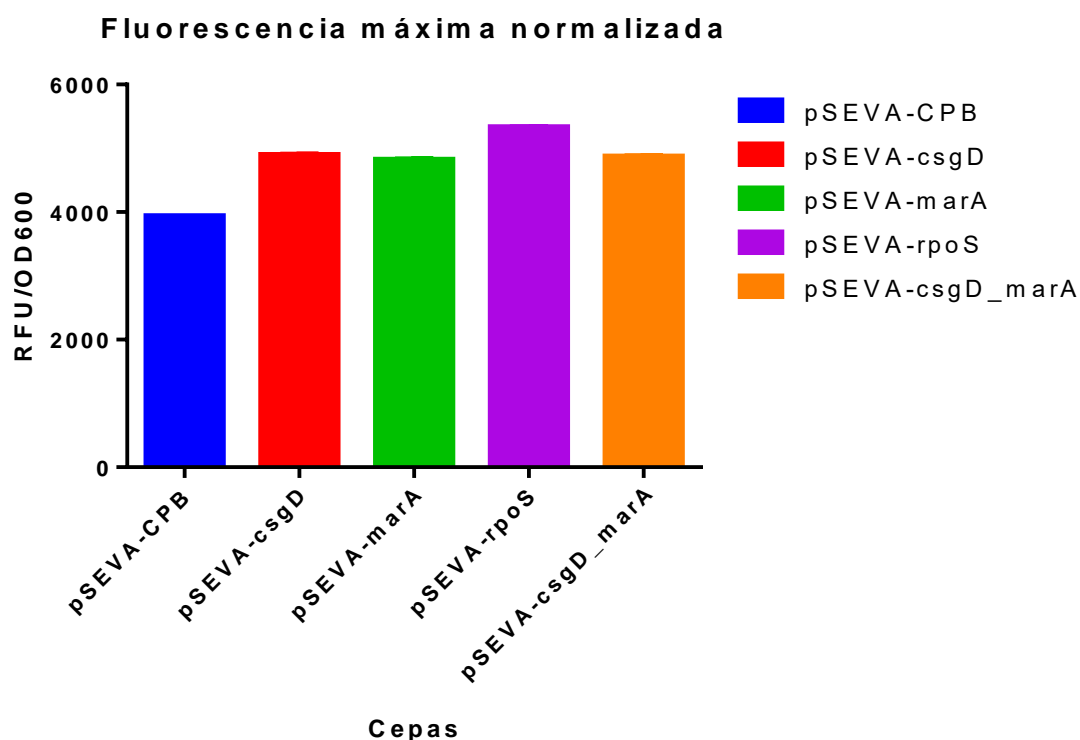
**Figura 4. 9.** Se grafican las líneas de isocosto de la fluorescencia teniendo en cuenta la tasa de producción de proteína, para cada uno de los gARN's en la cepa evaluada, al final de la cinética (24 horas).

La variante pSEVA/cgsD\_marA, al parecer mantiene un comportamiento bastante destacable en cada una de las condiciones analizadas en sentido general en este estudio. Tanto en el *estado estable* como al final de la cinética, sigue siendo la alternativa que al parecer se mantiene como una de las que mayores recursos celulares es capaz de liberar y por ende es capaz de emitir mayores niveles de fluorescencia, esto nos sugiere que el efecto de liberación

de carga proteómica en este caso posee un comportamiento aditivo en comparación al efecto observado de los gARN's involucrados actuando de manera individual.

#### 4.5 Efecto de los gARN's en la fluorescencia máxima, densidad óptica máxima y velocidad de crecimiento máxima.

Evaluamos el efecto del sistema CRISPRi en parámetros como la velocidad de crecimiento, la biomasa y la producción de proteína fluorescente máxima alcanzada en este estudio. Como se puede apreciar en la **figura 4.10**, en la cual se grafica la fluorescencia máxima normalizada alcanzada por las cepas durante la cinética.



**Figura 4. 10.** Se grafica la fluorescencia máxima normalizada alcanzada por las cepas durante el experimento cinético.

Analizando estos resultados podemos comprobar que las cepas transformadas con las variantes de los distintos gARN's mantienen un comportamiento similar entre sí, alcanzando valores de fluorescencia mucho más elevados que los alcanzados por la cepa control pSEVA/CPB. El porcentaje de incremento en la

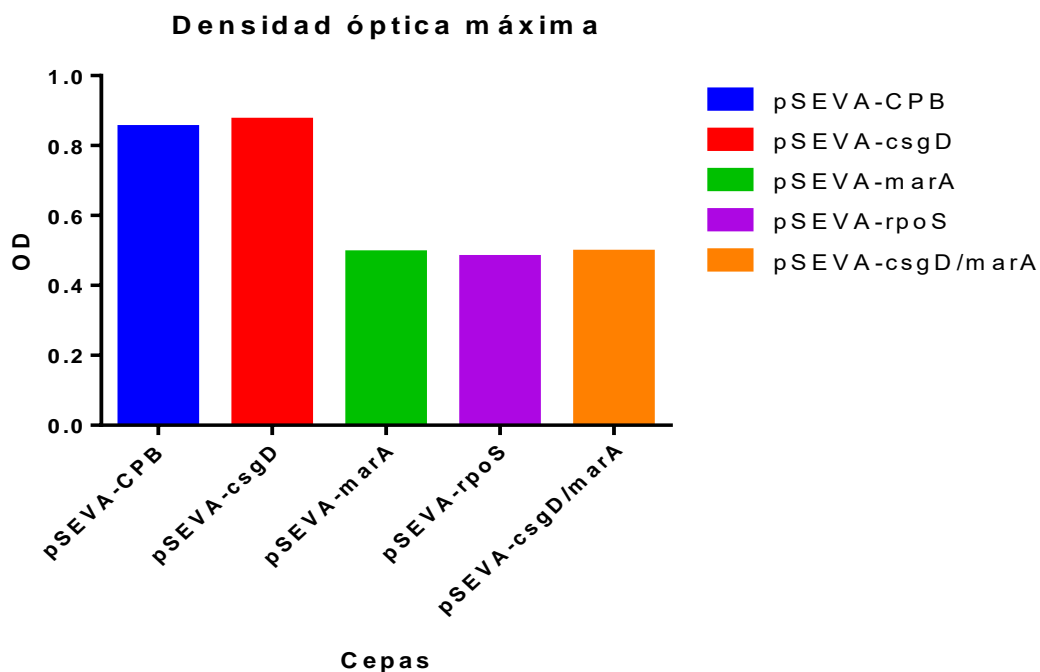


fluorescencia para las cepas pSEVA/csgD, pSEVA/marA, pSEVA/rpoS y pSEVA/csgD-marA con respecto a la cepa control fue del 23%, 22.5%, 35%, 25.4% respectivamente. Esto nos sugiere que de manera general los genes reguladores reprimidos, si generan una liberación en los recursos celulares, confirmándose los resultados que pudimos observar en el análisis efectuado en las líneas de isocosto.

Sin embargo, encontramos algunas particularidades en este análisis, primero esperábamos encontrar la mayor emisión de fluorescencia en la cepa transformada con la combinación doble de gARN's (pSEVA/csgD\_marA). Sustentamos la idea que el estado fisiológico de la cepa es muy importante para determinar los factores de transcripción a reprimir, debido a que las interacciones regulatorias a lo interno de la red de regulación transcripcional pueden cambiar. Otro factor a tener en cuenta es que si bien el sistema CRISPRi, posee muchas virtudes como sistema para interrogar las redes de regulación transcripcional y genes en sentido general, de manera multiplexiable, y sin causar modificaciones a la información genética del organismo, no es un sistema que permita el *knockout* de genes. A diferencia del sistema tradicional CRISPR/Cas9, por lo que las dinámicas de expresión de los genes afectados son diferentes, ya que nuestro sistema puede tener fugas y no estar reprimiendo constantemente los factores de transcripción. Debido a la propia naturaleza del mecanismo de acción del sistema CRISPRi de impedir la unión de la ARN polimerasa a un determinado gen, y de bloquear la transcripción completamente de un determinado gen, provocando la síntesis abortiva del transcrito. Sin embargo este fenómeno puede ocurrir en sentido contrario, dado que la ARN polimerasa puede en determinados momentos despegar la dCas9, y aunque se ha comprobado que ocurre con mucha menos frecuencia, sigue siendo un evento probabilístico, por lo cual no podemos descartar que pueda ocurrir en nuestras cepas.

En la **figura 4.11**, podemos observar la variación de la densidad óptica entre las cepas del experimento. Resulta interesante que la cepa control (pSEVA/CPB) resultó obtener mayores lecturas de densidad óptica, estos resultados nos sugiere que si bien estas cepas poseen la capacidad de crecer más que las cepas transformadas con las diferentes variantes de nuestro vector. Se comprobó en nuestras cepas que su capacidad de expresión de proteína fluorescente es bastante baja en comparación con el comportamiento de las

variantes con factores de transcripción reprimidos obtenidas durante el experimento. Registrándose una reducción en la densidad óptica para las cepas pSEVA/marA, pSEVA/rpoS y pSEVA/csgD-marA de un 42%, 43% y 41% respectivamente. La cepa pSEVA/csgD sin embargo mostró un ligero incremento en la densidad óptica con un 2.4% en relación con la cepa control (pSEVA/CPB). Según el trabajo de Ceroni *et al.*, (2017), la expresión de algunas construcciones sintéticas en *E. coli*, pueden inducir la sobreexpresión de factores de transcripción como el factor sigma 32 relacionado con estrés térmico, el cual se ha observado una sobreexpresión usando la célula censa la sobreexpresión proteica. Esto pudiese sugerir que nuestro vector pudiese a la vez estar reprimiendo los factores de transcripción seleccionados y liberando los recursos celulares. En teoría también pudiesen estar induciendo la expresión de otros factores de transcripción que no necesariamente se tengan en cuenta para su represión y que a su vez estos inducirían la expresión de múltiples genes, cuya expresión a su vez estén “robando” esos recursos celulares.



**Figura 4. 11.** Se grafican los valores de densidad óptica máxima alcanzados en el experimento cinético.

Estudios posteriores pudiesen involucrar el estudio del cambio del transcriptoma, o incluso de aquella fracción del transcriptoma con niveles de

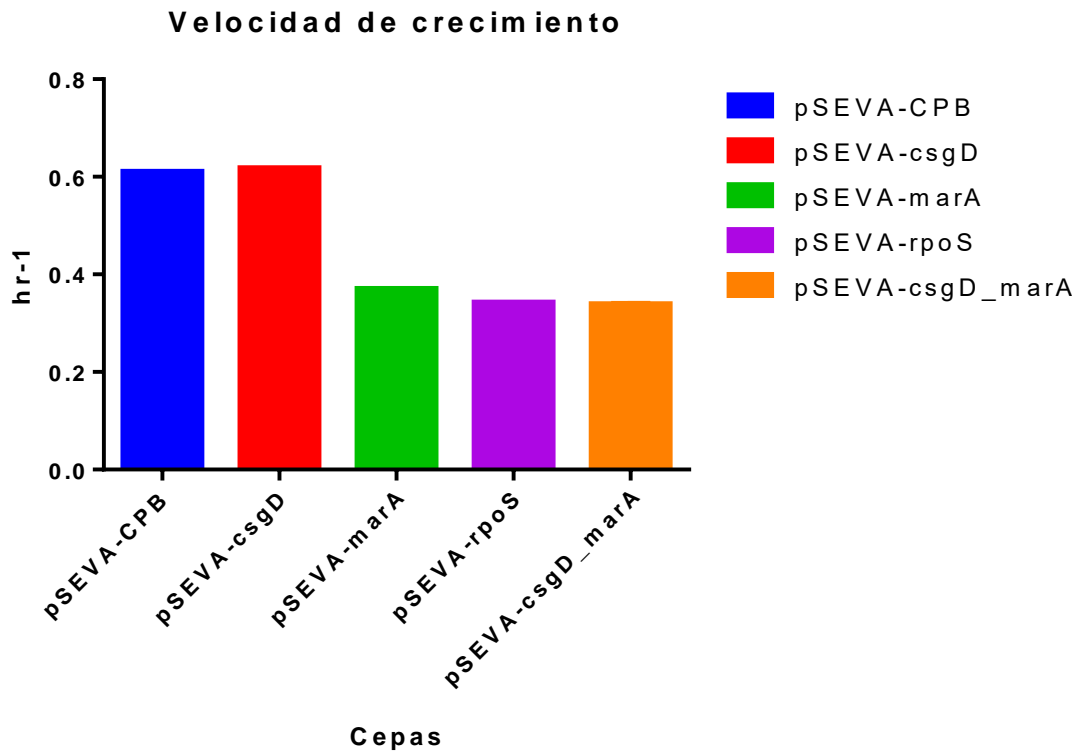
expresión y traducción elevados utilizando técnicas como el *Ribosome profiling* en la cual se puede inferir de manera bastante eficiente los niveles de expresión reales de los genes. Esta técnica pudiera utilizarse en *E.coli* en presencia del vector utilizado en este estudio, para determinar qué respuesta provoca en la red de regulación transcripcional y poder determinar si esta respuesta vale la pena modificarla.

Otra alternativa sería el diseño y modificación de una versión de este vector inducible por ejemplo a anhidrotetraciclina, u otro inductor, que permita regular a voluntad la expresión de la proteína dCas9. De esta manera se logra reducir los efectos tóxicos que pudiese generar en la célula y también la carga metabólica permanente que posee un sistema de expresión constitutivo. Otra ventaja importante de utilizar un sistema de expresión inducible sería potenciar todavía más la característica de la temporalidad de la represión sobre los factores de transcripción, permitiendo un control más sintonizable y por tanto un mayor control sobre las variables analizadas.

En la **figura 4.12** podemos observar el comportamiento de la velocidad de crecimiento de las cepas con diferentes variantes de gARN's. Registrándose una disminución en la velocidad de crecimiento para las cepas pSEVA/marA, pSEVA/rpoS y pSEVA/csgD-marA del 39%, 43% y 44% respectivamente. Una vez más la cepa pSEVA/csgD, logró mostrar un ligero incremento en la velocidad de crecimiento de un 1% en relación a la cepa control pSEVA/CPB.

En esta figura se puede observar un comportamiento similar a lo observado en la gráfica que describe la densidad óptica. La cepa con una mayor velocidad de crecimiento fueron las cepas pSEVA/CPB. Estos resultados parecen sugerir que si bien la célula es capaz de poseer niveles de expresión modestos de proteína heteróloga, esta a su vez prioriza la expresión de proteínas relacionadas con el crecimiento. Sin embargo como regla general las cepas con las variantes de gARN's diseñados para reprimir los factores de transcripción seleccionados alcanzan una menor densidad óptica y muestran una reducción en la velocidad de crecimiento.

Pero al introducir modificaciones a la red de regulación transcripcional celular con la intención de liberar recursos celulares, estos recursos pueden ser utilizados en la expresión de nuevas funciones, pero el costo es someter a la célula a un crecimiento más lento y limitado.



**Figura 4. 12.** Se grafican las velocidades de crecimiento máximas alcanzadas por las cepas en el experimento cinético.

Aunque los niveles de expresión de proteína fluorescente observados a diferentes concentraciones de inductor claramente sugieren que las variantes de gARN's empleadas en este trabajo cumplen con su objetivo y liberan una cantidad de recursos celulares suficientes para superar a la cepa control a pesar de la carga intrínseca del sistema, no deja de ser cierto que este sistema provee a la célula con un costo de crecimiento a tener en cuenta. Varios autores ya han proveído pistas sobre algunas de las causas que pueden influir en este aspecto (Gyorgy *et al.*, 2015), dentro de los factores más importantes que pueden influir están el número de copias del plásmido, la fortaleza del promotor empleado, y sobre todo la fortaleza del sitio de unión del ribosoma (RBS). Otros factores que pueden influir en este comportamiento son la disponibilidad de otros compuestos como el factor de elongación Tu (EF-Tu). Este factor forma un complejo con los tARN's encargados de cargar los diferentes aminoácidos hacia los ribosomas, formando un complejo terciario en conjunto con el GTP (EF-Tu/aminoácido/GTP) y otros complejos macromoleculares como las 21 diferentes tARN's sintetasas (Klumpp *et al.*, 2013), pudieran estar influyendo en el comportamiento

observado. Siguiendo esta línea de pensamiento podemos sugerir que factores como algunos de los mencionados anteriormente, están presentes en el plásmido pSEVA63-Dual, este vector posee sitios de unión del ribosoma muy fuertes, además posee un origen de replicación con un relativo bajo número de copias (alrededor de 18 copias por célula) y posee la presencia de promotores fuertes. Resultando válido suponer, basándonos en la literatura y en nuestros resultados que este sistema es capaz de “secuestrar” gran parte de la maquinaria transcripcional y traduccional de la célula, redirigiéndola a la sobreexpresión de proteína fluorescente, pero con un costo, la limitación de esta misma maquinaria para favorecer la expresión de genes relacionados con el crecimiento como pueden ser transportadores y porinas, entre otros factores.

## V. Conclusiones

- Se logró implementar un sistema de represión genético usando CRISPRi.
- La represión de genes reguladores mediante el uso del sistema CRISPRi es efectiva, lográndose una producción de proteínas fluorescentes mayor que la cepa control.
- Las variantes pSEVA/csgD, pSEVA/marA y pSEVA/rpoS produjeron considerablemente más proteína fluorescente que la cepa control.
- La combinación doble pSEVA/csgD\_marA, fue la de mejor comportamiento en sentido general a lo largo de este estudio, tanto en el *estado estable* como en el final de la cinética, demostrándose que la liberación de recursos celulares puede ser aditiva hasta cierto punto.
- La represión del factor sigma *rpoS* produjo una mayor fluorescencia que la combinación doble de gARN's en el final de la cinética. Lo que soporta la idea de que el momento fisiológico de la célula en el que están más activos estos genes reguladores juega un papel importante en su capacidad de liberación de recursos celulares.
- La disponibilidad de los recursos celulares puede ser modificada, mediante la optimización del proteoma, generando un incremento en el rendimiento de construcciones sintéticas empleadas en este estudio.

## VI. Perspectivas

- En estudios posteriores pudiera emplearse una variante inducible de este sistema CRISPRi que permita controlar de manera dinámica la represión de los genes reguladores; de esta manera se reduciría el efecto tóxico de la dCas9 y permitiría el estudio durante diferentes fases del estadio fisiológico.
- Una alternativa potencialmente interesante constituye incluir genes estructurales dentro de la lista de genes candidatos a reprimir, en aras de incrementar la liberación de recursos celulares.
- El empleo de técnicas más cuantitativas como el *Ribosome Profiling*, permitiría determinar de manera más eficiente los niveles de transcripción y traducción reales de los genes, esta información pudiera ser utilizada para refinar el proceso de selección de posibles genes a reprimir, priorizando aquellos que posean mayores niveles de expresión y no sean genes esenciales en determinada condición, haciendo la liberación de recursos celulares significativamente mayor.
- Esta metodología puede ser extrapolable a otros microorganismos de interés biotecnológico, permitiendo explorar las potencialidades de estos, para su implementación como organismos *chassis* emergentes de aplicación en la biotecnología.

## VII. Bibliografía

- Borkowski Olivier, Ceroni Francesca, Stan Guy-Bart, and Ellis Tom (2016). Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology*, 33: 123-130.
- Cardinale Stefano and Arkin Adam Paul (2012). Contextualizing context for synthetic biology –identifying causes of failure of synthetic biological systems. *Biotechnology JouARNI* 7 856-866.
- Vigouroux Antoine, Oldewurtel Enno, Cui Lun , Teeffelen Sven van and Bikard David (2017). Engineered CRISPR-Cas9 system enables noiseless, fine-tuned and multiplexed repression of bacterial genes.
- Hawkins John S, Wong Spencer, Peters Jason M, Almeida Ricardo, and Qi Lei S. (2015). Targeted Transcriptional Repression in Bacteria Using CRISPR Interference (CRISPRi). *Methods Molecular Biology*, 1311: 349-362
- Ceroni Francesca, Algar Rhys, Stan Guy-Bart & Ellis Tom (2015). Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature Methods* 12 415-418.
- O'Brien Edward J., Utrilla Jose, Palsson Bernhard O. (2016) Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments. *PLOS Computational Biology* 12(6)
- Qi Lei S., Larson Matthew H., Gilbert Luke A., DouADN Jennifer A., Weissman Jonathan S., Arkin Adam P., and Lim Wendell A. (2013). Repurposing CRISPR as an ARN-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* 152: 1173-1183.
- Utrilla Jose, O'Brien Edward J., Chen Ke, Armenta-Medina Dagoberto, Feist Adam M., Palsson Bernhard O. (2016). Global Rebalancing of Cellular Resources by Pleiotropic Point Mutations Illustrates a Multi-scale Mechanism of Adaptive Evolution. *Cells Systems* 2: 260-271.
- Glasscock Cameron J., Lucks Julius B., and DeLisa Matthew P. (2016). Engineered Protein Machines: Emergent Tools for Synthetic Biology. *Cell Chemical Biology* 23.
- Hutchison III Clyde A., Chuang Ray-Yuan, Noskov Vladimir N., Assad-Garcia Nancy, Deerinck Thomas J., Ellisman Mark H., Gill John, Kannan



- Krishna, Karas Bogumil J., Ma Li, Pelletier James F., Qi Zhi-Qing, Richter R. Alexander, Strychalski Elizabeth A., Suzuki Sun Lijie Yo, Tsvetanova Billyana, Wise KimS., Smith Hamilton O., Glass John I., Merryman Chuck, Gibson Daniel G., Venter J. Craig (2016). Design and synthesis of a minimal bacterial genome. *Science*, 351.
- Cress Brady F., Toparlak . . Duhan, Guleria Sanjay, Lebovich Matthew, Stieglitz Jessica T., Englaender Jacob A., Jones J. Andrew, Linhardt Robert J., and Koffas Mattheos A. G. (2015). CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in *E. coli*. *ACS Synthetic Biology* 4, 987-1000.
  - Yuan Lu (2017). Cell-free synthetic biology: Engineering in an open world. *Synthetic and systems biotechnology*, 2: 23-27.
  - DouADN Jennifer A. and Charpentier Emmanuelle(2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, vol 346.
  - Shalem Ophir, Sanjana Neville E, Hartenian Ella, Shi Xi, Scott David A., Mikkelsen Tarjei, Heckl Dirk, Ebert Benjamin L., Root David E., Doench John G., Zhang Feng (2013). Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science Express*
  - Puchta Holger( 2016). Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Current opinion in Plant Biology* 36, 1-8.
  - Kim Seong Keun, Seong Wonjae, Han Gui Hwan, Lee Dae-Hee and Lee Seung-Goo (2017). CRISPR interference-guided multiplex repression of endogenous competing pathway genes for redirecting metabolic flux in *Escherichia coli*. *Microbial Cell Factories*, 16:188.
  - Wu Junjun, Du Guocheng, Chen Jian, Zhou Jingwen (2015). Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*. *Science Reports*.
  - Jones Daniel Lawson, Leroy Prune, Unoson Cecilia, Fange David, Ćurić Vladimir, Lawson Michael J., Elf Johan (2017). Kinetics of dCas9 target search in *Escherichia coli*. *Science* 357: 1420-1424.

- Acar Murat, Mettetal Jerome, van Oudenaarden Alexander (2008). Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics* 40(4)
- Liu Q, Schumacher J, Wan X, Lou C, Wang B.( 2018) Orthogonality and Burdens of Heterologous and Gate Gene Circuits in E. coli. *ACS Synth Biol.*;7(2):553-564.
- Ceroni F, Furini S, Te G, et al.( 2017) Burden-driven feedback control of gene expression.1-17.
- Chen P, Vecchio D Del. A (2018). Model for Resource Competition in CRISPR-Mediated Gene Repression.
- Randall LP, Woodward MJ.(2002) The multiple antibiotic resistance (mar) locus and its significance. *Res Vet Sci.* 72(2):87-93.
- Maira-Litran T, Allison DG, Gilbert P. (2000)Expression of the multiple antibiotic resistance operon ( mar ) during growth of Escherichia coli as a biofilm. *JApplMicrobiol.*88(2):243-247.
- Griffith KL, Shah IM, Wolf RE. (2004)Proteolytic degradation of Escherichia coli transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol Microbiol.* 51(6):1801-1816.
- Hirsch M, Elliott T. (2002). Role of ppGpp in rpoS Stationary-Phase Regulation in Escherichia coli Role of ppGpp in rpoS Stationary-Phase Regulation in Escherichia coli. *J Bacteriol.*;184(18):5077-5087.
- Perni Stefano, Preedy Emily, Landini Paolo and Prokovich Polina.(2016) Influence of csgD and ompR in Nanomechanics, Adhesion Forces and Curli Properties of *E.coli*. *ACS Synth Biol* 32 7965-7974
- Ogasawara Hiroshi, Yamamoto Kaneyoshi, Ishihama Akira.(2011) Role of the Biofilm Master Regulator CsgD in Cross-Regulation between Biofilm Formation and Flagellar Synthesis. *JouARNI of Bacteriology* 193(10) 2587-2597
- Grantcharova Nina, Peters Verena, Monteiro Claudia, Zakikhany Katherina, Romling Ute. (2010) Biestable Expression of csgD in Biofilm Development of *Salmonella enterica* Serovar Typhimurin. *JouARNI of Bacteriology* 192(2) 456-466

- Liu Zhen, Niu Hua, Wu Shuvan and Huang Rui.(2014) CsgD regulatory network in bacterial trait-altering biofilm formation. Nature Emerging Microbes and Infections 3
- Gyorgy Andras, Jimenez Jose, Yazbek Jhon, Huang Hisin-Ho, Chung Hattie, Weiss Ron, Del Vecchio Domitilla. (2015) Isocost lines Describe the celular Economy of Genetics Circuits. Biophysical JouARNI 109 639-646