



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

“Caracterización de los efectos de mutaciones originadas durante un proceso de evolución adaptativa por crecimiento rápido en glucosa en cepas de *Escherichia coli* carentes del sistema PTS”

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M. en C. César A. Aguilar Martínez

TUTOR PRINCIPAL:

[Dr. Francisco G. Bolívar Zapata](#)

[Instituto de Biotecnología - UNAM](#)

MIEMBROS DEL COMITÉ TUTOR

[Dr. Sergio Manuel Encarnación Guevara](#)

[Maestría y Doctorado en Ciencias Bioquímicas](#)

[Dr. Ricardo Grande Cano](#)

[Maestría y Doctorado en Ciencias Bioquímicas](#)

Cuernavaca, Morelos. Enero, 2015



Universidad Nacional
Autónoma de México



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.



El presente trabajo se realizó en el departamento de Ingeniería Celular y Biocatálisis del Instituto de Biotecnología de la Universidad Nacional Autónoma de México, bajo la dirección del Dr. Francisco Gonzálo Bolívar Zapata.

César Augusto Aguilar Martínez fue financiado por el CONACyT a través de su programa de becas de posgrado, así como por el proyecto 740 Ing. Cel. Biocatálisis. Agradezco al Programa de Apoyo para Estudios del Posgrado (PAEP) por el apoyo económico otorgado para congresos.



Agradezco el apoyo técnico de:

Dra. Noemí Flores, M. C. Ramón de Anda, M. C. Georgina Hernández, C. D. Mercedes Enzaldo, Dr. Ricardo Grande, M. C. Francisco Riveros Mc-Kay, Dra. Diana Sahonero, Dra. Isabel López y la Sra. Aurelia Ocampo.

A mi madre, porque hizo de mi lo que soy.

A mi padre por estar siempre conmigo.

A Hugo, porque me ha enseñado tanto.

A Tere, por todo su amor.

A Andrea por serlo todo.

A mi nuevo amor.

**Sin prisa, pero sin descanso.
Johann Wolfgang Goethe (1749-1832)**

Agradecimientos:

- Al Dr. Francisco Bolívar Zapata, por darme la oportunidad de ser parte de su grupo de trabajo; por todo su apoyo y la infinita confianza que depositó en mí desde el primer momento; pero sobre todo, por los aquellos momentos de arduo trabajo, discusión y reflexión que enriquecieron mi formación.
- A los Doctores Sergio Encarnación y Ricardo Grande, que formaron parte del comité tutorial y contribuyeron con sus comentarios y observaciones a mejorar la calidad y los alcances del proyecto.
- Al Dr. Adelfo Escalante, por el apoyo que me brindó en todo momento, por su ayuda, sus consejos y sobre todo por su amistad.
- A Mercedes Enzaldo y Aurelia González, porque creo que su trabajo en el laboratorio es indispensable. Todo el apoyo que recibí de ellas resulta invaluable.
- A los técnicos del laboratorio: la Dra. Noemí Flores, el M.C. Ramón de Anda, la M.C. Georgina Hernández y la Q.I. Luz María Martínez; su apoyo y disposición me ayudaron a llevar a buen puerto este trabajo.
- A los Doctores Guillermo Gosset y Alfredo Martínez, porque sus observaciones y consejos durante los seminarios me ayudaron a incrementar la calidad de este proyecto.
- A Sonia Caro y a Delia Caro, ya que sus labores administrativas aceleraron los tiempos de espera en cuestiones académicas y experimentales.
- A la Dra. Andrea Sabido, porque crecimos juntos en lo académico y lo sentimental. Sus observaciones siempre fueron las más atinadas porque conocía las dos partes más importantes de mi proyecto: los objetivos y a mí.
- A todos mis compañeros de laboratorio, los que estuvieron y los que están, ya que son personas de un alto nivel académico, que te hacen mejorar en todos aspectos. Afortunadamente todos ellos me brindaron su ayuda y amistad. De entre ellos quisiera agradecer especialmente a los siguientes: Dr. Juan Carlos Sigala, Dr. Eugenio Meza, Dr. Iván Muñoz, Dr. Marco T. Fernandez, M.C. Larisa Cortés, M.C. Hannali Cuevas, M.C. Grecia Fuentes y al IBt Arturo Montes de Oca; gracias por su apoyo y su amistad.
- A la Universidad Nacional Autónoma de México.

MIEMBROS DEL JURADO

Presidente: Dr. José Luis Puente (IBT, UNAM)

Secretario: Dr. Daniel Segura González (IBT, UNAM)

Vocal: Dr. Miguel Angel Cevallos Gaos (CCG, UNAM)

Vocal: Dr. Otto Geiger (CCG, UNAM)

Vocal: Dra. Katy Juárez (IBT, UNAM)

Indice

I.	Resumen/Abstract	Pag. 15
II.	Introducción	Pag. 18
III.	Antecedentes	Pag. 19
	A. Ingeniería de vías metabólicas.....	Pag. 20
	B. Transcripción.....	Pag. 20
	C. Evolución.....	Pag. 21
	1. Evolución Adaptativa: La selección natural.....	Pag. 22
	2. Evolución Neutra: La deriva génica.....	Pag. 22
	D. Evolución bacteriana.....	Pag. 23
	1. Mutaciones inducidas por estrés.....	Pag. 23
	2. Respuesta SOS.....	Pag. 23
	3. Respuesta de estrés general.....	Pag. 24
	4. Respuesta estricta (Stringent response).....	Pag. 25
	E. Experimentos de Evolución Adaptativa.....	Pag. 26
	F. Secuenciación de DNA.....	Pag. 28
	G. Sistema de transporte de azúcares dependiente de PEP (PTS).....	Pag. 29
	1. Componentes del sistema PTS.....	Pag. 30
	2. Represión catabólica.....	Pag. 30
	H. Nodo PEP-PYR en el metabolismo central.....	Pag. 32
	I. Evolución adaptativa como herramienta en la IVM para la producción de compuestos aromáticos...Pag. 33	
	1. Características de las cepas PTS ⁻	Pag. 35
	a) Cepa PB11.....	Pag. 35
	b) Cepa PB12.....	Pag. 36
	c) Cepa PB13.....	Pag. 39
IV.	Propósitos y alcances de este trabajo	Pag. 39
V.	Objetivo general	Pag. 40
	A. Objetivos específicos.....	Pag. 40
VI.	Resultados y discusión	Pag. 40
	A. Cambios genéticos clave en la recuperación del crecimiento en glucosa en las cepas PTS ⁻Pag. 40	
	1. Determinación del genoma de la cepa JM101 y PB11.....	Pag. 40
	2. Determinación de las mutaciones en la cepa PB12 y PB13.....	Pag. 43
	3. Fragmento cromosomal ausente.....	Pag. 44
	4. Análisis bioinformático de la región cromosomal deletada.....	Pag. 45
	5. Pérdida del fragmento cromosomal en otras cepas aisladas.....	Pag. 45
	6. Divergencia evolutiva entre las cepas PB12 y PB13.....	Pag. 46
	7. Genes ausentes en la cepa PB12 y PB1.....	Pag. 47

a)	<i>rppH</i>	Pag. 48
b)	Operón <i>aas-lpIT</i>	Pag. 48
c)	<i>galR</i>	Pag. 49
B.	Mecanismos adaptativos en cepas PTS ⁻ creciendo en glucosa como única fuente de carbono.....	Pag. 49
1.	Contribución de los genes ausentes en la recuperación del crecimiento.....	Pag. 49
2.	Caracterización de cepas $\Delta rppH$ en medio mínimo M9 con glucosa como única fuente de carbono.....	Pag. 52
3.	Crecimiento de las cepas $\Delta rppH$ en acetato y en la mezcla glucosa-acetato.....	Pag. 53
4.	Determinación de los niveles de transcrito de genes del metabolismo central en cepas $\Delta rppH$	Pag. 56
a)	Glicólisis.....	Pag. 58
b)	TCA y cadena respiratoria.....	Pag. 59
c)	Gluconeogénesis, ciclo del glioxilato y <i>poxB</i>	Pag. 60
d)	Regulón de galactosa.....	Pag. 60
e)	Metabolismo del tetrafosfato de guanosina (ppGpp).....	Pag. 61
f)	Genes ausentes en el cromosoma de las cepas PB12 y PB13.....	Pag. 62
g)	Genes reguladores.....	Pag. 63
5.	Las cepas PB12 y PB13 carecen del ciclo de la fosfatidiletanolamina (PE).....	Pag. 64
6.	Mutaciones puntuales en los genes <i>dhaM</i> y <i>glpT</i>	Pag. 70
7.	Las cepas evolucionadas carecen del gen <i>mutH</i>	Pag. 73
a)	Implicaciones de una tasa de mutación elevada.....	Pag. 75
8.	Otras mutaciones relevantes en la cepa PB12.....	Pag. 76
a)	Mutaciones puntuales no sinónimas en genes implicados en funciones de regulación.....	Pag. 76
(1)	Gen <i>arcB</i>	Pag. 77
(2)	Gen <i>barA</i>	Pag. 78
(3)	Gen <i>rpoD</i>	Pag. 79
(4)	Gen <i>rna</i>	Pag. 80
b)	Mutaciones puntuales no sinónimas en genes implicados en funciones metabólicas.....	Pag. 80
c)	Mutaciones no sinónimas en genes con función desconocida.....	Pag. 81
(1)	Los genes <i>yjU</i> y <i>rssA</i> poseen un dominio "patatin like".....	Pag. 81
(2)	<i>ypdA</i>	Pag. 82
d)	Inactivaciones de genes relevantes de función desconocida que sufrieron mutaciones en la cepa PB12.....	Pag. 85
C.	Diseño de nuevas estrategias de producción en cepas PTS ⁻	Pag. 86
1.	¿Por qué la cepa PB12 es tan buena productora de compuestos aromáticos?.....	Pag. 86
2.	Nuevas estrategias de producción utilizando cepas de E. coli PTS ⁻	Pag. 88

VII.	Conclusiones	Pag. 89
VIII.	Perspectivas	Pag. 91
IX.	Metodología Experimental	Pag. 92
	A. Cepas bacterianas y plásmidos.....	Pag. 92
	B. Medios de cultivo y condiciones para todas las determinaciones del crecimiento y parámetros cinéticos.....	Pag. 93
	C. Componentes del medio mínimo M9.....	Pag. 93
	D. Caracterización de todas las cepas empleadas en este estudio.....	Pag. 93
	E. Extracción de DNA de las cepas JM101, PB11, PB12 y PB13 para la secuenciación de su genoma.....	Pag. 94
	F. Secuenciación comparativa (Roche NimbleGen Inc.).....	Pag. 94
	G. Secuenciación Masiva.....	Pag. 94
	1. Illumina Gallx.....	Pag. 94
	2. Ion Torrent PGM.....	Pag. 95
	H. Ensamble del genoma e identificación de variantes por Winter Genomics (WG).....	Pag. 95
	1. Datos obtenidos con el equipo de Illumina GAllx.....	Pag. 95
	2. Datos obtenidos con el equipo Ion Torrent PGM.....	Pag. 96
	3. Confirmación de las mutaciones puntuales.....	Pag. 96
	I. Resecuenciación de la región cromosomal perdida.....	Pag. 96
	J. Generación de todas las mutantes en las distintas cepas.....	Pag. 96
	K. Extracción de RNA y síntesis de cDNA en las cepas JM101, JM101 Δ rppH, PB11, PB11 Δ rppH, PB12 y PB13.....	Pag. 97
	L. Determinación de las tasas de mutación en las diferentes cepas de <i>E. coli</i>	Pag. 98
	M. Marcaje “ <i>in vivo</i> ” de las cepas JM101, PB11, PB12, PB13 y sus derivadas carentes del gen <i>fadD</i> con [¹⁴ C] acetato y análisis de composición lipídica de membrana.....	Pag. 98
X.	Material suplementario	Pag. 100
	A. Tablas.....	Pag. 100
	B. Figuras.....	Pag. 120
XI.	Referencias	Pag.122
XII.	Anexos	Pag. 136

Índice de Tablas

Tabla 1.- Parámetros cinéticos de las cepas incluidas en éste estudio evaluados en medio mínimo M9 con 2 g/L de glucosa.

Tabla 2.- Velocidades específicas de crecimiento de diferentes cepas generadas en este estudio utilizando medio mínimo M9 y diferentes sustratos.

Tabla 3.- Velocidades específicas de crecimiento en derivadas de las cepas JM101, PB11, PB12 y PB13 carentes de los genes *glpQ*, *glpT* y *fadD*, utilizando medio mínimo M9-Glc 2g/L.

Tabla 4.- Velocidades específicas de crecimiento en derivadas de las cepas JM101, PB11, PB12 y PB13 carentes del gen *dhaM* utilizando medio mínimo M9-Glc 2 g/L.

Tabla 5.- Velocidades específicas de crecimiento de diferentes derivadas de las cepas JM101, PB11 y PB12 carentes de los genes *yjiU*, *rssA* y *ypdA* utilizando medio mínimo M9-Glc 2 g/L.

Tabla A.- Todas las cepas empleadas en este estudio.

Tabla S1.- Diferencias en la cepa JM101 con respecto a la cepa MG1655 (Mutaciones puntuales no sinónimas en la cepa JM101).

Tabla S2.- Diferencias en la cepa JM101 con respecto a la cepa MG1655 (Mutaciones puntuales sinónimas en la cepa JM101).

Tabla S3.- Diferencias en la cepa JM101 con respecto a la cepa MG1655 (Mutaciones en regiones no codificantes en la cepa JM101).

Tabla S4.- Diferencias entre la cepa PB12 y JM101. A) Se muestran las 23 mutaciones puntuales no sinónimas en genes estructurales presentes en la cepa PB12 al compararla con la cepa silvestre JM101. Los asteriscos indican las 21 mutaciones detectadas por ambos métodos/compañías, Roche NimbleGen Inc. (RN) y Winter Genomics (WG). Adicionalmente, 16 mutaciones sinónimas en diferentes genes también fueron detectadas por ambos métodos (Tabla S3). Mutaciones en 14 genes fueron corroboradas mediante secuenciación por Sanger, los cuales se indican con el símbolo +. B)

Se muestran los genes ausentes en la cepa PB12, producto de la delección cromosomal ocurrida durante el proceso de evolución adaptativa.

Tabla S5.- Diferencias entre la cepa JM101 y PB12. Se muestran las 16 mutaciones puntuales sinónimas en la cepa PB12 al compararla con la cepa silvestre JM101.

Tabla S6.- Diferencias entre la cepa JM101 y PB12. Se muestran las mutaciones en regiones no codificantes en la cepa PB12 al compararla con la cepa silvestre JM101.

Tabla S7.- Diferencias entre la cepa JM101 y PB13. Se muestran las mutaciones puntuales no sinónimas en la cepa PB13 al compararla con la cepa silvestre JM101.

Tabla S8.- Diferencias entre la cepa JM101 y PB13. Se muestran las mutaciones puntuales sinónimas en la cepa PB13 al compararla con la cepa silvestre JM101.

Tabla S9.- Resultados del análisis estadístico para determinar las tasas y frecuencias de mutación en las cepas JM101, PB11, PB11 Δ Reg, PB12 y PB13.

Tabla S10.- Valores relativos de transcrito de las cepas PB11, PB12, JM101 Δ rppH y PB11 Δ rppH comparados con la cepa JM101. Para cada gen en todas las cepas, el nivel de transcripción del gen de la cepa silvestre JM101 se empleó como control para normalizar los datos, teniendo JM101 siempre un valor de RT-qPCR de uno. De esta forma los datos de éste estudio se reportan como valores de expresión relativa, comparados al nivel de expresión de JM101. Los resultados presentados son el promedio de al menos tres mediciones independientes de los valores de expresión para cada gen. Los valores se obtuvieron a partir de diferentes cDNA's generados al menos de dos muestras de biorreactores independientes. Los valores de expresión por RT-qPCR obtenidos para cada gen difieren menos del 30%. La desviación estándar para cada gen se muestra como: \pm (valor). * Indica valores previamente reportados [8,108,117].

Tabla S11.- Se muestran las secuencias nucleotídicas de los oligos empleados en este estudio con diversos fines.

Índice de Figuras

Figura 1.- Vía de degradación de mRNA's dependiente del extremo 5', en donde participa la pirofosfohidrolasa RppH y la RNasa E.

Figura 2.- Mecanismos moleculares asociados a la mutagénesis inducida por estrés.

Figura 3.- Principales sistemas de transporte en *E. coli*.

Figura 4.- Mecanismos moleculares de regulación durante la represión catabólica.

Figura 5.- Principales vías de utilización de PEP en el metabolismo central.

Figura 6.- Selección de cepas evolucionadas en un cultivo continuo a partir de la cepa PB11.

Figura 7.- Esquema del metabolismo en las cepas PB11 y PB12, el cual se desprende de datos obtenidos por RMN y RT-qPCR para determinar flujos de carbono y niveles de expresión de genes del metabolismo central respectivamente.

Figura 8.- Esquema de los cambios genéticos en el cromosoma de las cepas PB12 y PB13, donde se muestra la posición estimada de los genes donde se originaron las mutaciones puntuales no sinónimas así como de la delección del fragmento cromosomal. En negritas se muestran las mutaciones compartidas en ambos genomas.

Figura 9.- Mapa comparativo de las cepas JM101 y PB12 realizado por la empresa Nimblegen. Inc. donde se muestran dos delecciones en la cepa PB12; una correspondiente al operón PTS, y otra a una región del cromosoma ubicada en el minuto 64. Ambos fragmentos también se encontraron ausentes en el cromosoma de la cepa PB13.

Figura 10.- Comprobaciones mediante PCR de la región del cromosoma faltante en la cepa PB12.

Figura 11.- Comprobaciones mediante PCR de la región del cromosoma faltante en varias cepas evolucionadas aisladas.

Figura 12.- Esquema de la divergencia evolutiva entre las cepas PB12 y PB13.

Figura 13.- Eliminación de un fragmento cromosomal de 10,328 pb alrededor del minuto 64 (3.0 millones de pb) en la cepa PB12 y PB13 durante el proceso de evolución adaptativa. En el nuevo arreglo se observa la fusión de los genes *ptsP* y *galR* por su parte media; ambos genes carecen de la primera mitad, incluidas sus secuencias promotoras.

Figura 14.- Ciclo de recambio de la fosfatidiletanolamina, donde están involucradas las proteínas LpIT y Aas.

Figura 15.- A. Velocidades específicas de crecimiento en cepas derivadas de la JM101 y PB11 generadas en éste estudio. **B.** Gráfica comparativa de la recuperación del crecimiento en las diferentes cepas derivadas de la PB11. **C.** Esquema de los genes eliminados en las cepas evolucionadas después del proceso de evolución adaptativa.

Figura 16.- Esquema del metabolismo central con metabolitos clave, genes y proteínas involucrados en su transformación.

Figura 17.- Ciclo de recambio y de degradación de la fosfatidil etanolamina (PE).

Figura 18.- Operón *dhaKLM*, donde se señala la posición de la mutación puntual presente en el gen *dhaM* y el mecanismo de fosforilación de la dihidroxiacetona por conducto de la dihidroxiacetona cinasa (DhaK) a partir de PEP.

Figura 19.- Esquema de la estructura primaria de la proteína ArcB, así como las diferentes configuraciones que adopta el homodímero y que determinan su nivel de actividad. Se señala también la posición donde se podría estar generando un tercer puente disulfuro, producto de la mutación puntual.

Figura 20.- A. Esquema de regulación entre los sistemas BarA/UvrY y Csr. Con un asterisco en color rojo se muestra la posición de la mutación en la cepa PB12.

Figura 21.- Modelos de la estructura secundaria y terciaria de la proteína Rnal. En ambos esquemas se identifican los dominios estructurales de la proteína, así como la mutación puntual originada en la posición 90 de la proteína.

Figura 22.- Modelo de la estructura terciaria de las proteínas YjjU y RssA predicho mediante Phyre2, el cual se basó en la identidad que poseen con la proteína VipD de *Legionella pneumophila*. Los diferentes colores indican ocho diferentes niveles de confianza que denotan la seguridad del plegamiento predicho a lo largo de la proteína.

Figura 23.- Modelo de la estructura terciaria de las proteínas YjjU y RssA predicho mediante Phyre2, el cual se basó en la identidad que poseen con la proteína VipD de *Legionella pneumophila*. Los diferentes colores en la estructura indican el grado de sensibilidad a una mutación a lo largo de la proteína. Dentro de las estructuras se señala la posición y la naturaleza de la mutación encontrada en estas proteínas en la cepa PB12. En las gráficas, tomando en cuenta los 20 aminoácidos y siguiendo el mismo patrón de colores, se señala la preferencia del residuo señalado en cada proteína, así como el efecto probable de una mutación.

Figura S1: Secuencia nucleotídica del rearreglo cromosomal después del proceso de evolución adaptativa en las cepas PB12 y PB13.

Figura S2: Curvas de amplificación para el gen *ihfB* en las diferentes cepas de este estudio.

Índice de Gráficas

Gráfica 1.- Mutaciones puntuales encontradas en la cepa JM101 al compararla con la cepa silvestre MG1655. N/A: Regiones no codificantes.

Gráfica 2: Cinéticas de crecimiento en medio mínimo M9-Glc 2g/L de diferentes cepas de éste estudio

Gráfica 3: Consumo de glucosa en diferentes cepas de este estudio utilizando medio mínimo M9-Glc 2 g/L .

Gráfica 4.- Niveles de expresión relativos de genes del regulón *gal*.

Gráfica 5.- Niveles de expresión relativos de genes del metabolismo del ppGpp.

Gráfica 6.- Niveles de expresión relativos de genes en el fragmento cromosomal ausente.

Gráfica 7.- Niveles de expresión relativos de genes reguladores

Gráfica 8.- Distribución de fosfolípidos de membrana en las cepas de este estudio a DO de 0.6 y 1.2. PE, Fosfatidiletanolamina, PG, Fosfatidilglicerol, CL, Cardiolipina.

Gráfica 9.- Acumulación de ácidos grasos libres en las diferentes cepas de este estudio y sus derivadas carentes del gen *fadD*, después de 24, 48 y 72 horas de cultivo. Los porcentajes se calculan a partir del total de especies lipídicas presentes en membrana y periplasma.

Gráfica 10.- Tasas y frecuencias de mutación de las diferentes cepas empleadas en este estudio estimadas con el test de fluctuación de Luria-Delbruck, el cual se basó en la capacidad de generar resistencia a estreptomycin debido a una mutación puntual en el gen *rpsL*. Los análisis para las tasas de mutación se realizaron empleando los métodos de Lea-Coulson (LC) y Ma-Sandri-Sarkar (MSS-MLE).

I. Resumen / Abstract

Resumen: El trabajo de este grupo de investigación se ha centrado en la construcción y caracterización de cepas de *Escherichia coli* carentes del sistema de transporte de carbohidratos dependiente de fosfoenolpiruvato (sistema PTS), el cual es el principal sistema de transporte y fosforilación de glucosa en esta bacteria. Cepas carentes del sistema PTS (cepa PB11), logran incrementar al doble el fosfoenolpiruvato generado durante la glucólisis, el cual puede canalizarse mediante ingeniería de vías metabólicas (IVM) hacia la síntesis de compuestos aromáticos; no obstante, su crecimiento en glucosa se ve seriamente afectado. Como una estrategia de IVM, se llevó a cabo un proceso de evolución adaptativa por crecimiento rápido en glucosa a partir de la cepa PB11 utilizando medio mínimo con glucosa como única fuente de carbono. Dos de las cepas resultantes, aisladas a las 120 y 200 hrs, incrementaron su velocidad específica de crecimiento un 400% y 500% respectivamente, comparadas con la cepa PB11. La secuenciación del genoma de estas cepas evolucionadas denominadas PB12 y PB13 permitió establecer su historia evolutiva y las adaptaciones metabólicas adquiridas para superar la ausencia del principal sistema de transporte de glucosa.

Uno de los principales eventos durante el proceso de evolución ocurrió alrededor de la doceava generación. Este evento comprendió la pérdida de un fragmento cromosomal de 10,328 pb, donde se localizaban, entre otros, los genes *rppH*, *galR*, *aas*, *lpIT* y *mutH*. La ausencia de este fragmento cromosomal en las cepas PB12 y PB13 incrementó la tasa de mutación debido a la eliminación del gen *mutH*, permitiendo generar variabilidad, así como la posibilidad de explorar varias soluciones adaptativas a causa de la ausencia del sistema PTS. Al mismo tiempo, la ausencia del fragmento cromosomal permitió recuperar la capacidad de asimilación de glucosa debido a la ausencia de los genes *rppH*, *galR* and *lpIT-aas*.

La ausencia del factor de transcripción GalR permitió la desrepresión del regulón de galactosa en estas cepas evolucionadas, cuyo gen más importante fue *galP*, el cual codifica para la permeasa de galactosa GalP que es capaz de transportar glucosa. Como resultado de la ausencia del gen *rppH*, que codifica para la pirofosfohidrolasa RppH, las cepas evolucionadas incrementaron la vida media de diversos mensajeros, entre ellos probablemente el del gen *glk*, que codifica para la glucocinasa Glk, permitiendo la fosforilación de la glucosa entrante. Así mismo, la ausencia de RppH pudo haber incrementado la vida media de genes glucolíticos, incluido el gen *galP*, favoreciendo el transporte y el metabolismo de la glucosa. Esta estrategia evolutiva incrementó el "fitness" 2.6 veces y se considera el evento adaptativo más importante.

Además de los “beneficios” derivados de la ausencia de los genes *galR* y *rppH*, la ausencia del operón *aas-lpIT*, el cual codifica para dos componentes esenciales en el ciclo de recambio de la fosfatidiletanolamina, produjo un incremento de 1.5 veces en el “fitness”. La utilización de la vía de degradación de fosfolípidos como adaptación metabólica a la ausencia del ciclo de recambio, optimizó el metabolismo de las cepas evolucionadas, probablemente por la eliminación del ciclo fútil que representaba el recambio de PE, así como a la reasimilación de los productos de degradación (glicerol-3-fosfato y ácidos grasos).

El conocimiento derivado de este trabajo permite un mejor entendimiento de las bases genéticas de la adaptación, la plasticidad y la fisiología de las cepas evolucionadas. Adicionalmente, esta información también será útil en el diseño y la optimización de estrategias de evolución adaptativa e IVM para el mejoramiento en la redirección de los flujos de carbono hacia la síntesis de compuestos aromáticos empleando cepas PTS⁻.

Abstract: Our group has been working on the construction and characterization of *Escherichia coli* strains that lack the phosphoenolpyruvate: carbohydrate phosphotransferase system (PTS), which is the major bacterial component involved in glucose transport and its phosphorylation. Strains lacking PTS (PB11 strain), in spite of growing very slow on glucose as the only carbon source, accumulates high amounts of phosphoenolpyruvate, which can be diverted towards the synthesis of aromatic compounds. As a metabolic engineering strategy, an adaptive laboratory evolution (ALE) process for the selection of faster growing derivatives of the PB11 strain was carried out in a fermentor in minimal medium with glucose as the sole carbon source. Two of the resulting strains isolated at 120 and 200 hr increased their growth 400% and 500% respectively as compared with the predecessor strain. The whole genome sequencing of these two strains named PB12 and PB13 allowed establishing their evolutive history and the metabolic adaptations in order to overcome the major bacterial component involved in glucose transport.

One of the first adaptive events during the ALE, occurred around the first 12 generations. This event comprises the loss of a 10,328 bp chromosomal fragment on which *rppH*, *galR*, *aas*, *lpIT* and *mutH* among other genes were located. The absence of this chromosomal fragment in the PB12 and PB13 strains, causes an increased mutation rate due to the *mutH* elimination, allowing variability and the possibility to explore several adaptive solutions in order to overcome the PTS lacking condition. At the same time, the absence of the chromosomal fragment was the adaptive solution to recover glucose assimilation capacity due to the lack of the *rppH*, *galR* and *lpIT-aas*, which in turn improve fitness in the evolved strains.

The absence of the transcription factor GalR, allows the derepression of the entire galactose regulon in these evolved strains. One of the most important derepressed genes is the *galP* gene, which codes for the galactose permease GalP and is capable of transporting glucose. In addition, as result of the lack of the *rppH* gene, which codes for a pyrophosphohydrolase RppH, the cells were able to increase the lifetime of several transcripts, among them probably the *glk* mRNA, which codes for the glucokinase Glk, allowing the glucose phosphorylation after its internalization through the GalP transporter. Furthermore, the lack of RppH could increases the glycolytic mRNA's lifetime, including *galP*, favoring glucose uptake and its further metabolism. These adaptive mechanisms improve cell fitness 2.6-fold and are considered the main consequences of the first adaptive event.

Besides the benefits of the lack of *galR* and *rppH* genes, the absence of the *aas-1p1T* operon which codes for the main components of the phosphatidylethanolamine (PE) turnover metabolism produces a 1.5-fold fitness increase. The utilization of the phospholipid degradation pathway as an alternative mechanism, improves the carbon metabolism, probably by the elimination of a futile cycle and the re-assimilation of the degradation products: glycerol 3-phosphate and fatty acids.

This knowledge, allows a better understanding of the genetic basis of growth adaptation, plasticity and the physiology of this evolved *E. coli* strains. Certainly this information will also be useful in the design of improved adaptive laboratory evolution and metabolic engineering strategies for enhancing carbon diversion into the aromatic pathway utilizing PTS⁻ strains.

II. Introducción:

Actualmente una gran cantidad de productos químicos y combustibles se producen principalmente a partir de compuestos fósiles, los cuales juegan un papel muy importante en el desarrollo social y económico del planeta. No obstante, con la creciente preocupación de la sustentabilidad, la liberación de dióxido de carbono a la atmósfera, y en general el impacto negativo al ambiente por el uso de productos petroquímicos y combustibles fósiles, las opciones biotecnológicas de producción a partir de fuentes renovables de carbono son una alternativa ideal [1,2]. La importancia de estas nuevas alternativas va en aumento, debido a que en años recientes, el potencial económico de la industria biotecnológica se ha convertido en una contribución importante para la economía mundial. Hasta el 2011, su valor se estimó en cerca de 500 mil millones de dólares [2].

Dentro del gran abanico que ofrece la naturaleza, la biotecnología ha explorado varias alternativas para llevar a cabo procesos industriales de desarrollo sustentables; siendo las bacterias organismos con un gran potencial para llevarlos a cabo. Durante los últimos 50 años, se han realizado una gran cantidad de estudios en la bacteria *Escherichia coli*, ya que posee características que la convierten en un organismo ideal para desarrollar procesos de producción a gran escala: tiempos cortos de duplicación, lo que implica tasas de crecimiento rápidas y fácil obtención de cultivos de alta densidad celular, bajos costos de producción, conocimiento detallado de su metabolismo y lo más importante, la disponibilidad de excelentes herramientas moleculares para su modificación genética. Por esta razón, *E. coli* es el organismo modelo para estudios biotecnológicos, donde ya se han obtenido mediante ingeniería de vías metabólicas (IVM) excelentes resultados en la producción de diversos metabolitos [3].

Los intermediarios y productos finales de la vía de los aminoácidos aromáticos (AA), son metabolitos por los cuales se tiene un gran interés, ya que debido a sus propiedades y aplicaciones éstos pueden ser utilizados como precursores para sintetizar otros compuestos relevantes en la industria química y farmacéutica. Entre los ejemplos más importantes podemos mencionar a la fenilalanina, la cual es materia prima para la elaboración del edulcorante dietético aspartame [4]. Recientemente, el shikimato (SHK), un intermediario de la vía de los AA ha cobrado una importancia considerable, debido a su empleo como precursor de la molécula oseltamivir fosfato, comercializada por la compañía Roche con el nombre de Tamiflu, para ser empleado como antiviral para la prevención y el tratamiento en infecciones de influenza estacional tipo A y B, influenza aviar H5N1 de 1997 y recientemente del virus de la influenza A/H1N1 [5].

En *E. coli*, el fosfoenolpiruvato (PEP) es uno de los precursores más importantes en la biosíntesis de los compuestos aromáticos, sin embargo, sólo una mínima parte del PEP se canaliza hacia la síntesis dichos compuestos [6]. Estas características celulares impiden que de manera natural se sinteticen altas concentraciones de compuestos aromáticos derivados de este precursor. Desde hace varios años, en el laboratorio del Dr. Francisco Bolívar Zapata, se realizan estudios encaminados al entendimiento del metabolismo y la fisiología de cepas de *E. coli* modificadas genéticamente con el objetivo de incrementar las concentraciones de fosfoenolpiruvato (PEP) en el interior celular, las cuales tienen como fin último la producción de metabolitos de la vía de los AA. Una de las estrategias llevadas a cabo en el laboratorio, fue la eliminación del principal sistema de transporte y fosforilación de glucosa (PTS), ya que consume el 50% del PEP intracelular. Sin embargo, las cepas carentes de este sistema, no son adecuadas para los propósitos de producción, ya que su crecimiento en glucosa está afectado de manera importante [7,8]. A partir de una cepa PTS⁻, denominada PB11, y por un proceso de evolución adaptativa por crecimiento rápido en glucosa, se generaron las cepas PB12 y PB13, las cuales recuperaron parcialmente su capacidad de asimilación de glucosa e incrementaron un 400% y un 500 % respectivamente su velocidad específica de crecimiento (μ) con respecto a la cepa PB11. Por medio de la IVM realizada en la cepa PB12, ha sido posible redirigir parte del PEP no utilizado por PTS hacia la síntesis de compuestos aromáticos a partir de glucosa con altos rendimientos [9,10,11].

Existen varios reportes que indican que cepas de *E. coli* logran incrementar su velocidad de crecimiento, como resultado de mutaciones específicas [12,13,14]. Por tal motivo, con el objetivo de conocer a detalle los cambios genéticos surgidos durante el proceso de evolución adaptativa, que ocasionaron un incremento en el crecimiento de las cepas PB12 y PB13, se secuenció su genoma mediante dos tecnologías distintas.

En este trabajo se presentan todos los cambios genéticos que se generaron en las dos cepas evolucionadas así como el análisis y la caracterización de los más relevantes, lo que llevó a la identificación de aquellas mutaciones que les permitió crecer más eficientemente utilizando la glucosa como única fuente de carbono [15].

III. Antecedentes:

La evolución es el resultado de la adaptación continua de un organismo a ambientes cambiantes o diferentes, basada en cambios genéticos yacentes. Algunas modificaciones en el genoma de un individuo traen consigo la posibilidad de conquistar o sobrevivir en nuevos ambientes. De este modo, la adaptabilidad de un microorganismo se incrementaría, cuanto mayor fuera la capacidad de un individuo o de una población para producir variabilidad genética, aunque con el riesgo de generar mutaciones deletéreas que pongan en riesgo su sobrevivencia. La evolución microbiana depende por

tanto, de dos fuerzas contrapuestas: por un lado, el mantenimiento de la información genética, y por otro lado, un cierto nivel de variación genética que le permita adaptarse a nuevos ambientes. Este equilibrio de fuerzas es la clave de la evolución microbiana [16]. Cualquier situación de estrés para la bacteria (falta de nutrientes, luz ultravioleta, antibióticos, etc.) rompe este equilibrio, poniendo en marcha mecanismos que incrementan la tasa de mutación [17].

Estas capacidades intrínsecas de adaptación bacteriana, pueden ser aprovechadas en biotecnología, para optimizar características de interés [15,18,19], lo que consolida a la evolución adaptativa como una poderosa herramienta para la IVM.

A. Ingeniería de vías metabólicas:

La IVM es una rama de la ingeniería genética que usa principios y técnicas de varias ciencias, cuyo objetivo principal estriba en redirigir, interrumpir e incrementar los flujos de carbono y las propiedades celulares, mediante la interrupción, modificación y/o introducción de nuevas vías metabólicas, así como ampliar el rango de sustratos que se utilizan, con el objetivo de aumentar la producción de metabolitos de interés para varios tipos de industrias (farmacéutica, alimenticia, petroquímica etc.) [20].

B. Transcripción:

En ambientes cambiantes, los organismos vivos deben adaptarse rápidamente para sobrevivir, por lo que es esencial una regulación lo suficientemente fina que controle la homeostasis. En respuesta a estímulos ambientales o al estrés, las complejas redes de regulación ajustan el estado metabólico celular a través de la modulación de la expresión génica, la cual es esencialmente determinada por tres factores: a) La eficiencia de la transcripción del DNA a mRNA, b) la estabilidad del mRNA y c) la tasa de traducción del mRNA a proteínas [21]. De esta forma, la degradación del mRNA afecta directamente la expresión génica, limitando el número de veces que el mensajero puede ser traducido a proteína [22]. Este impacto puede incrementarse con la vida media intrínseca de los mRNA's, la cual difiere hasta en dos órdenes de magnitud en una célula [23].

En *Escherichia coli*, la vida media promedio del mRNA es de 5 minutos; sin embargo la vida media de mensajeros individuales varía desde algunos segundos hasta varias horas [24]. En *E. coli.*, la ribonucleasa más importante involucrada en este proceso es la endonucleasa RNasa E, codificada por el gen *rne*. La RNasa E corta el RNA de cadena sencilla en los segmentos ricos en AU. Los fragmentos resultantes se degradan a mononucleótidos por digestión endo y exonucleolítica [25]. La velocidad de degradación depende esencialmente, en la marcada preferencia de la RNasa E por mRNA's monofosforilados en su extremo 5' [26,27]; mientras que la disponibilidad de este tipo de transcritos depende de la pirofosfohidrolasa de RNA RppH, codificada por el gen *rppH*, que cataliza la

hidrólisis del mRNA trifosfatado, generando RNA's monofosfatados en el extremo 5' (Fig. 1). Este proceso es funcionalmente similar al "decapping" de mRNA's eucariontes [28], y dirige en gran medida la degradación de los mRNA's; lo cual en consecuencia, determina la vida media de dichos mensajeros y por consiguiente la cantidad de mensajeros disponibles para su posterior traducción a proteína.

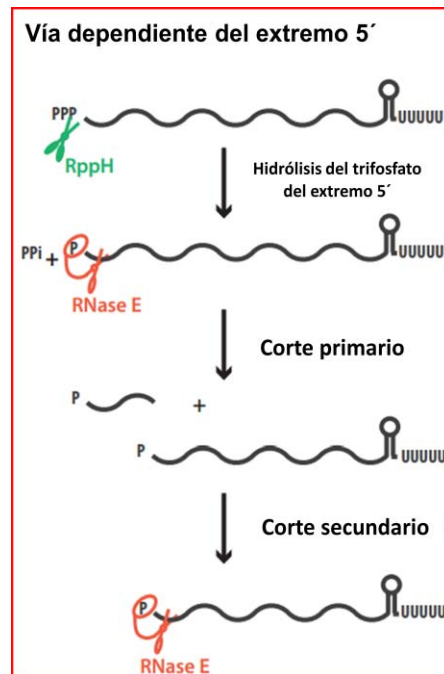


Fig. 1.- Vía de degradación de mRNA's dependiente del extremo 5', en donde participa la pirofosfohidrolasa RppH y la RNasa E (Modificado de Bouvier & Carpousis 2011) [27].

C. **Evolución:**

El término evolución se empleó por primera vez en el siglo XVIII por el biólogo suizo Charles Bonnet en su obra "*Consideration sur les corps organisés*" para describir el conjunto de transformaciones o cambios a través del tiempo que originaron la diversidad de formas de vida que existen sobre la Tierra [29]. No obstante, a pesar de que la evolución se percibe como un proceso continuo de transformación de las especies a través de cambios fenotípicos producidos en sucesivas generaciones, existen cambios que no necesariamente se ven reflejados en características visibles; de hecho, el más mínimo cambio en el genoma de un organismo independientemente de sus efectos, puede ser considerado como cambio evolutivo, siempre y cuando éste se fije en la población. De este modo, la evolución se define como el cambio en las frecuencias alélicas de una población a través del tiempo, lo que deja de manifiesto que el término evolución, se refiere a los cambios que se generan

sobre el genoma de los individuos a causa de mutaciones, pero se define en función de su distribución en las poblaciones [30].

La evolución, puede llevar a la generación de nuevas especies, a la adaptación a distintos ambientes o a la aparición de novedades evolutivas, lo cual depende en gran medida de factores como el ambiente en el que se desenvuelven los organismos y de los efectos que generan las mutaciones [31]. Todos estos factores son tomados para definir los dos mecanismos principales que dirigen la evolución: la selección natural y la deriva génica.

1. Evolución adaptativa: La selección natural

La selección natural consiste en un cambio en la frecuencia de alelos de una población, producida por la reproducción diferencial debida a una mejora en sus capacidades reproductivas y/o de supervivencia. Este proceso se da en una población cuando se cumplen tres condiciones: (1) variación genética entre los individuos de una población inducida por mutaciones; (2) eficacia biológica diferencial asociada a la variación; es decir, ciertos fenotipos o variantes están asociados a una mayor descendencia y/o una mayor supervivencia; y (3) la herencia de la variación, que requiere que la variación fenotípica se deba, al menos en parte, a una variación genética subyacente que permita la transmisión de los fenotipos seleccionados a la siguiente generación. Si en una población de organismos se dan estas tres condiciones, entonces se genera un cambio en la composición genética de la población (evolución) por selección natural. La selección natural es, por lo tanto, el proceso que resulta de las tres premisas citadas [31].

2. Evolución neutra: La deriva génica

La deriva génica consiste en un cambio en la frecuencia de alelos de una población, producida por una reproducción diferencial aleatoria. Una mutación que produzca un alelo inocuo (ni dañino ni benéfico) no se verá seleccionada ni positiva ni negativamente. Por lo tanto, la frecuencia del alelo en la población dependerá únicamente del azar. Por cuestión meramente estadística, la deriva génica tiende a perder los alelos menos frecuentes, inclinándose a la población hacia la homogeneidad genética. En consecuencia, en poblaciones pequeñas -por ejemplo poblaciones fundadoras- los efectos son mucho más marcados; lo que conduce a la fijación de caracteres que no sean adaptativos [32].

De forma independiente, Kimura (1968) y King & Jukes (1969) [33,34], formularon la teoría del neutralismo en evolución molecular. Esta teoría plantea que la inmensa mayoría del cambio molecular es en términos adaptativos neutro, es decir, que la mayor parte de variaciones producen proteínas

que no funcionan ni peor ni mejor que sus predecesoras, por lo que no suponen una mayor o menor adaptación del organismo.

Consecuentemente, la selección natural no puede trabajar sobre estas variantes, por lo que sería la deriva genética el único fenómeno que supondría una fuerza de cambio a nivel molecular. De esta forma, el aumento o disminución de determinados alelos en la población se debería exclusivamente al azar, en forma de deriva genética, y no a la presión selectiva [32,33,34].

En este punto, es importante señalar que la teoría únicamente postula este fenómeno a nivel molecular, es decir, no lo amplía a nivel macroscópico, donde sigue asumiendo que la selección natural es el principal motor evolutivo. Por otro lado, el neutralismo tampoco niega la intervención de la selección natural a nivel molecular, dado que ésta actuaría como filtro para las variaciones dañinas, eliminándolas rápidamente. El único punto en el que discrepa del seleccionismo es en que la fijación de variantes moleculares beneficiosas sería un evento muy poco frecuente [33,34].

D. Evolución bacteriana:

1. *Mutaciones inducidas por estrés*

El proceso evolutivo de un ser vivo se acelera cuanto mayor es su capacidad para producir variabilidad genética, ya sea por mutación, recombinación ó delección. Sin embargo, cuanto mayor sea esta capacidad, mayor también será el riesgo de acumular mutaciones deletéreas. En un controversial estudio publicado en 1988 [35], se describían experimentos donde se sugería que una población bacteriana bajo estrés, podía incrementar su tasa de mutación y dirigir la variabilidad hacia aquellos genes que le ayudarían a contender con el estrés en cuestión. Después de 26 años de investigación, la evidencia sugiere que la variabilidad genética es un proceso altamente regulado, de tal manera que las poblaciones bacterianas tienden a mantener una baja tasa de mutación; aunque con un porcentaje variable de individuos con tasas de mutación superior a la frecuencia modal del resto de la población; siendo este escenario una población en equilibrio. La falta de nutrientes, el frío, el calor, exposición a luz ultravioleta ó antibióticos rompe este equilibrio poblacional e inducen respuestas que tienen consecuencias mutagénicas, donde algunas veces la variabilidad generada al azar, pareciera ser dirigida debido a la selección natural (Fig. 2) [36].

2. *Respuesta SOS*

El daño al DNA que resulta de la exposición a agentes mutagénicos (radiación α , β y γ , luz UV, compuestos desaminantes, alquilantes etc), detona en *E. coli* y otras bacterias una respuesta

bastante efectiva denominada, respuesta SOS, la cual minimiza las posibles consecuencias letales de tales daños [37,38]. La exposición a estos agentes mutagénicos, genera DNA de cadena sencilla (ssDNA), resultado de: 1) daño directo de los mutágenos 2) daño indirecto cuando el DNA es reparado, o 3) daño indirecto debido a una replicación ineficiente. El ssDNA es reconocido por la proteasa RecA, y forma complejos nucleoprotéicos, que inactivan al represor LexA, el cual se encarga de reprimir alrededor de 30 genes en *E. coli* [39,40]. Entre los genes inducidos durante la respuesta SOS, se encuentran aquellos que codifican para DNA polimerasas que tienen la habilidad de replicar DNA dañado (Pol II, Pol IV y Pol V), las cuales toman temporalmente el lugar de la principal DNA polimerasa implicada en replicación (Pol III) [41,42,43,44]. Bajo estas condiciones, las polimerasas IV y V incrementan la tasa de mutación, incluso en ausencia de DNA dañado en alrededor de 3 y 10 veces respectivamente [45]. En células en donde no se induce la respuesta SOS, estas polimerasas generan mutaciones espontáneas; sin embargo contribuyen muy poco para generar variabilidad [46,47,48,49]; no obstante se ha observado que la sola sobreexpresión del gen *dinB* (que codifica para la DNA Pol IV) incrementa la frecuencia de mutación hasta 100 veces [50,51].

3. Respuesta de estrés general

Durante el crecimiento exponencial y bajo condiciones aeróbicas, el factor sigma 70, también conocido como RpoD, transcribe la mayoría de los genes. Cuando *E. coli* entra a la fase estacionaria de crecimiento debido a la limitación en las fuentes de carbono, se induce la respuesta de estrés general [52]. El regulador maestro de esta respuesta es el factor sigma RpoS, también llamado sigma 38, sigma S o KatF, el cual se encarga de transcribir alrededor de 200 genes [53,54]. La reducción en el crecimiento, altas presiones osmóticas, disminuciones en el pH y variaciones en la temperatura son factores que también activan la respuesta general de estrés, la cual está estrictamente controlada a varios niveles e interacciona con otras redes de regulación global. Este complejo sistema de regulación asegura que la célula esté preparada con una variedad de contingencias en presencia de algún estrés.

Existen dos vías conocidas en las que la respuesta general de estrés puede incrementar la tasa de mutación. La primera implica la sobreexpresión del gen *dinB*, que codifica para la DNA polimerasa IV. Después de 10 horas de haber entrado en la fase estacionaria, los niveles de Pol IV se incrementan alrededor de tres veces, mientras que 48 horas después de estar en fase estacionaria, las concentraciones se incrementan entre 30 y 50 veces. Este incremento es independiente de la respuesta SOS, ya que la inducción en la transcripción de *dinB* bajo estas circunstancias esta mediada únicamente por RpoS [55].

La segunda vía se da a través de la subexpresión de enzimas involucradas con el sistema de reparación del DNA MMR (Mismatch Repair). En *E. coli* y algunas bacterias, el complejo MMR reconoce bases mal apareadas en el DNA durante la replicación y escinde la nueva hebra recién sintetizada; la hebra vieja entonces sirve como molde para que se re-sintetice una nueva hebra de DNA. *E. coli* normalmente discrimina entre la hebra nueva (que es la más probable en contener el error) y la hebra templado (hebra correcta) mediante el estado de metilación, ya que únicamente la nueva hebra carece de dicha modificación [56].

Aunque el sistema MMR continúa activo durante la fase estacionaria [57,58], los niveles de los transcritos de las enzimas MutS y MutH implicadas en este sistema se expresan menos gracias a RpoS [59,60]. La evidencia indica que bajo estas circunstancias la tasa de mutación incrementa [61,62,63,64]. Cabe señalar que la disminución de las funciones del sistema MMR no genera mutaciones “*de novo*”, sino que preserva las mutaciones generadas naturalmente durante la replicación.

4. Respuesta estricta (*stringent response*)

Cuando *E. coli* y otras bacterias tienen escasez de aminoácidos o se da una baja considerable en las fuentes de carbono, la célula experimenta por un lado, una disminución en la síntesis de tRNA's y rRNA's, así como en la síntesis de proteínas; y por otro lado, un incremento en el recambio protéico, así como la inducción de RpoS y enzimas involucradas en la biosíntesis de aminoácidos. Esta reacción conocida como respuesta estricta (“*stringent response*”), es mediada por dos metabolitos: el pentafofato de guanosina y el tetrafofato de guanosina (pppGpp y ppGpp respectivamente) [65,66]. Ambos metabolitos (con ayuda del co-efector DskA), regulan la transcripción uniéndose a la RNA polimerasa y cambiando la distribución de su actividad en varios promotores [67], favoreciendo aquellos reconocidos por el factor RpoS. Aquí es importante señalar que la actividad de RpoS se optimiza con la presencia de (p)ppGpp [68]. De este modo la respuesta general de estrés (dependiente de RpoS), y la respuesta estricta (dependiente de los metabolitos (p)ppGpp) interactúan de forma concertada de acuerdo al tipo de estrés y no son excluyentes [69,70].

Durante la transcripción se reclutan sistemas de reparación de DNA debido a que en este proceso, se generan secciones de DNA con hebras separadas (“burbujas”), y las hebras de DNA de cadena sencilla son más susceptibles a mutaciones por diferentes tipos de daño (p. ej. la citocina se desamina y forma uracilo 100 veces más rápido cuando el DNA se encuentra en cadena sencilla) [71]. Se ha reportado que la alta transcripción de los genes incrementa la aparición de mutaciones en *E. coli* [72,73,74,75], *Bacillus subtilis* [76] y en *Saccharomyces cerevisiae* [77]. De este modo se

plantea que durante la respuesta estricta en donde se induce la transcripción de muchos genes (p. ej. para la biosíntesis de aminoácidos), se generen mutaciones que en cierto modo estarían dirigidas solo a los genes que se inducen durante este tipo de respuesta [78,79,80].

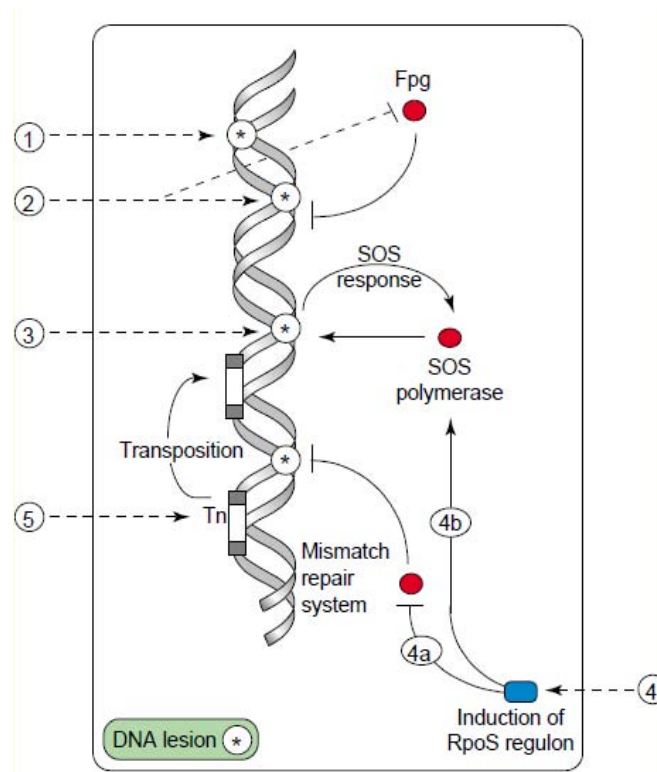


Fig. 2.- Mecanismos moleculares asociados a la mutagénesis inducida por estrés. 1) Agentes mutagénicos 2) Agentes medioambientales 3) Radiación UV 4) Condiciones medioambientales 4a) Inhibición de los sistemas de reparación 4b) Inducción de la respuesta SOS 5) Transposones (Tomada de Tenailon et al. 2004) [17].

E. Experimentos de evolución adaptativa:

E. coli ha resultado de particular importancia para aplicaciones biotecnológicas, ya que su utilización en la producción de compuestos empleados por diferentes tipos de industrias ha derivado en buenos resultados [81,82], sobre todo porque la catálisis biológica ofrece muchas ventajas sobre la síntesis química, entre ellas, la producción estereo selectiva de compuestos químicos [83].

Como complemento a la IVM, la biología sintética y la biología de sistemas son campos emergentes que están resultando exitosos para el diseño y la implementación de procesos de producción [84,85,86,87]. Otra de las herramientas en la IVM, la ingeniería genética clásica, se ha complementado con la selección artificial de microorganismos, basada en procesos de evolución adaptativa con el objetivo de obtener cepas robustas, con potencial para ser modificadas y optimizadas en la producción de algún compuesto [88,89,90].

Existen tres formas de llevar a cabo un experimento de evolución adaptativa:

- a) Placa: Los experimentos en placa consisten en la propagación continua en medio sólido de colonias específicas, seleccionadas de forma arbitraria durante un periodo de tiempo determinado. Como resultado de este tipo de procedimiento se genera una selección artificial de clones mediante cuellos de botella, lo cual disminuye la competencia. La dinámica genética en este tipo de procesos de adaptación ocasiona la acumulación de mutaciones a lo largo del tiempo, favoreciendo aquellas que disminuyen el “fitness” y generan por tanto microorganismos medianamente adaptados y muchas veces genéticamente inestables [91].
- b) Cultivo continuo: Este tipo de evolución es la más ampliamente utilizada, y consiste en el empleo de un quimiostato con capacidad mínima de 1L, en donde se alimenta medio fresco a distintas velocidades de acuerdo a los objetivos del proceso. Usualmente este tipo de experimentos implica dos fases, cuyo primer paso es la generación de biomasa (que no es exenta a los mecanismos evolutivos), para posteriormente iniciar con el cultivo continuo a una tasa de dilución específica, consiguiendo mantener las condiciones de crecimiento estables durante largos periodos de tiempo. La dinámica genética en este tipo de procesos es dirigida por la selección natural, donde se favorece la interferencia clonal, ocasionando por tanto la acumulación de mutaciones benéficas y eliminando casi por completo la aparición de mutaciones deletéreas, lo cual a su vez, desemboca en la generación de diversidad genética, representada por diversas poblaciones que están muy bien adaptadas [91].
- c) Transferencia serial: Consiste en la propagación serial de inóculos en matraz a lo largo de un periodo de tiempo específico. Como resultado de este tipo de evolución se obtienen individuos bien adaptados, ya que a lo largo del cultivo se genera diversidad y por tanto competencia. Al igual que en los experimentos en quimiostato la selección natural dirige la dinámica genética durante el proceso [91].

Es importante comentar que el resultado de las diferentes metodologías, así como la dinámica genética durante los procesos, está estrechamente relacionado con la arquitectura genética de los microorganismos, lo cual ocasiona que diferentes especies o cepas tomen distintos “camino evolutivos” (“fitness landscape”) aún bajo condiciones idénticas. Esto genera en consecuencia un mayor grado de complejidad al momento de interpretar sucesos y de encontrar explicaciones tanto a procesos adaptativos, como a mecanismos evolutivos y en consecuencia impide la extrapolación directa de resultados.

F. Secuenciación de DNA:

La secuenciación del DNA y su disciplina conocida como genómica, son áreas relativamente nuevas, las cuales son el resultado de la combinación de la biología molecular y la química de nucleótidos. Ambas disciplinas nacieron en la década de los 50's; sin embargo fue hasta la década de los 70's, cuando Frederick Sanger en el Medical Research Council (MRC), en Cambridge Inglaterra, inició con la investigación acerca de un método para secuenciar moléculas de DNA; y no fue sino hasta 1977, cuando publicó el primer método de secuenciación, el cual consistía en añadir, durante la polimerización de una cadena de DNA, nucleótidos modificados químicamente [92]. Estos nucleótidos tenían un átomo de hidrógeno en lugar de un grupo hidroxilo en su extremo 3', lo que impedía la adición de un nuevo nucleótido en la cadena, deteniendo la elongación. De este modo, mezclando por separado análogos de los cuatro nucleótidos en cuatro reacciones independientes, se generaban al azar cadenas de diferentes tamaños; las reacciones se cargaban en geles de poliacrilamida y una vez terminada la electroforesis, el gel se ponía en contacto con una película fotográfica de autorradiografía, ya que contenían ³²P debido a la adición de ATP marcado en la reacción. En el gel, las diferentes moléculas de DNA se separaban de acuerdo a su tamaño (de la más larga arriba, a la más corta abajo); de este modo era posible identificar a simple vista la secuencia nucleotídica de acuerdo a la migración de las moléculas y al tipo de nucleótido añadido en la reacción [92].

Los instrumentos automáticos actuales de secuenciación masiva han logrado incorporar avances tecnológicos en cuanto a miniaturización e informática a métodos más sofisticados de secuenciación construidos a partir del método de Sanger. Gracias a estos equipos ha sido posible conocer el genoma de organismos modelo (*Escherichia coli*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Mus musculus*, *Arabidopsis thaliana*, *Zea mays*), incluido el ser humano; no obstante desde el año 2005, nuevos equipos han logrado disminuir los costos y acortar los tiempos de secuenciación, permitiendo incrementar el número de genomas producidos con una precisión cada vez mayor [93].

Actualmente debido a la creciente expansión de equipos y tecnologías de secuenciación masiva, ha sido posible llevar a otro nivel los estudios de evolución, donde la genómica comparativa ha emergido como una poderosa herramienta para analizar la evolución bacteriana, la cual se da en periodos de tiempo relativamente cortos. Esta estrategia ha permitido entender el mecanismo de evolución en escenarios específicos tales como emergencias sanitarias, adaptación a ambientes perturbados ó eventos adaptativos durante cultivos bacterianos que poseen capacidades de producción [12,13].

1. Componentes del sistema PTS:

La enzima I (EI) y la proteína de histidina (HPr) son proteínas citoplásmicas que participan en la fosforilación de todos los de carbohidratos PTS y se denominan proteínas generales de PTS. Por otra parte, el conjunto de enzimas II (EII) es específico para determinados carbohidratos. Estas proteínas pueden estar unidas a la membrana y tener tres dominios (A, B y C), o estar constituidas de dos o más proteínas, de las cuales al menos una está unida a membrana [97].

En el caso del transporte y la fosforilación de la glucosa, la cascada de señalizaciones inicia con la enzima EI (codificada por el gen *ptsI*), la cual se autofosforila utilizando fosfoenolpiruvato (PEP) y transfiere el grupo fosfato a la proteína Hpr (codificado por el gen *ptsH*). Posteriormente se transfiere el fosfato de la proteína Hpr al componente IIA (codificado por el gen *crr*), que a su vez lo transfiere al complejo formado por la proteína IIBC (codificado por el gen *ptsG*). Este complejo funciona como una permeasa, fosforilando de manera concomitante al carbohidrato transportado. Además del transporte y fosforilación de azúcares, el sistema PTS tiene un papel fundamental en la distribución de flujos de carbono, así como en la regulación transcripcional global, que controla el consumo preferencial de glucosa sobre otras fuentes de carbono [97].

2. Represión catabólica:

En *E. coli* y otras bacterias cercanas filogenéticamente, una fuente de carbono rápidamente metabolizable como la glucosa, inhibe la síntesis de enzimas involucradas en el metabolismo de otras fuentes de carbono. Este hecho involucra al menos cuatro fenómenos relacionados fisiológicamente pero diferentes en cuanto a su cinética y mecanismo, denominados: represión permanente, represión transitoria, exclusión del inductor y expulsión del inductor.

La represión permanente y transitoria se refieren a formas de represión de larga y corta duración que ocurren en respuesta a la adición de glucosa al medio de cultivo de un microorganismo; ambos fenómenos son independientes de la concentración intracelular de los inductores. Por su parte, la exclusión y la expulsión del inductor se refieren a la inhibición de la toma del inductor hacia la célula y la estimulación de la liberación del inductor del citoplasma respectivamente. Estos cuatro fenómenos se engloban en el término “represión catabólica” debido a que causan la represión de la síntesis de ciertas enzimas en respuesta a la presencia de una fuente de carbono y de energía alternativa en el medio de cultivo [98]. El sistema PTS es capaz de modular coordinadamente las concentraciones de AMPc y de los inductores intracelulares que controlan la represión catabólica y la exclusión del inductor. El mecanismo de control inicia con la proteína efectora, en este caso, la enzima IIA^{Glc} del sistema PTS, la cual se encuentra desfosforilada cuando la glucosa está presente en el medio; en este estado, la proteína se une a varias permeasas no PTS, inhibiendo el consumo de azúcares no PTS (Fig. 4). Del mismo modo, la proteína puede unirse a proteínas solubles implicadas en vías

catabólicas como la enzima glicerol cinasa, inhibiendo su actividad [99]. Adicionalmente, la forma no fosforilada del componente IIB^{Glc} une al represor Mlc, disminuyendo la represión sobre los genes *ptsHI*, *ptsG*, *mlc*, *manXYZ* y *malT* (Fig. 4) [100]. Cuando no hay glucosa en el medio el componente IIA^{Glc} y IIB^{Glc} se mantienen principalmente en su forma fosforilada. La proteína IIA^{Glc-P} se une a la enzima adenilato ciclasa induciendo su actividad biosintética de AMPc. Bajo estas circunstancias la concentración de AMPc se incrementa e interacciona con la proteína CRP; este complejo AMPc-CRP se encarga de inducir los genes que estaban bajo represión catabólica (Fig. 4) [101]. Por su parte, la proteína IIB^{Glc-P} pierde su capacidad de secuestrar a Mlc, por lo tanto este factor de transcripción puede unirse a secuencias operadoras blanco, ocasionando la represión de genes involucrados en el consumo de glucosa [100].

En el esquema de regulación anterior, la desfosforilación de IIA^{Glc-P} en respuesta a la presencia de glucosa, da como resultado la inactivación coordinada de la enzima adenilato ciclasa (represión permanente) y la inhibición de las actividades de permeasas y enzimas catabólicas que generan los inductores intracelulares de varios operones (exclusión del inductor). La reducción intracelular simultánea de AMPc y del inductor provoca un control dual sobre la transcripción de genes catabólicos; el AMPc actúa a través de CRP, mientras que la respuesta a varios inductores es mediada al menos por tres tipos de factores de transcripción: represores, activadores y anti terminadores [98].

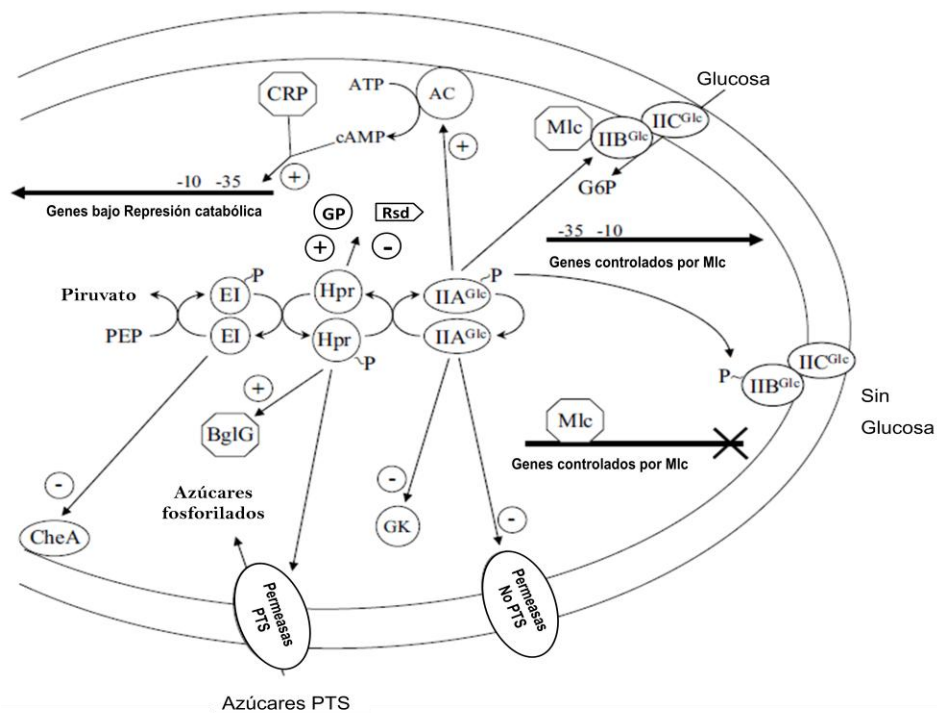


Fig. 4.- Mecanismos moleculares de regulación durante la represión catabólica (Modificada de Gosset 2005) [94].

Las proteínas El y Hpr también tienen funciones de regulación. En su estado no fosforilado, Hpr activa a la proteína glucógeno fosforilasa (GP), mientras que su forma fosforilada tiene un efecto similar sobre BglG, el cual es un factor de transcripción que induce la expresión del operón *bgl*, el cual codifica para enzimas involucradas en el consumo y catabolismo de β -glucósidos (Fig. 4) [102,103].

Estudios recientes demuestran que Hpr puede considerarse como un regulador global, ya que es un antagonista del factor anti-sigma Rsd, el cual se encarga de secuestrar al factor sigma 70 impidiendo su unión con la polimerasa, y favoreciendo que otros factores sigma (sigma S principalmente) formen el complejo de transcripción. En presencia de glucosa, Hpr se encuentra esencialmente en su forma desfosforilada; en este estado es capaz de unirse al factor Rsd, impidiendo que sigma 70 sea secuestrado y por lo tanto favoreciendo su actividad. Cuando no existe glucosa en el medio la forma fosforilada de Hpr^P, no es capaz de unirse a Rsd, por tal motivo este factor anti-sigma puede secuestrar a sigma 70 y, en consecuencia, se favorece la transcripción de genes dependientes del factor sigma S [104].

Por otra parte, la forma no fosforilada de El es capaz de unirse a la proteína CheA, inhibiendo su autofosforilación, lo que impacta sobre los mecanismos de desplazamiento en *E. coli* durante la quimiotaxis [105].

H. Nodo PEP-PYR en el metabolismo central

En *E. coli*, el nodo PEP-PYR (fosfoenolpiruvato-piruvato) es de suma importancia, ya que ambos metabolitos están involucrados en diversos procesos metabólicos ya que son precursores en la biosíntesis de distintos metabolitos, entre ellos, de los compuestos aromáticos. No obstante, sólo una mínima parte del PEP producido mediante la glucólisis (alrededor del 3%), se canaliza a producir dichos compuestos, pues el restante 97% del PEP se utiliza por otras enzimas en distintas reacciones metabólicas: un 16% del PEP lo utiliza la fosfoenolpiruvato carboxilasa (Ppc) para generar oxaloacetato; un 15% lo utiliza las piruvato cinasas (PykA y PykF) para generar piruvato, un 16% por la acción de la UDP-N-acetil-glucosamina-enol-piruvil-transferasa para formar N-acetil- α -D-glucosamina-1-fosfato, y el restante 50% por el sistema de transporte de azúcares dependiente de PEP ó sistema PTS (Fig. 5) [6].

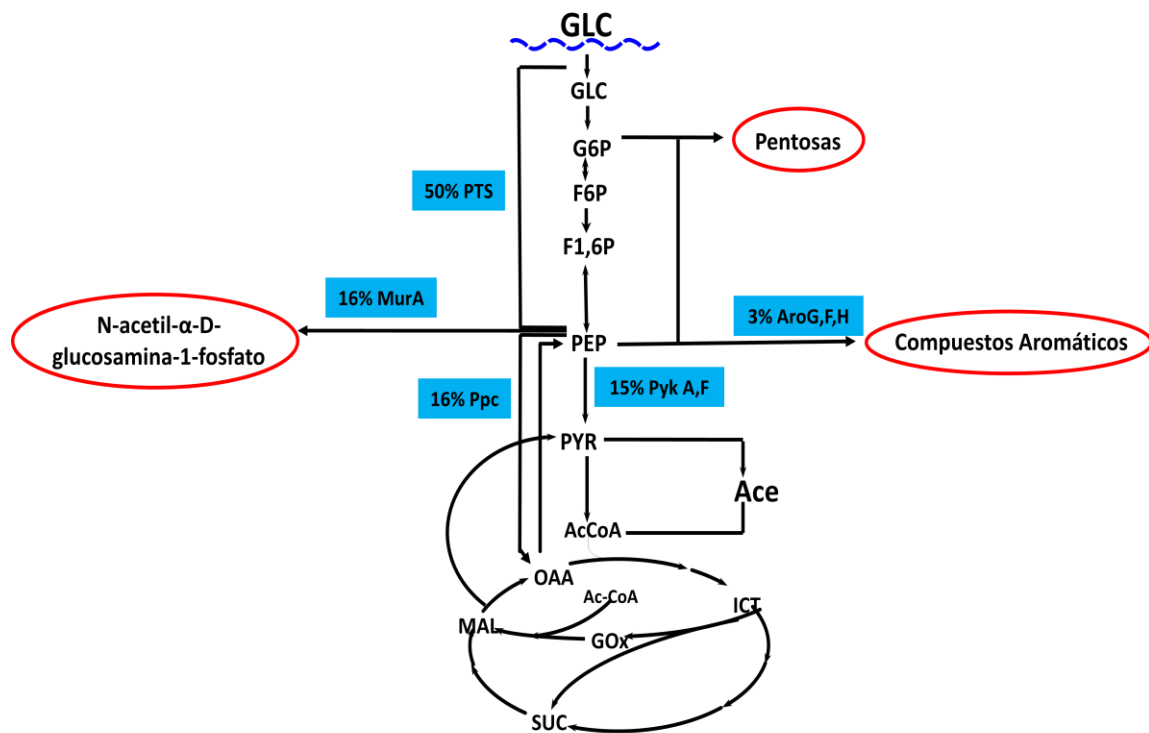


Fig. 5.- Principales vías de utilización de PEP en el metabolismo central.

Estas características le impiden a la célula tener una mayor cantidad de PEP para incrementar la síntesis de metabolitos de interés comercial derivados de este precursor cuando *E. coli* utiliza la glucosa como fuente de carbono. Por esta razón, se han realizado una serie de esfuerzos encaminados a la comprensión y la manipulación de las vías metabólicas que tienen que ver con la síntesis, el catabolismo y la utilización del PEP en diferentes especies y variedades de bacterias.

I. Evolución adaptativa como herramienta en la IVM para la producción de compuestos aromáticos:

Una de las estrategias realizadas en el laboratorio para incrementar la poza intracelular de PEP para utilizarlo en la síntesis de compuestos aromáticos, fue la inactivación del principal sistema de transporte de glucosa (PTS) en la cepa JM101, codificado por el operón *ptsHlcr*. Esto en teoría duplicaría la disponibilidad de PEP para la biosíntesis de aminoácidos aromáticos [7,10,106], lo que dio origen a la cepa PB11. La capacidad de asimilación y de utilización de glucosa de esta cepa se afectó seriamente, ya que se observó una disminución del 85% en su μ con respecto a la cepa silvestre ($\mu=0.7 \text{ hr}^{-1}$ vs $\mu=0.1 \text{ hr}^{-1}$) [8,107], por lo que no fue adecuada para los propósitos de producción. Para contender con dicho problema, se aprovechó la capacidad de adaptación bacteriana y a partir de la cepa PB11, se realizó un experimento de evolución adaptativa por crecimiento rápido

en glucosa, que hizo posible la generación de cepas capaces de crecer más rápidamente, utilizando la glucosa como fuente de carbono.

Para este experimento, la cepa PB11 se creció en medio mínimo M9 con 2 g/L de glucosa, en fermentador de 1L (cultivo lote). Bajo esas condiciones se genera una presión de selección que favorece mutantes que adquieren la capacidad de crecer más rápidamente que la cepa original (PB11). El cultivo se mantuvo hasta alcanzar la fase estacionaria, para posteriormente comenzar a alimentar el fermentador con medio nuevo a la misma concentración (cultivo continuo), a tasas de dilución con tres tiempos de retención cada una, que se fueron incrementando gradualmente (Fig. 6). Este proceso, permitió el aislamiento de mutantes de acuerdo a su μ . Se aislaron seis mutantes (dos colonias de cada una de las diluciones 0.4, 0.6 y 0.8 hr^{-1}) y dos se seleccionaron para su posterior caracterización: la cepa PB12 ($\mu=0.42 \text{ hr}^{-1}$), aislada a las 120 hrs, y la cepa PB13 ($\mu=0.5 \text{ hr}^{-1}$) aislada a las 200 hrs (Fig. 6) [8,107,108]. Mediante la expresión e inactivación de genes clave en la cepa PB12, la cual recuperó parcialmente su capacidad de asimilación de glucosa e incrementó su crecimiento un 400% con respecto a la cepa PB11 [8,107,108], ha sido posible redirigir parte del PEP no utilizado por PTS hacia la síntesis de compuestos aromáticos, lo que resultó en la generación de una cepa sobreproductora de SHK con el mayor rendimiento reportado hasta el momento [11].

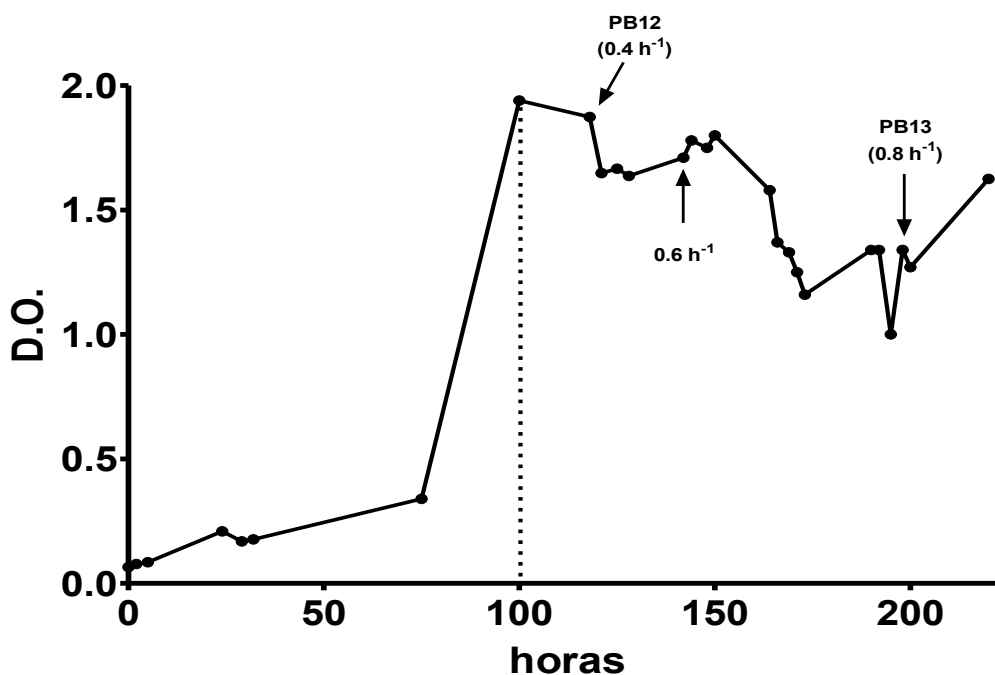


Fig. 6.- Selección de mutantes PTS- Glc⁺ de un cultivo continuo a partir de la cepa PB11 (PTS⁻ Glc⁻). Las flechas indican el tiempo de aislamiento de varias cepas Glc⁺ incluyendo la PB12. La línea punteada indica el final del cultivo en lote y el inicio del cultivo continuo. Las tasas de dilución ($D=\text{h}^{-1}$) 0.4, 0.6 y 0.8. se especifican en la imagen.

1. Características de las cepas PTS:

Para establecer las características de cada una de las cepas generadas (PB11, PB12 y PB13), se han llevado a cabo una serie de estudios encaminados a la comprensión de su metabolismo. De este modo, se han realizado análisis de flujos de carbono mediante resonancia magnética nuclear [109] y análisis de expresión de genes mediante RT-qPCR [8]. Con los estudios realizados hasta el momento, se ha demostrado que debido a la ausencia del sistema PTS, estas cepas carecen de represión catabólica; por ello son capaces de co-utilizar glucosa y otras fuentes de carbono y en ellas coexisten tanto el metabolismo glucolítico como el gluconeogénico [8]. Cabe señalar que la cepa PB12 ha demostrado ser una buena productora de AA; por lo que los esfuerzos para entender el metabolismo se han centrado principalmente en las cepas PB11 y PB12, dejando de lado el análisis y la caracterización de la cepa PB13.

a) Cepa PB11:

El comportamiento fisiológico de esta cepa, a pesar de crecer en glucosa como única fuente de carbono en altas concentraciones (2 g/L), es similar al que se presenta cuando se tienen en el medio de cultivo sustratos pobres en esqueletos de carbono ó cuando las fuentes de carbono son escasas, es decir, una condición de estrés permanente que activa vías gluconeogénicas, anapleróticas y de reciclamiento de carbono, ocasionado en parte a la pérdida de su capacidad para reconocer y transportar la glucosa. Adicionalmente, la ausencia de PTS ha ocasionado efectos pleiotrópicos, principalmente en la regulación, involucrada entre otras cosas en la jerarquización de la asimilación de fuentes de carbono a metabolizar [110]. Esto podría explicar el incremento en la transcripción de muchos genes relacionados con el transporte y utilización de otras fuentes de carbono (galactosa, maltosa, fucosa, arabinosa, xilosa, ribosa, acetato, glicerol, glicerol-3-fosfato) [111]. Cabe señalar que esta cepa no sobreexpresa los genes involucrados en el transporte y catabolismo de glucosa. Por otra parte, algunos genes involucrados en la asimilación de otro tipo de nutrientes (amonio, sulfato, ácido galacturónico y diversos nucleótidos), también están sobreexpresados. La mayoría de las características observadas se pueden explicar por la inducción de mecanismos activados durante el estado de hambruna ("scavenging") que se presenta en esta cepa; sin embargo, el sentido biológico de algunos de ellos aún no se comprenden del todo, debido a que algunos de estos sistemas sirven para mantener el equilibrio osmótico celular [111].

A pesar de observarse una sobreexpresión 12.4 veces mayor del gen que codifica para la permeasa de galactosa (*galP*), la capacidad de transporte de glucosa en la cepa PB11 es 10 veces menor que el de la cepa silvestre JM101 (10.6 mmol Glc·h⁻¹·gDW vs 1.1 mmol Glc·h⁻¹·gDW). La permeasa de galactosa es el principal sistema de transporte de glucosa en ausencia del sistema PTS [6], no

obstante este sistema no está acoplado a la fosforilación del carbohidrato transportado, por lo que la activación de las moléculas internalizadas debe darse por cinasas solubles en el citoplasma. La principal encargada de dicha fosforilación es la glucocinasa codificada por el gen *glk*, cuyo transcrito se encuentra a los mismos niveles que los encontrados en la cepa silvestre. Por esta razón, es probable que la limitante en esta cepa no sea el transporte de la glucosa, sino su fosforilación; de hecho la expresión de *glk* en plásmido incrementa al doble la μ de esta cepa [108].

Con respecto a los flujos de carbono, esta cepa destina alrededor de un 40% del carbono consumido hacia la vía glucolítica, y una cantidad superior al 50% hacia la ruta de las pentosas fosfato (Fig. 7) [109]. Esta cepa muestra también un mayor flujo de carbono a través del ciclo de los ácidos tricarbónicos (TCA) con respecto a la cepa silvestre, lo que supondría una mayor conversión de glucosa a CO₂; sin embargo, se ha observado que la mayor parte del carbono que entra a TCA en la cepa PB11, se recicla por la vía del glioxilato, cuyos genes se encuentran altamente sobreexpresados, impidiendo la pérdida de carbono como CO₂ en la parte baja del TCA, cuyos genes no se sobreexpresan (Fig. 7) [109,111]. Es probable que la capacidad reducida de asimilación de glucosa en esta cepa ocasione concentraciones bajas de intermediarios clave y la distribución de los flujos de carbono, esté influida no solo por mecanismos de respuesta a estrés, sino que también sean reflejo de la afinidad de las enzimas por su sustrato en nodos del metabolismo, donde 2 ó más enzimas compiten por el mismo sustrato [109].

b) Cepa PB12:

En la cepa PB12 a pesar de carecer del sistema PTS, no se presentan mecanismos de respuesta a estrés tan severos como la inducción generalizada de los sistemas específicos para la asimilación de otras fuentes de carbono, con excepción del regulón *gal* (la mayoría de los genes involucrados en la asimilación y catabolismo de la galactosa están sobreexpresados). No obstante es bastante notable que más del 90% de los genes del metabolismo central (glucólisis, gluconeogénesis, vía de las pentosas fosfato, TCA, vía del glioxilato y genes de respiración) analizados mediante RT-qPCR, presentan elevados niveles de transcrito en esta cepa (Fig. 7).

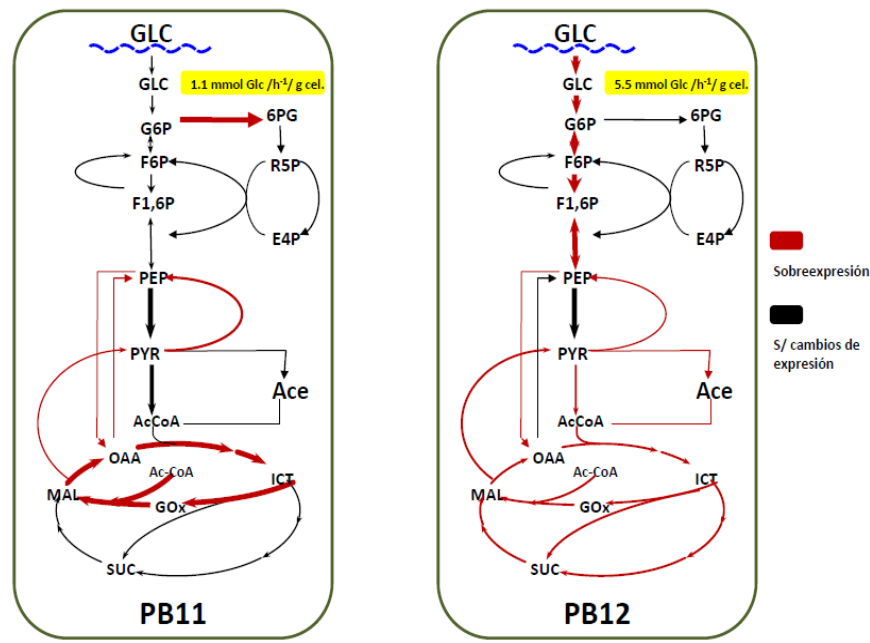


Fig. 7.- Esquema del metabolismo central en las cepas PB11 y PB12. Las flechas indican la dirección de la reacción, mientras que su grosor estima el flujo de carbono destinado. El color rojo indica incrementos en la transcripción de los genes que codifican para las enzimas que catalizan las reacciones indicadas. Este esquema se desprende de datos obtenidos por RMN y RT-qPCR para determinar flujos de carbono y niveles de expresión de genes del metabolismo central respectivamente.

Esta cepa también tiene niveles elevados de transcrito del gen *galP* (13.1 veces más con respecto a la cepa JM101), lo que permite un eficiente transporte de glucosa, aunque a diferencia de la cepa PB11, esta cepa sobreexpresa el gen *glk*, alrededor de 2.2 veces más con respecto a la cepa silvestre, lo que le permite fosforilar de forma más eficiente a la glucosa, lo que resulta en una asimilación 5 veces mayor con respecto a la cepa PB11 ($5.5 \text{ mmol Glc} \cdot \text{h}^{-1} \cdot \text{gDW}$ vs $1.1 \text{ mmol Glc} \cdot \text{h}^{-1} \cdot \text{gDW}$), aunque dos veces menor con respecto a la cepa silvestre ($10.6 \text{ mmol Glc} \cdot \text{h}^{-1} \cdot \text{gDW}$ vs $5.5 \text{ mmol Glc} \cdot \text{h}^{-1} \cdot \text{gDW}$). La cepa PB12 cuenta con un metabolismo bastante glucolítico; de hecho, el flujo de carbono que destina hacia esta vía es superior al 90%, y solo un 5% del carbono se destina hacia la vía de las pentosas fosfato (Fig. 7). Como se mencionó con anterioridad, en la cepa PB12 se presentan niveles de transcrito elevados en prácticamente todos los genes que codifican para enzimas involucradas con la glucólisis, lo que permite un mejor aprovechamiento de la glucosa; de hecho, en esta cepa ocurre un sobreflujo de carbono por la vía glicolítica que conlleva a la formación de acetato como producto de desecho (36%). La sobreexpresión de los genes glucolíticos en la cepa PB12, se observó tanto en fuentes ricas de carbono (glucosa), como en fuentes pobres de carbono (acetato) [8,112], lo que sugiere que durante el proceso de evolución se pudieron generar una o varias mutaciones responsables de este fenómeno.

En cualquier cepa, la reducción en el flujo de carbono hacia la vía de las pentosas tiene consecuencias sobre la fisiología celular, debido a que bajo estas condiciones de crecimiento, esta vía es la principal proveedora de NADPH [113,114]. No obstante, la cepa PB12 al parecer es capaz de suplir esta disminución de poder reductor, sobreexpresando los genes que codifican para las enzimas málicas (*maeB* y *sfcA*), logrando canalizar un mayor flujo de carbono por esta vía anaplerótica la cual también es capaz de regenerar NADPH, y por ello compensar en cierta medida la demanda de este cofactor [109,112].

El flujo de carbono a través de las enzimas PykA y PykF, las cuales forman piruvato a partir de PEP, se incrementa de un 40% en la cepa silvestre hasta un 128% en las cepas PB11 y PB12, lo que demuestra que la eliminación del sistema PTS ocasionó un incremento en la disponibilidad del PEP intracelular (Fig. 7) [109]. En el caso de la cepa PB12, una mayor internalización de glucosa vía GalP-Glk, permite tener una mayor concentración de PEP que no se utiliza por el sistema PTS, y que puede ser canalizado hacia la vía de los compuestos aromáticos.

Se ha especulado que la sobreexpresión de los genes del TCA y aquellos involucrados en respiración, podría deberse a un mal funcionamiento del sistema de dos componentes ArcA/B, como resultado de una mutación encontrada en el gen *arcB*, que cambia un residuo de tirosina por una cisteína en la posición 71 de la proteína [8].

Se ha detectado también, una mutación en el gen que codifica para el factor sigma RpoS, el cual está involucrado en la transcripción de varios genes que se expresan en fase estacionaria y/o en condiciones de estrés por limitación de carbono. Esta mutación puntual genera un codón de paro en la posición 98 del gen, lo que produciría un péptido de 32 aminoácidos, en lugar de la proteína completa. No obstante, es muy probable que esta mutación se revierta en la cepa JM101 y sus derivadas por la acción del tRNA^{Gln}, codificado por el gen *supE*, ya que estas cepas poseen una mutación en el tercer nucleótido del anticodón de este tRNA, la cual ocasiona en cierta proporción, la inserción de una glutamina cuando existe un codón de paro durante la traducción [115].

Se ha establecido que la secuencia promotora que reconocen los factores sigma 70 y sigma S, son muy parecidos; incluso se ha comprobado que algunos promotores pueden ser reconocidos por ambos factores sigma [116]. En ese sentido, se ha sugerido que el factor sigma 70 de la cepa PB12 ha suplido parcialmente algunas de las funciones llevadas a cabo por el factor sigma S; de hecho la inactivación de éste último en la cepa PB12 afecta en un 10% su crecimiento; contrario a lo que sucede con la cepa PB11, donde el crecimiento se ve afectado en un 50% [117].

c) Cepa PB13:

La cepa PB13 se aisló en la hora 200 durante el experimento de evolución adaptativa, y presenta una μ de 0.49 h^{-1} , lo que representa un incremento cercano al 500% con respecto a la cepa PB11. Experimentos de actividad enzimática para la cinasa Glk en esta cepa demuestran niveles elevados, superiores incluso a los presentes en la cepa PB12 (127 vs $136 \text{ nmol/min}^{-1}/\text{mg de proteína}^{-1}$) [109]. Adicionalmente, cuenta con una capacidad de asimilación de glucosa casi 8 veces mayor con respecto a la cepa PB11, y 1.5 veces mayor con respecto a la cepa PB12 ($8.4 \text{ mmol Glc}\cdot\text{h}^{-1}\cdot\text{gDW}$); por lo que se especula que también internaliza la glucosa mediante GalP y la fosforila mediante Glk. Mediante un análisis de flujos de carbono por RMN, utilizando ^{13}C , se comprobó la gran capacidad glucolítica de esta cepa, la cual canaliza cerca de un 90% del carbono hacia la glucólisis, mientras que el flujo restante se envía hacia la vía de las pentosas fosfato [109]. A pesar de que ambas cepas evolucionadas (PB12 y PB13) tienen una distribución del flujo de carbono muy similar, la cepa PB13 logra internalizar una mayor cantidad de glucosa, lo cual genera un sobreflujo de carbono en la vía glucolítica, que a su vez ocasiona la producción de acético en concentraciones similares a las generadas por la cepa silvestre JM101 (56%) [109]. Se ha realizado ingeniería de vías metabólicas para la producción de compuestos aromáticos en esta cepa, no obstante la obtención de rendimientos aceptables, la cepa PB12 ha demostrado ser más eficiente [118], debido muy probablemente a la gran cantidad de carbono que es canalizado para la generación de acético.

IV. Propósitos y alcances de este trabajo:

El propósito principal de este trabajo, fue identificar los cambios genéticos surgidos durante el proceso de evolución que dio origen a las cepas PB12 y PB13, y que le confirieron la capacidad de metabolizar de mejor forma la glucosa en ausencia de su principal sistema de transporte. No obstante, la realización de esta tarea no se limita únicamente al entendimiento de aspectos metabólicos, ya que derivado de este conocimiento, tendremos la capacidad de incursionar en la resolución de varios cuestionamientos que tienen que ver con aspectos de ciencia básica y aplicada. En el primer caso, los datos que de aquí se derivaron, permitirán un mejor entendimiento de las bases genéticas de la adaptación, la plasticidad y la fisiología de estas cepas evolucionadas, que podrán hacerse extensivas, hasta cierto punto, a los procesos evolutivos dirigidos por selección natural que se originan en ambientes de estrés por limitación de fuentes de carbono y que hacen posible la supervivencia bacteriana.

Dentro de la otra vertiente, el conocimiento aquí generado, será útil en el diseño de nuevas estrategias de IVM para optimizar la canalización del flujo de carbono hacia la vía de compuestos aromáticos en cepas PTS.

V. Objetivo general:

Evaluar el efecto de las mutaciones presentes en las cepas PB12 y PB13 sobre su crecimiento y su perfil de transcripción.

A. Objetivos específicos:

1. Secuenciar el genoma de las cepas JM101, PB11, PB12 y PB13.
2. Identificar los cambios genéticos ocurridos durante el proceso de evolución adaptativa en las cepas PB12 y PB13.
3. Establecer el efecto de esas mutaciones sobre el crecimiento y el perfil de transcripción de las cepas PB12 y PB13..

VI. Resultados y discusión:

A. Cambios genéticos clave en la recuperación del crecimiento en glucosa en las cepas PTS:

1. Determinación del genoma de la cepa JM101 y PB11:

Debido a que no se contaba con la secuencia del genoma de la cepa silvestre JM101; para la determinación de las mutaciones en el genoma de las cepas PB12 y PB13, fue necesario secuenciar primero el genoma de las cepas PB11 y JM101. La determinación del genoma de la cepa PB11 se realizó con el objetivo de determinar si todos los cambios encontrados en las cepas evolucionadas eran producto del proceso de evolución adaptativa, o existían algunos cambios genéticos ocurridos en la cepa PB11 debidos principalmente al manejo y propagación de los inóculos.

La secuenciación de las cepas JM101 y PB11 se llevaron a cabo por dos métodos de secuenciación masiva: utilizando las tecnologías de Ion torrent PGM e Illumina Inc. GAIIx, realizadas por la UUSMD (IBt, UNAM) y Winter Genomics. Inc.

Una vez obtenida la secuencia genómica de la cepa JM101, se comparó con el de la cepa silvestre MG1655 con el objetivo de encontrar diferencias entre ambas cepas. Como primer paso se identificaron las características de la cepa JM101 expresadas en el genotipo ya reportado: [*glnV44*

(*supE44*), *thi*, $\Delta(\textit{lac-proAB})$ / F' (*traD36 proAB+ lacIq lacZ Δ M15)] [119,120,121]; estas características se revisaron con los datos de secuenciación y todas se corroboraron, con excepción de una de ellas: la mutación que le confiere la capacidad de suprimir codones del tipo ámbar fue re-anotada, ya que dicha mutación se identificó en el gen *glnX* y no en el gen *glnV* como se reportó en el artículo original del 1979.*

Adicionalmente, se detectaron varias diferencias (mutaciones puntuales no sinónimas), entre los genomas de la JM101 y la cepa MG1655, una de ellas lo bastante relevante como para incorporarla en la descripción del genoma. Esta mutación se detectó en la posición 98 del gen *rpoS* (Tabla S1), la cual ocasiona el cambio de un residuo de glutamina por un codón de término (CAG – UAG). Como consecuencia de esta mutación se produce una proteína incompleta de 32 aminoácidos en lugar de la proteína normal que cuenta con 330 residuos; sin embargo, como se mencionó anteriormente, la cepa JM101 tiene la capacidad de suprimir codones de término (mutaciones ámbar) originados por errores en la replicación [122]. Por lo anterior, es probable que se genere una menor cantidad de RpoS funcional, pero lo suficiente para activar genes altamente dependientes de este factor sigma [117]. Previo a la secuenciación del genoma de las cepas de este estudio, la mutación ya se había detectado en el genoma de las cepas PB12 y PB13, motivo por el que se pensaba que estaba presente únicamente en estas cepas y que se había originado durante el proceso de evolución [117]. Es importante señalar que existen varios reportes donde describen la aparición de mutaciones que ocasionan en última instancia la disminución en las concentraciones intracelulares de este factor sigma. Dichas mutaciones son ocasionadas por causas diversas (situaciones de estrés, métodos de almacenamiento y manejo de inóculos), tanto en cepas silvestres como en cepas de laboratorio [123,124,125,126,127,128]. Por tal motivo y debido a que en el artículo original, en donde se describe el genotipo de la cepa [119], no se hace referencia a esta característica; debe tomarse con precaución su incorporación al genotipo de la cepa JM101, ya que dicha mutación pudo haberse generado durante la propagación a partir de la cepa original generada por Messing. A continuación se enumeran las características genotípicas de la cepa JM101 reportadas, corregidas (*supE*) ó añadidas (*rpoS*):

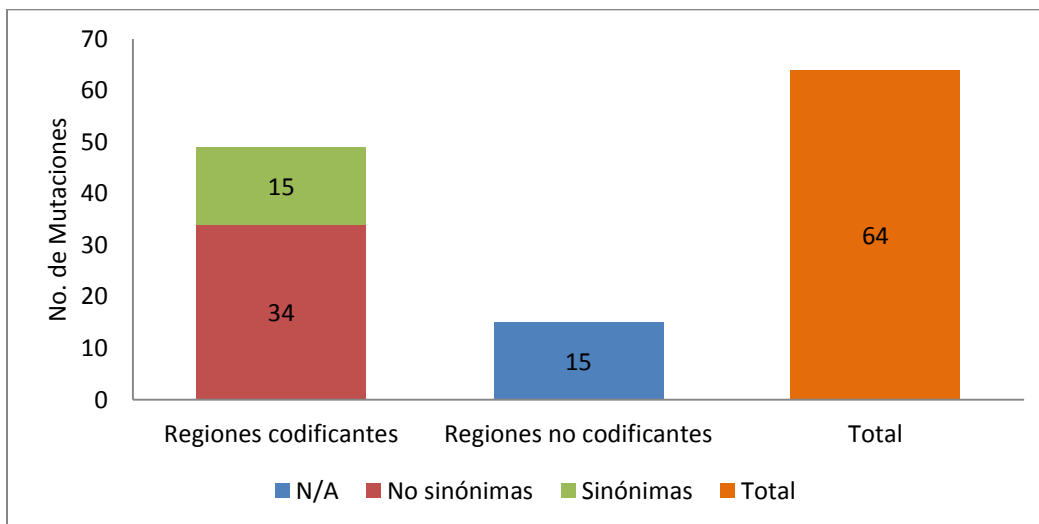
- I. *glnX35* (*supE*): Mutación en el nucleótido 35 del gen que codifica para el tRNA_{glu} que ocasiona la supresión de codones ámbar (UAG).
 - II. $\Delta(\textit{lac-proAB})$: Carece del fragmento cromosomal *lac-proAB*.
 - III. *rpoS*(33am): Mutación puntual en el codón 33 que cambia un codón CAG (glutamina) a un codón de término UAG (ámbar).
- F': Posee el plásmido F.

- IV. *traD36*: Mutación en el gen *traD* que ocasiona la eliminación del factor de transferencia (ayuda a transferir el plásmido F a otra célula).
- V. *proAB*⁺: Contiene el operón *proAB* (síntesis de prolina).
- VI. *lacI*_q: Sobreproducción del represor del operón de lactosa (LacI); la secuencia promotora -35 tiene una mutación puntual que lo ocasiona: GCGCAA-----GTGCAA.
- VII. *lacZ*ΔM15: Delección parcial del gen *lacZ*, que permite una α complementación de la proteína β-galactosidasa (se eliminó del aa 11 al 41).

Por lo tanto, la nueva descripción del genotipo de la cepa JM101 es la siguiente:

glnX35(supE35), thi, Δ(lac-proAB), rpoS(33am) / F' (traD36 proAB+ lacI_q lacZΔM15).

Adicionalmente, como resultado de la secuenciación se detectaron 64 mutaciones puntuales en la cepa JM101 al comparar su secuencia contra la cepa silvestre MG1655, incluyendo las características arriba mencionadas. De las 64 mutaciones puntuales, 49 de ellas se dieron en regiones codificantes, 15 de ellas fueron sinónimas (Tabla S2) y 34 no-sinónimas (Tabla S1; Gráfica 1). Las 15 restantes se originaron en regiones no codificantes (Gráfica 1; Tabla S3), entre ellas la mutación puntual en la región promotora del gen *lacI*. Cabe señalar que las variaciones reportadas en este trabajo en la cepa JM101, se identificaron por ambos métodos de secuenciación, por tal motivo se decidió no corroborarlas mediante secuenciación por Sanger, a excepción de la mutación en el gen *rpoS*, la cual se re-secuenció por este método.



Gráfica 1.- Mutaciones puntuales encontradas en la cepa JM101 al compararla con la cepa silvestre MG1655. N/A: Regiones no codificantes.

2. Determinación de las mutaciones en la cepa PB12 y PB13:

La secuenciación de la cepa PB12 se realizó mediante dos métodos de secuenciación distintos: secuenciación comparativa (CGS), llevada a cabo por la compañía Roche NimbleGen Inc. (RN); y mediante secuenciación masiva, utilizando la tecnología de Illumina Inc. GAlx, con la cual se identificaron 23 mutaciones puntuales no sinónimas y 16 mutaciones puntuales sinónimas en la cepa PB12 (Fig. 8; Tabla S4; Tabla S5). Entre los genes que sufrieron mutaciones puntuales no sinónimas se encontraron 7 genes con funciones desconocidas; 12 genes involucrados en el metabolismo y transporte y 4 genes involucrados con funciones de regulación (Fig. 8; Tabla S4).

Con respecto a la cepa PB13, se secuenció su genoma mediante la tecnología de Illumina Inc. utilizando el equipo GAlx, y mediante el equipo PGM Ion Torrent (Life technologies). Ambas secuenciaciones se realizaron en la UUSMD (IBt, UNAM), y el análisis de los resultados los realizó la compañía Winter Genomics Inc. (WG) Gracias a ambos estudios se detectaron 14 mutaciones puntuales no sinónimas y 8 mutaciones puntuales sinónimas (Fig. 8; Tabla S7,S8).

Adicionalmente a las mutaciones puntuales, en ambas cepas se detectó la ausencia de un fragmento cromosomal de alrededor de 10 kb's (Fig. 8; Tabla S4 y S7); la delección del fragmento cromosomal, y dos mutaciones puntuales no sinónimas, ocurridas en los genes *glpT* y *dhaM* son cambios genéticos que ambas cepas evolucionadas comparten (Fig. 8; Tabla S4 y S7).

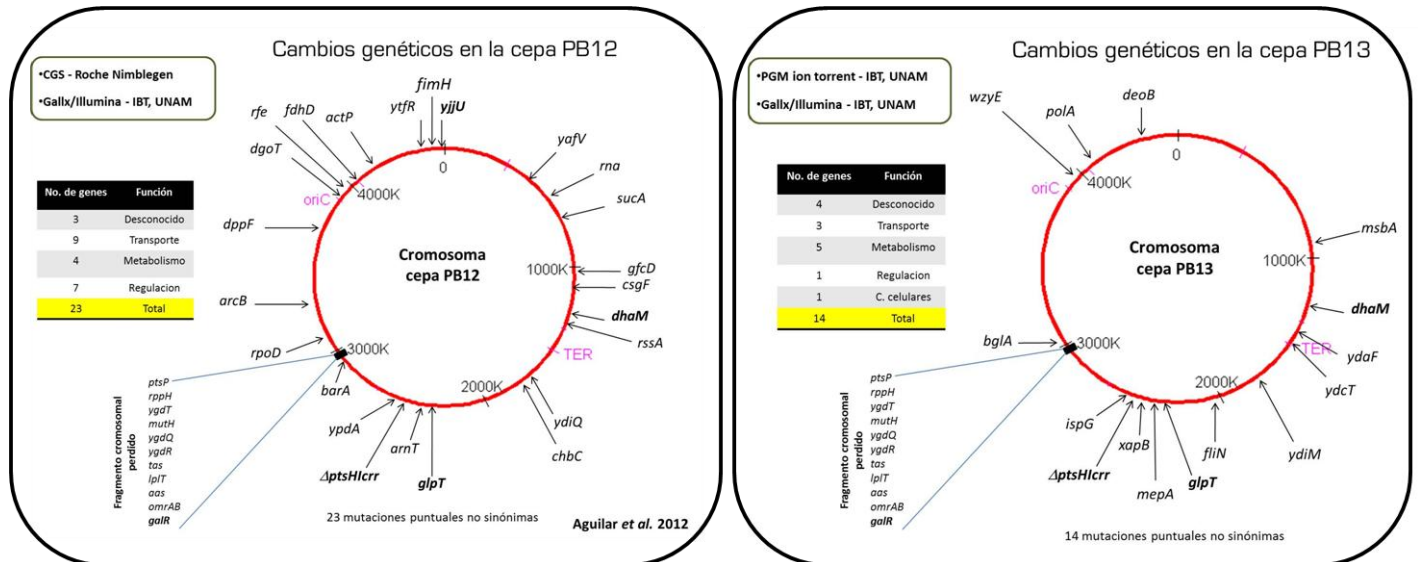


Fig. 8.- Esquema de los cambios genéticos en el cromosoma de las cepas PB12 y PB13, donde se muestra la posición estimada de los genes donde se originaron las mutaciones puntuales no sinónimas así como de la delección del fragmento cromosomal. En negritas se muestran las mutaciones compartidas en ambos genomas.

3. Fragmento cromosomal ausente:

Además de las mutaciones puntuales, se detectó la ausencia de un fragmento cromosomal de 10,328 pb en el genoma de las cepas evolucionadas, en el que residían 12 genes (*rppH*, *ygdT*, *muth*, *ygdQ*, *ygdR*, *tas*, *lplT*, *aas*, *omrA*, *omrB*, *ptsP* and *galR*) (Fig. 9 Tabla S4 y S7).

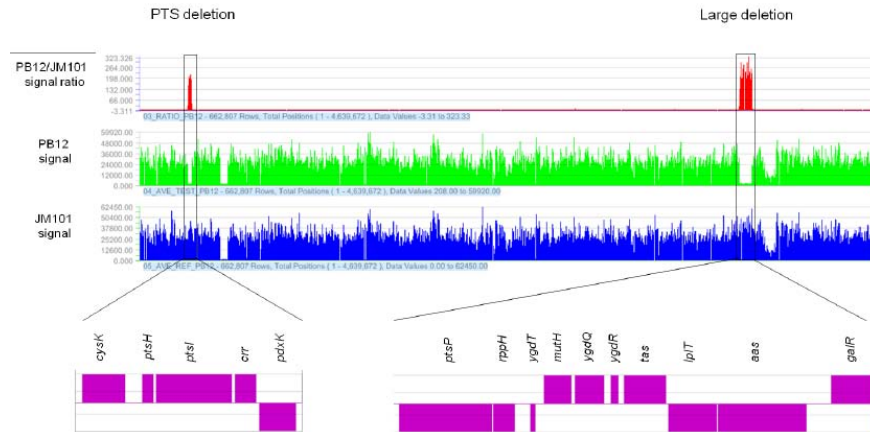


Fig. 9.- Mapa comparativo de las cepas JM101 y PB12 realizado por la empresa Nimblegen. Inc. donde se muestran dos deleciones en la cepa PB12; una correspondiente al operón PTS, y otra a una región del cromosoma ubicada en el minuto 64. Ambos fragmentos también se encontraron ausentes en el cromosoma de la cepa PB13.

Mediante PCR se confirmó la ausencia del fragmento (Fig. 10) y fue posible establecer la nueva organización del cromosoma, que ocasionó la fusión de los genes *ptsP* y *galR*. La secuencia nucleotídica de la región cromosomal en las cepas PB12 y PB13 se muestra en la Fig. S1.

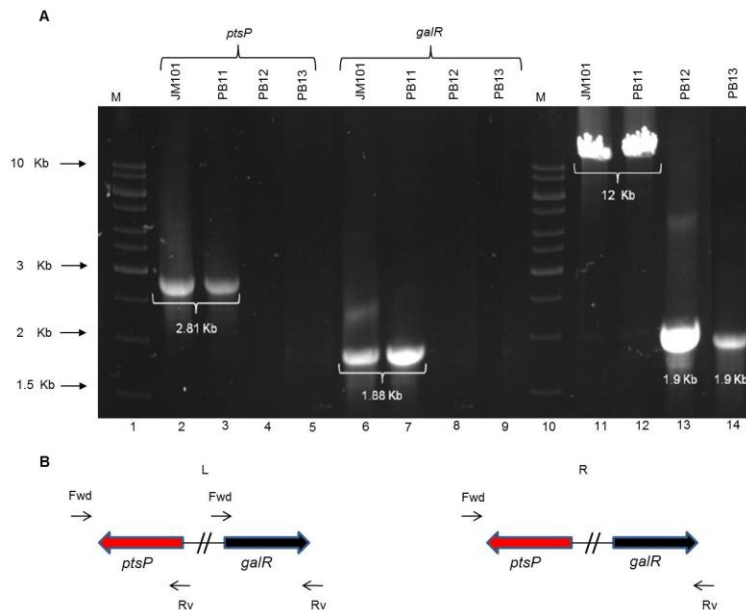


Fig. 10.- Comprobaciones mediante PCR de la región del cromosoma faltante en la cepa PB12. **A.** Se amplificaron los genes *ptsP* y *galR* en las cepas JM101, PB11, PB12 y PB13 carril 1 marcador, carril 2, 3, 4, y 5 amplificación del gen *ptsP* en las cepas JM101, PB11, PB12 y PB13 respectivamente. No se observó amplificación para la cepa PB12 y PB13. Carriles 6, 7, 8 y 9 amplificación del gen *galR* en las cepas

JM101, PB11, PB12 y PB13 respectivamente. No se observó amplificación para la cepa PB12 y PB13. Carril 10, marcador, carril 11, 12 y 13 y 14 amplificación de la región utilizando los oligos ptsP-fwd y galR-rv en las cepas JM101, PB11, PB12 y PB13 respectivamente. **B.** Esquema de los genes que flanquean la región cromosomal y los oligos utilizados para las amplificaciones.

4. *Análisis bioinformático de la región cromosomal deletada:*

Los elementos genéticos denominados secuencias de inserción (IS), son pequeñas secuencias de DNA que contienen información genética relacionada con su transposición y su regulación [129]. La inserción de un elemento IS en ocasiones llega a generar inactivación génica y fuertes efectos polares [130,131,132]. Adicionalmente, los elementos IS actúan como sustrato para las vías de recombinación, ocasionando rearrreglos cromosomales tales como inversiones y deleciones, y se consideran de gran importancia en la evolución bacteriana, dada la alta tasa y diversidad de mutaciones que se generan [133,134]. Con notables excepciones; la mayoría de las IS's exhiben secuencias invertidas repetidas (IR) de entre 10 y 40 pb. Otra característica principal de los elementos IS's, es que en la inserción, la mayoría generan secuencias cortas directas repetidas (DR) en el DNA blanco. La longitud de las DR es entre 2 y 14 pb, y es característica de cada elemento de inserción.

Para averiguar si este fenómeno de adaptación ocurrido en las cepas evolucionadas donde se perdió un fragmento cromosomal se debió a la presencia de secuencias IS en regiones cercanas al fragmento perdido, se realizó un análisis bioinformático de dichas regiones que abarcó 3 kb de longitud hacia los extremos de la región en cuestión (17.4 Kb de longitud); con el objetivo de encontrar posibles secuencias IS (en la cepa silvestre), o remanentes de estas secuencias (en la cepa PB12 y PB13). Para esta tarea se utilizó el software de uso libre "IS Finder" (www-is.biotoul.fr) el cual analiza y localiza IS's, IR's y DR's de la secuencia problema; identificando para cada caso, de que elemento móvil se trata (existen cientos de elementos móviles, distribuidos en más de 20 familias). Durante la búsqueda no se encontró ninguna coincidencia significativa para ningún elemento en ambas secuencias.

5. *Pérdida del fragmento cromosomal en otras cepas aisladas durante el proceso de evolución:*

Como ya se mencionó, a partir de la cepa PB11 y por un proceso de evolución adaptativa se aislaron varias cepas capaces de crecer más rápidamente utilizando la glucosa como única fuente de carbono. Se sabe poco de la fisiología de las otras cepas aisladas durante el proceso de evolución y hasta ahora no se han llevado a cabo estudios sobre las mutaciones que se generaron en las cepas aisladas restantes. Con el objetivo de conocer más acerca de estas cepas, se procedió a determinar mediante comprobaciones por PCR y por análisis de transcripción mediante RT-qPCR, la presencia

del fragmento cromosomal en 4 cepas, las cuales se aislaron en las diluciones 0.4 (1), 0.6 (2) y 0.8 (1) hr⁻¹ denominadas: 10Glc, 32Glc, 40Glc y 99Glc respectivamente [107,117].

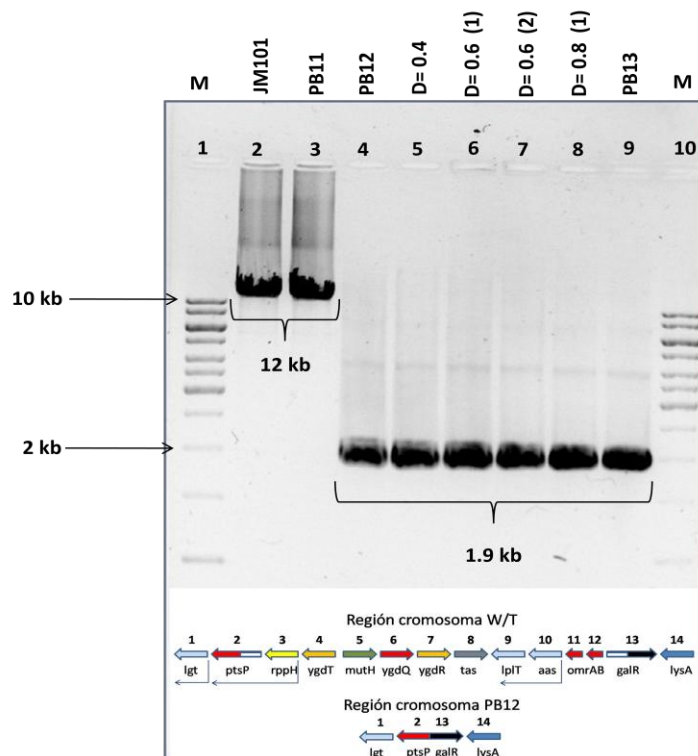


Fig. 11.- Comprobaciones mediante PCR de la presencia de la región del cromosoma faltante en varias cepas evolucionadas aisladas. El fragmento cromosomal consta de 10,328 pb donde residían 12 genes; dos de estos genes (*ptsP* y *galR*), se fusionaron por su parte media (esquema debajo del gel). Cepas JM101 y PB11 (carril 2 y 3), PB12 (carril 4), 10Glc (carril 5), 32Glc (carril 6), 40Glc (carril 7), 99Glc (carril 8), PB13 (carril 9).

Los resultados revelaron que todas las cepas aisladas a partir de la hora 120 (incluida la cepa PB12) y hasta la hora 200 (incluida la cepa PB13), carecen de este fragmento cromosomal (Fig. 11), sugiriendo que este fenómeno de delección se presentó temprano en la evolución y se fijó como una mutación benéfica para un mejor crecimiento. Adicionalmente, se secuenciaron los genes *glpT* y *dhaM* en estas cepas mediante el método de Sanger, donde se observó que las mutaciones puntuales en ambos genes también se encuentran presentes en todas las cepas aisladas.

6. Divergencia evolutiva entre las cepas PB12 y PB13:

Los resultados de la secuenciación de las cepas evolucionadas PB12 y PB13, así como la búsqueda de la región cromosomal en las cepas aisladas en etapas intermedias, sugiere que uno de los eventos tempranos durante el cultivo fue la escisión del fragmento cromosomal, lo cual podría generar ventajas adaptativas a la célula. Es difícil determinar si las mutaciones puntuales en los genes *glpT* y *dhaM* surgieron antes o después de la pérdida del fragmento cromosomal, sin embargo

es probable que al menos la mutación en el gen *glpT* haya surgido después, pues como se discutirá más adelante, no se descarta que esta mutación pueda generar ventajas para contender con los efectos ocasionados por la ausencia del ciclo de regeneración de fosfatidiletanolamina. De este modo, es posible establecer a la pérdida del fragmento cromosomal como el primer evento durante el proceso de evolución, seguido de las mutaciones puntuales en los genes *glpT* y *dhaM*. A partir de ese momento, las cepas PB12 y PB13 tomaron caminos evolutivos diferentes, acumulando diversas mutaciones hasta el momento de su aislamiento a las 120 y 200 horas respectivamente (Fig. 12).

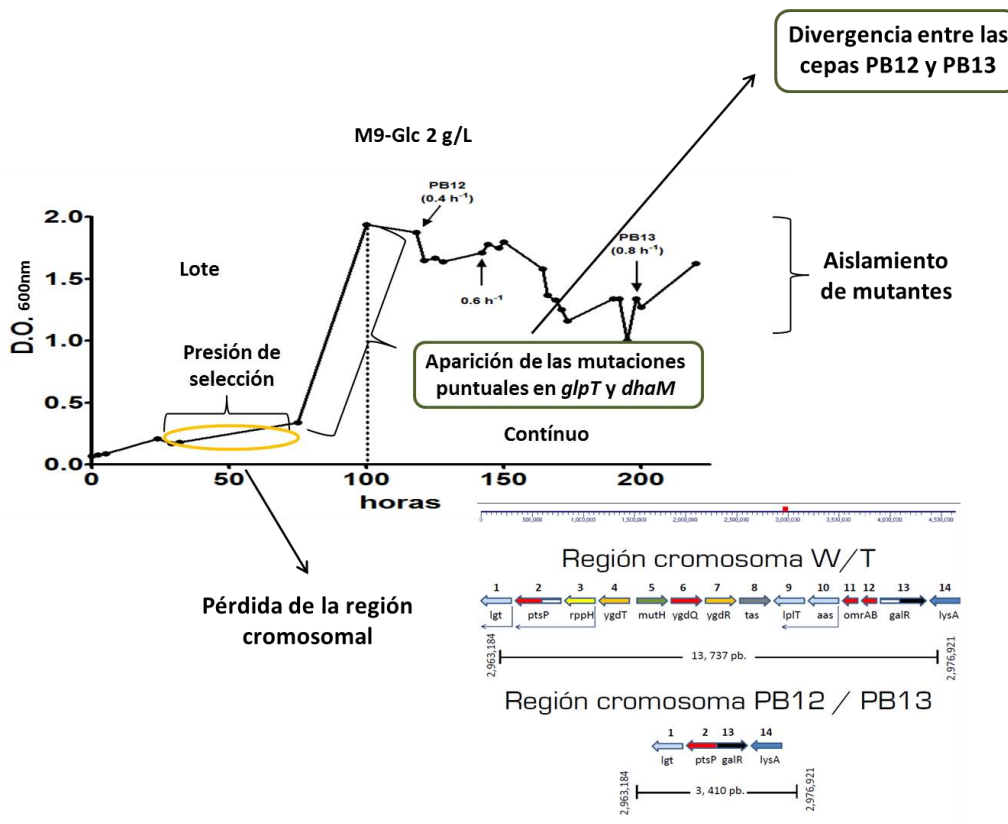
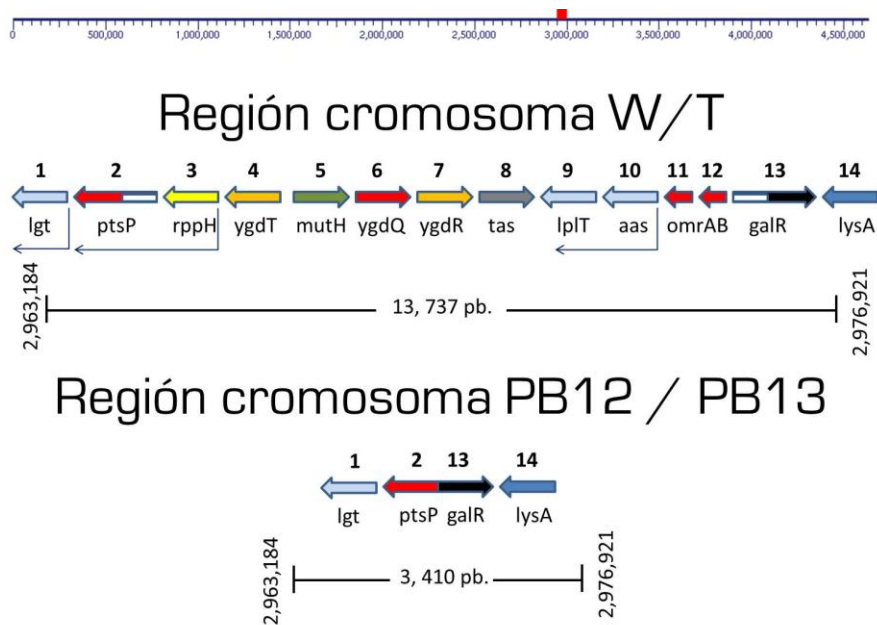


Fig. 12.- Esquema de la divergencia evolutiva entre las cepas PB12 y PB13

7. Genes ausentes en la cepa PB12 y PB13:

Con la pérdida del fragmento cromosomal, se eliminaron 12 genes, 10 de ellos en su totalidad, mientras que la de los genes flanqueantes se perdió únicamente la primera mitad, generando la fusión de éstos (*ptsP* y *galR*) por su parte media, eliminando en consecuencia las regiones promotoras de ambos genes (Fig. 13). Entre los genes que se consideraron más importantes, debido a su función y potencial efecto sobre el crecimiento fueron: *rppH*, *lpIT*, *aas* y *galR*.



Aguilar et al. 2012

Fig. 13.- Eliminación de un fragmento cromosomal de 10,328 pb alrededor del minuto 64 (3.0 millones de pb) en la cepa PB12 y PB13 durante el proceso de evolución adaptativa. En el nuevo arreglo se observa la fusión de los genes *ptsP* y *galR* por su parte media; ambos genes carecen de la primera mitad incluidas sus secuencias promotoras.

a) RppH:

El gen *rppH*, consta de 531 pb y codifica para una pirofosfohidrolasa de RNA. Esta proteína es la primera en la vía de degradación de los mRNA's, y se encarga de hidrolizar el pirofosfato del extremo 5' de los mRNA's, lo que genera mRNA's monofosfatados, que son un mejor sustrato para la RNasa E, la cual se encarga de continuar con la degradación de los mismos. La remoción del pirofosfato por tanto, es el paso limitante para la degradación [135]. Una débil unión por parte del ribosoma acelera el decaimiento de los mensajeros, independientemente de la actividad de RppH, lo que muestra la existencia de otros factores implicados en la estabilidad de los mensajeros que determina su vida media. El proceso llevado a cabo por esta pirofosfohidrolasa es similar al "decapping" en células eucariontes [28].

b) Operón *aas-lplT*:

El gen *lplT* consta de 1,194 pb y codifica para una proteína miembro de la superfamilia de facilitadores principales. LplT está embebida en la membrana interna, y actúa como transportador de lisofosfatidiletanolamina (liso-PE) desde el espacio periplásmico al citoplasma, proveyéndole sustrato a la enzima bifuncional Aas. El gen *aas* consta de 2,160 pb y codifica para la enzima Aas (2-acil-glicerofosfatidiletanolamina acil-transferasa / acil-ACP sintasa), la cual regenera

fosfatidiletanolamina (PE) a partir del liso-PE transportado por LpIT [136]. En *E. coli*, la PE constituye alrededor del 75-80%, del total de fosfolípidos de membrana, mientras que un 15-20% lo conforma el fosfatidilglicerol (PG), y un 6% corresponde a la cardiolipina (CL) [137]. Aas y LpIT son parte del ciclo de la 2-acilglicerofosfatidiletanolamina (ciclo de la liso-PE), en el cual se regenera constantemente este la PE, lo cual evita su biosíntesis durante el mantenimiento celular (Fig. 14). Ambos genes se encuentran juntos y conforman el operón *aas-lpIT* [136].

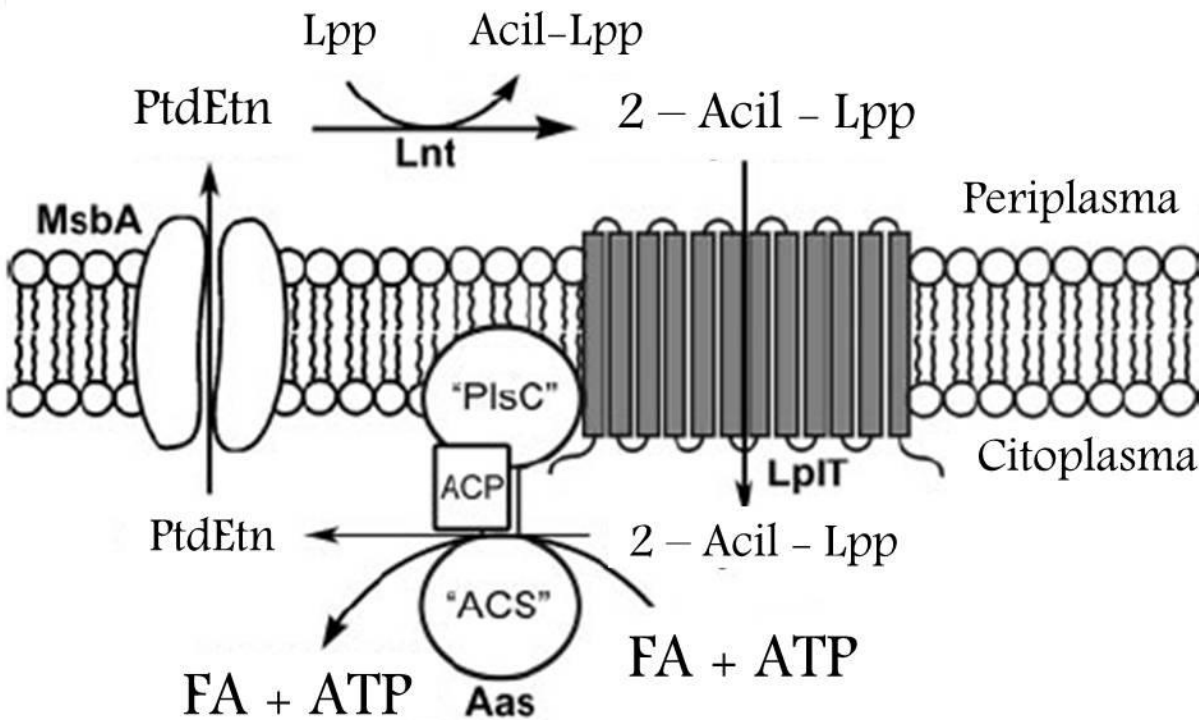


Fig. 14.- Ciclo de la 2-acilglicerofosfatidiletanolamina (Modificado de Harvat et al. 2005) [136]. Abreviaturas: Lpp, lipoproteína; acil-Lpp, liso-lipoproteína PtdEtn, fosfatidiletanolamina; 2-acil-GPE, liso-fosfatidiletanolamina; FA, ácidos grasos; ATP, adenosin trifosfato; AMP, adenosin monofosfato; PPI, pirofosfato inorgánico

c) GalR:

El gen *galR* consta de 1,032 pb, y codifica para el factor de transcripción GalR, que reprime la expresión de los operones involucrados en el transporte y el catabolismo de la D-Galactosa. La inactivación de este gen en *E. coli* incrementa los niveles de transcrito de los genes del regulón aún en ausencia del inductor (galactosa). No obstante, análisis de RT-qPCR realizados en las cepas PB11 y PB12 [8] demuestran que la mayoría de los genes involucrados con el regulón Gal se encuentran sobreexpresados en ambas cepas, entre ellos el gen *galP*, que codifica para la permeasa de galactosa, el cual además de transportar galactosa, es capaz de transportar glucosa, sobre todo en ausencia del sistema PTS.

B. Mecanismos adaptativos en cepas PTS⁻ creciendo en Glucosa como única fuente de carbono

1. Contribución de los genes ausentes en la recuperación del crecimiento:

Con el objetivo de establecer la contribución de la ausencia de algunos genes residentes en el fragmento cromosomal sobre el crecimiento, se determinó inactivar en la cepa PB11, los genes *rppH* y el operón *aas-lpIT*. Estos genes se eligieron de acuerdo a su función y el posible impacto sobre el crecimiento, tomando en cuenta también las características presentes en la cepa PB12. Cabe señalar que debido a que la cepa PB11 ya presenta las características observadas en cepas carentes del gen *galR*, no se generó esta mutante. De este modo se construyeron cepas derivadas de PB11 carentes del operón *aas-lpIT*, involucrado en el recambio de la fosfatidiletanolamina, y del gen *rppH*, implicado en la ruta de degradación de los mRNA, además de una doble mutante, carente tanto del operón *aas-lpIT* como del gen *rppH*, así como de una cepa PB11 carente de los 12 genes del fragmento cromosomal. El crecimiento de las cepas arriba descritas y a lo largo de este trabajo se realizó en matraces bafleados de 125 mL, con un volumen de trabajo de 25 mL de medio mínimo M9 con glucosa (2 g/L), a 37°C y 300 rpm.

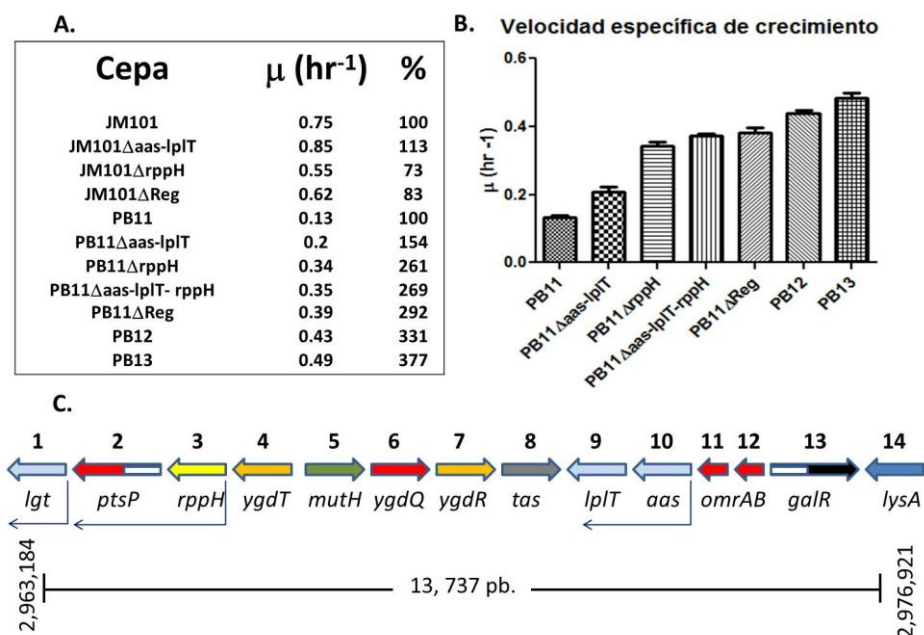


Fig. 15.- **A.** Velocidades específicas de crecimiento en cepas derivadas de la JM101 y PB11 generadas en éste estudio. **B.** Gráfica comparativa de la recuperación del crecimiento en las diferentes cepas derivadas de la PB11. **C.** Esquema de los genes eliminados en las cepas evolucionadas después del proceso de evolución adaptativa.

Como se puede apreciar en la Fig. 15, la inactivación del operón *aas-lpIT* incrementó la μ de la cepa PB11 de 0.13 a 0.2 hr⁻¹, lo que representa un incremento del 54%, mientras que la inactivación del gen *rppH* incrementó el crecimiento hasta $\mu=0.34$ hr⁻¹, lo que significa un incremento de 261%. A partir de la cepa PB11 Δ *aas-lpIT*, se procedió a inactivar el gen *rppH*, para producir la cepa PB11 Δ *aas-lpIT,rppH*; la cual presentó una μ similar a la de la cepa PB11 Δ *rppH* (Fig. 15). Lo anterior indica que la ausencia del gen *rppH* representa un mayor incremento en el “fitness” comparado con la ausencia del operón *aas-lpIT*. Por su parte, la generación de la cepa doble mutante ocasionó efectos epistáticos. Con el objetivo de determinar si existía algún incremento adicional en “fitness” debido a la ausencia de otro gen dentro de la región cromosomal en cuestión, se generó una cepa PB11 que al igual que las cepas evolucionadas carece de los 12 genes antes mencionados, denominada PB11 Δ Reg. Los resultados de la caracterización de la cepa PB11 Δ Reg muestran una μ de 0.39 hr⁻¹ lo cual representa un incremento del 292% con respecto a la cepa original PB11, hecho que sugiere un beneficio mínimo adicional sobre el crecimiento, además de aquel aportado por la ausencia de los genes *aas*, *lpIT* y *rppH*. Existe la posibilidad de que ese incremento adicional se deba a la ausencia del gen *galR*; no obstante las características esperadas para una cepa *galR*⁻ como lo es la sobreexpresión de los genes involucrados con el metabolismo de la galactosa, incluido el gen *galP* que codifica para la permeasa de galactosa, ya están presentes en la cepa PB11 [8]. Debido a lo anterior, es más plausible atribuirle este incremento, a los beneficios otorgados por la ausencia de uno o varios genes (independientemente de un probable efecto epistático); no obstante, de acuerdo a la función de las proteínas codificadas por los genes restantes, resulta complicado establecer cuales estarían otorgándole un beneficio en cuanto a crecimiento.

Otro aspecto relevante que se desprende de los resultados anteriores es que el incremento observado en la cepa PB11 Δ Reg, representa una recuperación del 91 y del 80 % con respecto al crecimiento de las cepas PB12 y PB13 respectivamente, lo que significa que las mutaciones puntuales en el genoma de las cepas evolucionadas aportan entre un 10% (en la cepa PB12) y un 20% (en la cepa PB13) en la recuperación del crecimiento.

Por otra parte resulta interesante el efecto pleiotrópico ocasionado por la inactivación de los genes *aas*, *lpIT* y *rppH* en la cepa silvestre. Por un lado, con la inactivación del operón *aas-lpIT* en la cepa JM101 se observa un ligero incremento en el crecimiento con una μ de 0.85 hr⁻¹, lo que significa un incremento del 13% con respecto al control (cepa JM101), lo que resulta en un efecto similar al observado para el caso de la cepa PB11. No obstante, la inactivación del gen *rppH* en esta misma cepa, ocasionó una disminución del 27% en su crecimiento ($\mu=0.55$ hr⁻¹). Tomando en cuenta la función llevada a cabo por la enzima RppH, es probable que su ausencia en la cepa JM101 ocasione

algún tipo de desequilibrio en los niveles de transcrito y por consiguiente en la fisiología celular, lo que ocasionando la disminución del crecimiento observada en esta cepa.

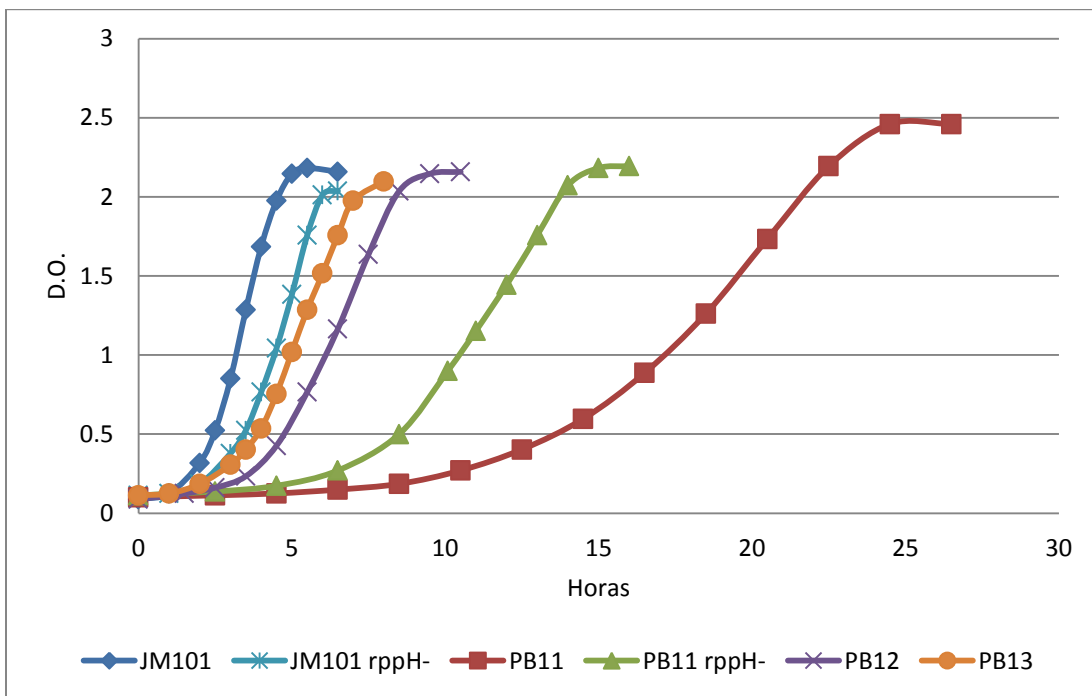
Una vez evaluado el crecimiento de todas las cepas generadas, se seleccionó a las mutantes en el gen *rppH* y sus controles (cepas JM101, JM101 Δ *rppH*, PB11, PB11 Δ *rppH* y PB12) para una caracterización a nivel transcriptómico principalmente, que generara un conocimiento más detallado y fuera la pauta para explicar el crecimiento diferencial presentado en los distintos fondos genéticos; lo anterior, debido al gran impacto de la ausencia del gen sobre el crecimiento, aunado a su función en la degradación de mRNA's y los altos niveles de transcrito observados en la cepa PB12. De esta forma, se realizaron análisis para la determinación del consumo de glucosa, caracterización del crecimiento en sustratos como acetato y glucosa - acetato, análisis transcriptómico de genes del metabolismo central, así como un análisis proteómico global; este último análisis, no se presenta en este escrito ya que aún no ha concluido.

2. Caracterización de cepas Δ rppH en medio mínimo M9 con glucosa como única fuente de carbono:

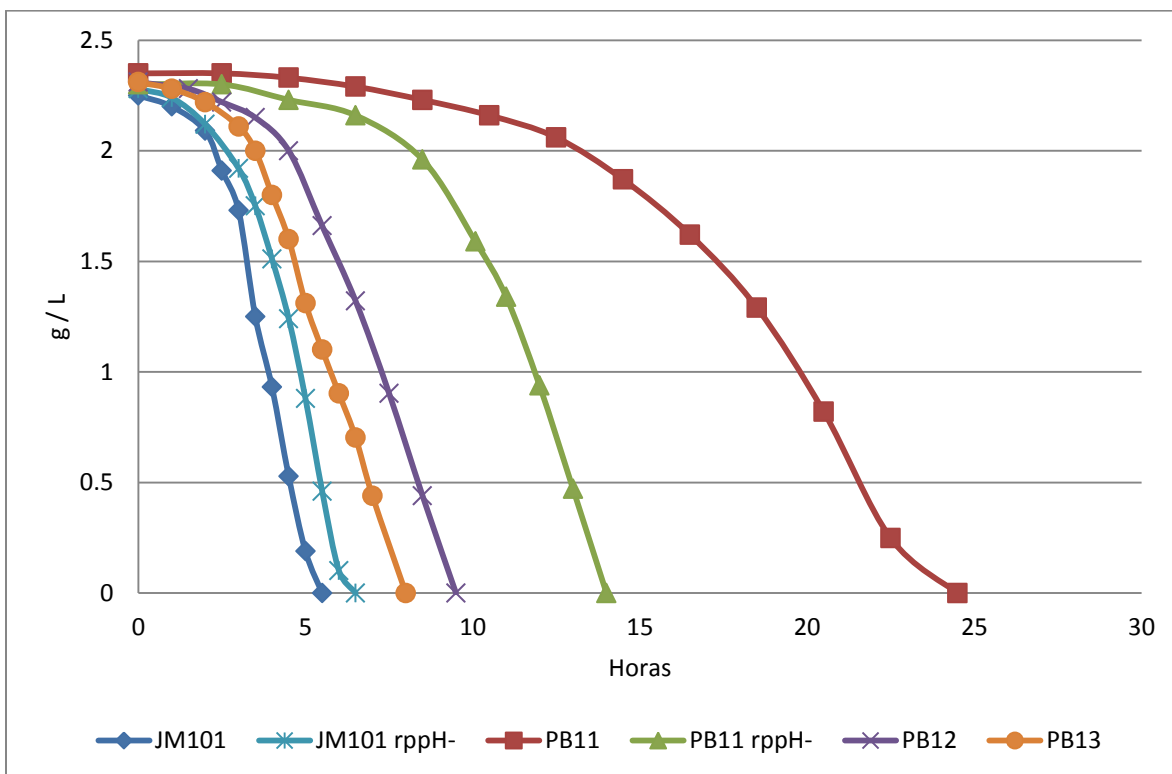
Para la caracterización de las cepas JM101 y PB11 inactivas en el gen *rppH* se procedió a la determinación de otros parámetros cinéticos (rendimiento biomasa-sustrato y velocidad específica de consumo de sustrato) que nos brindaran un mejor entendimiento de los alcances adaptativos de la ausencia del gen *rppH*. Para ello se realizaron cinéticas de crecimiento en glucosa en ambas cepas carentes del gen *rppH*, así como en las cepas parentales JM101 y PB11 y las evolucionadas PB12 y PB13. Los resultados se muestran en la tabla 1 y en las gráficas 2 y 3.

Cepas	Y x/s (g DCW / g Glc)	qs (g Glc / g DCW h)	qs (%)
JM101	0.36 ± 0.01	2.05 ± 0.12	100
JM101 Δ <i>rppH</i>	0.35 ± 0.01	1.7 ± 0.10	83
PB11	0.37 ± 0.00	0.48 ± 0.00	23
PB11 Δ <i>rppH</i>	0.39 ± 0.01	0.81 ± 0.03	39.5
PB12	0.34 ± 0.02	1.25 ± 0.02	61
PB13	0.32 ± 0.03	1.55 ± 0.09	75.6

Tabla. 1.- Parámetros cinéticos de las cepas incluidas en éste estudio evaluados en medio mínimo M9 con 2 g/L de glucosa.



Gráfica 2: Cinéticas de crecimiento en medio mínimo M9-Glc 2 g/L de diferentes cepas de éste estudio.



Gráfica 3: Consumo de glucosa en diferentes cepas de éste estudio utilizando medio mínimo M9-Glc 2 g/L.

Los valores de velocidad específica de consumo de sustrato (q_s) obtenidos nos indican una afectación de cerca del 80% en el consumo de glucosa cuando se inactiva el sistema PTS (cepa

PB11). Este efecto negativo disminuye con la inactivación del gen *rppH*, lo cual prácticamente duplica la qs. No obstante, este incremento no logra igualar las capacidades de las cepas evolucionadas PB12 y PB13, las cuales tienen una qs de 1.25 y 1.55 g Glc/g DCW h, lo cual representa un 61 y 75% respectivamente con respecto a la qs de la cepa silvestre JM101. Por otro lado, la inactivación del gen *rppH* en la cepa JM101 disminuye el consumo de glucosa cerca del 20%. Las diversas capacidades de consumo de sustrato de las diferentes cepas se hacen evidentes en la gráfica 3. Con lo anterior, es posible establecer que la capacidad de asimilación de glucosa que se perdió por la inactivación del sistema PTS, se recupera parcialmente por la ausencia del gen *rppH*, probablemente debido a que se favorece la producción de enzimas implicadas en dichas funciones.

3. Crecimiento de las cepas $\Delta rppH$ en acetato y en la mezcla glucosa - acetato:

Las cepas PTS⁻ son capaces de co-utilizar 2 o más sustratos aún en presencia de glucosa, debido a que carecen de represión catabólica [6,8,111]. Biotecnológicamente esta capacidad resulta favorable, ya que para la producción de diversos metabolitos o proteína recombinante es deseable que se consuman los sustratos rápida y eficientemente, lo que resulta en bajos costos en términos de productividad. En el caso particular de la cepa PB11, a pesar de ser capaz de co-utilizar glucosa y acetato presentes en el medio, se observa una marcada preferencia por consumir primero el acetato [8,112], debido a que carece de los medios para asimilar la glucosa (ausencia del sistema PTS). En el caso de las cepas PB12 y PB13, resulta interesante que sus características sean opuestas a las de la cepa PB11; ya que co-utilizan la glucosa y el acetato presentes, pero se observa una marcada preferencia por la glucosa [8,112], lo que pone de manifiesto el resultado del proceso de adaptación. El hecho de que la ausencia del gen *rppH* en la cepa PB11 resulte en una sustancial ganancia en el crecimiento, debido a un mejor catabolismo de la glucosa, probablemente también sea la causa de una disminución en su preferencia para asimilar el acetato cuando ambas fuentes de carbono están juntas en el medio de cultivo. Por tal motivo se procedió a determinar la μ de las cepas $\Delta rppH$ en acetato y glucosa-acetato como únicas fuentes de carbono. Los resultados se muestran en la Tabla 2.

Cepa	μ (hr ⁻¹)	%
Glucosa		
JM101	0.75	100
JM101 $\Delta rppH$	0.55	73
PB11	0.13	100
PB11 $\Delta rppH$	0.34	261
PB12	0.43	331
PB13	0.49	377
Acetato 3 g/L		
JM101	0.29	100
JM101 $\Delta rppH$	0.22	76
PB11	0.22	76
PB11 $\Delta rppH$	0.17	59
PB12	0.13	45
PB13	0.1	34.5
Glucosa 2 g/L y Acetato 1 g/L		
PB11	0.3	100
PB11 $\Delta rppH$	0.26	87
PB12	0.46	153
PB13	0.51	170

Tabla 2.- Velocidades específicas de crecimiento de diferentes cepas generadas en este estudio utilizando medio mínimo M9 y diferentes sustratos.

Como es posible observar, la inactivación del gen *rppH* en la cepa PB11 ocasiona una disminución superior al 20 % en el crecimiento cuando se utiliza acetato como única fuente de carbono. Lo mismo ocurre con las cepas evolucionadas, las cuales presentan un crecimiento muy disminuido ($\mu=0.13$ hr⁻¹ para la cepa PB12 y $\mu=0.1$ hr⁻¹ para la cepa PB13). Por otra parte, estos resultados nos indican que la ausencia del gen *rppH* es una mutación adaptativa, específica para crecer en glucosa, la cual tiene efectos pleiotrópicos. Es interesante observar que la eliminación de este gen en la cepa silvestre JM101, afecte el crecimiento de forma negativa cuando se utiliza glucosa o acetato como únicas fuentes de carbono (Tabla 2), lo cual sugiere la existencia de carga metabólica.

Como se verá más adelante en este documento, las disminuciones en cuanto a crecimiento esencialmente pueden explicarse debido a la presencia de carga metabólica producto de la ausencia de RppH. La carga metabólica se define como la cantidad de recursos que utiliza la célula para el mantenimiento, la replicación, la transcripción y la traducción de la información genética contenida en vectores sintéticos principalmente [138], donde se favorece la sobreproducción de ciertas proteínas mediante la sobreexpresión de genes específicos.

Entre los efectos de la carga metabólica se encuentran diversas alteraciones metabólicas y fisiológicas, que ocasionan entre otras cosas: disminuciones en las velocidades específicas de crecimiento, reducción en el rendimiento biomasa-sustrato, alteraciones del ciclo celular, activación de respuestas a diferentes tipos de estrés etc [139,140,141]. La eliminación del gen *rppH*, simula lo que ocurre cuando se expresan genes en plásmido (grandes concentraciones de mRNA que serán traducidos) lo que genera que una gran cantidad de recursos sean utilizados para la producción de proteína; aunque con la ausencia del gen *rppH*, este efecto se da de forma masiva (en una gran cantidad de genes). De este modo, las cepas carentes del gen *rppH*, presentan carga metabólica, lo que ocasiona menores velocidades de crecimiento con respecto a sus controles; aunque para el caso de las cepas evolucionadas creciendo en acetato como única fuente de carbono, el pobre crecimiento presentado se explica en función de su capacidad de adaptación a la glucosa producto de los cambios genéticos presentes en su genoma, ya que como se discutirá más adelante, esta cepa logró disminuir esta carga metabólica probablemente por mutaciones específicas. En ese mismo sentido, la utilización de ambas fuentes de carbono en el medio de cultivo, ocasionó una ligera disminución del crecimiento en la cepa PB11 Δ *rppH* (13%), comparada con la cepa PB11 lo que sustenta que la ausencia de RppH genera carga metabólica que va en detrimento del crecimiento.

Las cepas evolucionadas por su parte incrementan ligeramente su crecimiento cuando existen los dos sustratos en el medio ($\mu=0.46 \text{ hr}^{-1}$ en PB12 y $\mu=0.51 \text{ hr}^{-1}$ en PB13), lo que sugiere que la carga metabólica en ambas cepas ha disminuido, probablemente debido a mutaciones específicas. De esta forma, el crecimiento observado en glucosa y en la mezcla Glc-Ace para las cepas evolucionadas se debe a: 1) Ausencia del fragmento cromosomal (*rppH*) como cambio adaptativo para crecer cuando existe glucosa en el medio y 2) disminución de la carga metabólica producto de mutaciones específicas.

Con el objetivo de establecer los efectos de la ausencia del gen *rppH* sobre los transcritos de genes relevantes, así como poder explicar el incremento en las capacidades glicolíticas y por tanto los

beneficios en el crecimiento, se procedió a determinar los niveles de transcrito de las cepas JM101 Δ rppH y PB11 Δ rppH.

4. Determinación de los niveles de transcrito de genes del metabolismo central en cepas Δ rppH:

Para profundizar en el efecto de la ausencia del gen *rppH* sobre el perfil transcripcional, se analizaron los niveles de expresión mediante RT-qPCR de cerca de 100 genes involucrados en el metabolismo central de carbono en las cepas JM101 Δ rppH, PB11 Δ rppH y PB13 los cuales ya se habían determinado en las cepas PB11 y PB12 en medio mínimo con glucosa como única fuente de carbono [8]. Tomando en cuenta la función de la proteína RppH, se esperaba que con su ausencia se incrementara la vida media de los mensajeros y ello se viera reflejado en niveles mayores de transcrito en las cepas Δ rppH debido al diferencial entre la síntesis y la degradación del mRNA. Para el caso de la cepa PB13, era factible esperar valores similares a los ya reportados para la cepa PB12, aunque probablemente con algunas diferencias en genes involucrados en glucólisis que evidencien una mayor capacidad glicolítica.

Los resultados de los análisis se presentan en la Figura 16, y muestran que del total de genes analizados, el 73% en la cepa JM101 Δ rppH y el 92% en la cepa PB11 Δ rppH incrementaron su nivel de transcrito. Estos datos sugieren, por tanto, que la disminución del crecimiento en ambas cepas en medio mínimo suplementado con glucosa, acetato o la mezcla de ambas discutida en la sección anterior, se debe a la elevada carga metabólica que se genera en la célula cuando se modifica un elemento tan importante para la regulación transcripcional, ya que como se discutirá más adelante se llegaron a observar niveles de transcrito 40 veces superiores con respecto a la cepa control para algunos genes. Por su parte la cepa PB13 también tuvo niveles elevados de transcrito en la mayoría de los genes analizados, no obstante, dichos niveles fueron menores al compararlos con los observados en la cepa PB11 Δ rppH, y muy similares a los ya reportados previamente para la cepa PB12, lo que sugiere que alguna o algunas mutaciones están influyendo en disminuir dichos niveles de transcrito y por consiguiente, en aliviar la carga metabólica que se genera por la ausencia de RppH en las cepas evolucionadas.

isocitrato (ICT), glioxilato (GOx), α -cetoglutarato (α -KG), succinil-coenzyma A (SUC-CoA), succinato (SUC), fumarato (FUM), malato (MAL), oxaloacetato (OAA), 6-fosfogluconolactona (6PGNL), 6-fosfogluconato (6PGNT), ribulosa-5-fosfato (RU5P), ribosa-5-fosfato (R5P), xilulosa-5-fosfato (X5P), sedoheptulosa-7-fosfato (S7P), eritrosa-4-fosfato (E4P), 2-ceto-3-desoxi-6-fosfogluconato (KDPGNT), 3-desoxi-D-arabinoheptulosonato-7-fosfato (DAHP).

Es importante puntualizar que los altos niveles de transcrito encontrados en las cepas carentes del gen *rppH*, incluyendo las cepas evolucionadas, no se deben en la mayoría de los casos a una sobreexpresión, sino al incremento en la vida media de los mensajeros, los cuales no se están degradando a la misma velocidad que en la cepa silvestre. Por tal motivo, a pesar de que los niveles en la mayoría de genes reguladores no cambian de manera significativa, es aventurado establecer de forma precisa qué genes o grupos de genes tienen niveles elevados de expresión debido a un incremento en la vida media o a una sobreexpresión como tal; incluida la posibilidad de que para algunos casos los niveles de transcrito observados podrían deberse a ambas causas. Adicionalmente, la ausencia del sistema PTS para el caso de las cepas PB11, PB12 y PB13 es un factor más a considerar, ya que este sistema influye fuertemente sobre la regulación.

A continuación se describen los datos más relevantes de los análisis de expresión, los cuales se agrupan de acuerdo a la vía metabólica de la cual forman parte. Todos los datos se presentan en forma esquemática en la Figura 16, y de forma numérica en la Tabla S10.

a) Glucólisis:

Se observaron niveles de transcrito superiores entre dos y tres veces para varios de los genes de la vía glucolítica cuando se inactivó el gen *rppH* en la cepa silvestre JM101, alcanzando valores de hasta 5 y 7 veces para los genes *tpiA* y *gapA* respectivamente. Por su parte, la cepa PB11 Δ *rppH* también tuvo niveles elevados de transcrito para varios de los genes de esta vía, alcanzando valores de hasta 8 y 9 veces para los genes *pgi* y *tpiA* respectivamente. En el caso de la cepa PB13, los resultados son similares a los observados en la cepa PB12, a excepción de los genes *aceE* y *AceF*, los cuales se encuentran ligeramente más elevados en la cepa PB13.

Como ya se hizo mención, una de las principales características de la cepa PB11 son los niveles de transcrito de los genes involucrados con el catabolismo de la glucosa, los cuales están disminuidos, mientras que en la cepa PB12 (y ahora sabemos que también en PB13) ocurre lo contrario. No obstante, la diferencia crucial entre estas cepas es la capacidad de fosforilar la glucosa que se internaliza, pues el hecho de tener o no esta capacidad, es el origen de una gran cantidad de diferencias que se generan para poder contender con esta limitante. La enzima de la vía glucolítica encargada de llevar a cabo esta función en ausencia del sistema PTS se codifica por el gen *glk* y sus niveles de transcrito se encuentran elevados 3.4 y 6 veces en las cepas JM101 Δ *rppH* y PB11 Δ *rppH* respectivamente, lo que demuestra que la ausencia del gen *rppH* es responsable del incremento en la expresión de genes glucolíticos, incluido *glk*, y por lo tanto responsable de la característica de la cual

derivan las diferencias entre las cepas PB11 y las evolucionadas PB12 y PB13, las cuales incrementaron su nivel de transcrito para el gen *glk* 2.3 y 2.8 veces respectivamente.

Los datos observados denotan por tanto un incremento en los niveles de transcrito de genes involucrados con el catabolismo de glucosa (especialmente el gen *glk*), lo cual es congruente con los datos de consumo de glucosa cuyos valores se incrementan en las cepas PB11 Δ *rppH* y las evolucionadas PB12 y PB13; corroborando también que la pérdida del gen *rppH* fue un evento adaptativo que le confirió a la cepa PB11 la capacidad de internalizar y fosforilar la glucosa, así como su eficiente catabolismo.

Por otra parte, como ya se hizo mención, los niveles de transcrito de la cepa PB11 Δ *rppH* son mayores con respecto a los observados en las cepas evolucionadas. Originalmente se esperaba que los niveles de transcrito fueran muy similares, considerando que la cepa PB11 Δ *rppH* representa una cepa intermedia entre PB11 y las evolucionadas, sin embargo, es evidente que alguna o algunas mutaciones están influyendo hacia la disminución de los genes de la vía glucolítica.

b) TCA y cadena respiratoria:

Para el caso de los genes involucrados en el ciclo de los ácidos tricarboxílicos y en la cadena respiratoria, la cepa PB11 presenta menores niveles de transcrito para la mayoría de los genes con respecto al control (particularmente en aquellos involucrados con cadena respiratoria), y solo la transcripción de tres genes del TCA están incrementados 3, 4 y 5 veces (*fumA*, *fumC* y *acnA* respectivamente). Por otro lado, la cepa PB12 y PB13 tiene niveles similares de transcrito, los cuales están elevados entre 2 y 3 veces para la mayoría de los genes. No obstante, para la cepa PB12 los genes *fumA*, *fumC* y *acnA* poseen niveles entre 3.5 y 6 veces mayores.

Por su parte, la inactivación del gen *rppH* trajo consigo incrementos de entre 2 y 5 veces para la gran mayoría de los genes en la cepa JM101 Δ *rppH* y de entre 4 y 24 veces para todos los genes en la cepa PB11 Δ *rppH*. Con estos datos, es posible establecer que la inactivación del gen *rppH* trae consigo incrementos en los niveles de transcrito para ambos grupos de genes, no obstante es evidente que la inactivación del gen *rppH*, en un fondo PTS⁻ genera un efecto aún más acentuado sobre el incremento de los transcritos, donde también se observa una disminución en los niveles de transcrito en las cepas evolucionadas aunque de forma más drástica. Con respecto a las mutaciones que podrían estar llevando a cabo dicho efecto, resulta de particular interés aquella originada en la cepa PB12, sobre el gen *arcB*, de la cual se discute más adelante en este trabajo (Mutaciones puntuales no sinónimas en la cepa PB12). No obstante, es importante señalar en este momento, que los niveles de transcrito de la cepa PB13 también disminuyen, y son similares a los de la cepa PB12.

c) Gluconeogénesis, ciclo del glioxilato y *poxB*:

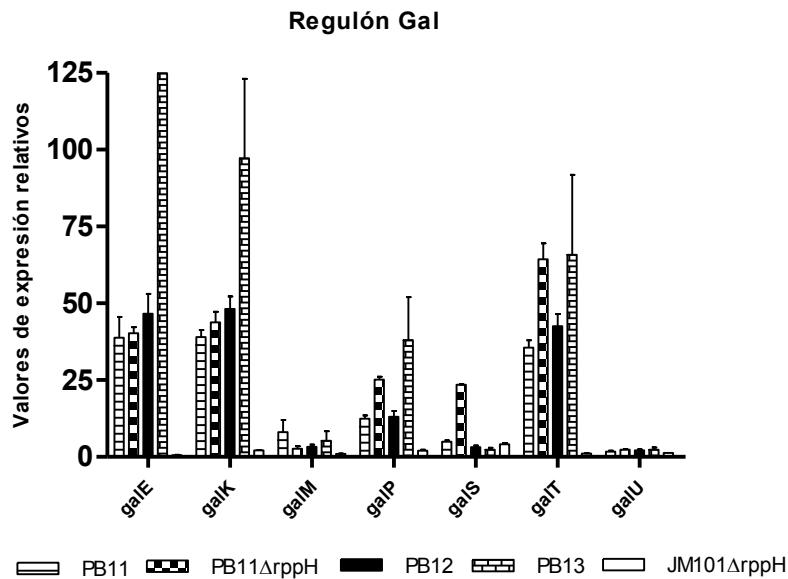
Este grupo de genes tiene niveles de transcrito de entre 2 y 6 veces mayores para la mayoría de los casos en las cepas PB11, PB12 y PB13. No obstante, existe una notable diferencia, en los genes involucrados con el ciclo del glioxilato, los cuales presentan para el caso de la cepa PB11 valores de entre 5 y 15 veces mayores, mientras que para las cepas evolucionadas los valores se encuentran en el rango de entre 2 y 4 veces superiores. Estos datos concuerdan con los análisis de flujos de carbono realizados en la cepa PB11, PB12 y PB13 donde se identificó un flujo elevado de carbono en el ciclo del glioxilato en la cepa PB11, como mecanismo de “scavenging” debido a la ausencia del sistema PTS. En lo que respecta a las cepas evolucionadas, los valores de este grupo de genes en general es ligeramente más elevado en la cepa PB12, lo que sugiere que en la cepa PB13 el efecto de la carga metabólica podría incluso ser menor.

Los efectos de la eliminación del gen *rppH* en las cepas JM101 y PB11, ocasionaron incrementos en los niveles de transcrito, los cuales se elevan entre 2 y 3 veces para el caso de la cepa JM101 Δ *rppH* (a excepción del gen *ppsA* el cual se incrementa 7 veces) y entre 4 y 28 veces en la cepa PB11 Δ *rppH*. Al igual que en el caso de los genes de la glucólisis, respiración y TCA, los niveles de transcrito de éste grupo de genes se ven disminuidos en las cepas evolucionadas al compararlos con los de la cepa PB11 Δ *rppH*.

d) Regulón de galactosa:

A excepción de lo observado para la cepa JM101 Δ *rppH*, donde los niveles de transcrito se elevaron alrededor de 2 veces con respecto al control, los genes del regulón de la galactosa presentan valores considerablemente elevados, y son los genes *galE*, *galK* y *galT*, los que presentaron los valores más altos (Gráfica 4). La explicación a los valores en este grupo de genes se deben a diferentes razones. Por un lado se observa que la ausencia de *rppH* sigue impactando sobre los niveles de transcrito (cepa JM101 Δ *rppH*), no obstante no es la única razón de este incremento, ya que la pérdida de la capacidad de asimilación de glucosa debida a la ausencia del sistema PTS (cepa PB11), ocasiona la inducción del regulón. En el caso de la PB11, los valores de los transcritos de *galE*, *galK*, *galT* y *galP* se elevaron a razón de 39, 39, 36 y 12 veces con respecto al control, lo que demuestra que en esta cepa se logra inducir este regulón (Gráfica 4). Esta inducción se describe como un mecanismo de “scavenging” en el cual la cepa sintetiza galactosa como inductor a partir de la escasa glucosa que logra internalizar [8,142]. Por tanto, la limitante para la asimilación de la glucosa no es su internalización, llevada a cabo por GalP, sino la fosforilación de la misma (*glk* presenta los mismos niveles de transcrito que la cepa control JM101). En cuanto a la cepa PB12 y PB13, los niveles de transcrito siguen siendo bastante elevados, y son muy similares a los presentes en la cepa PB11, no obstante, la razón de dichos valores es distinta, ya que aunado a la ausencia de *RppH*, ambas cepas

evolucionadas carecen también de GalR. Por lo tanto, los altos niveles de transcrito se deben a la incapacidad de reprimir dichos genes, aunado al posible efecto de la ausencia de RppH, e independientemente de la inducción del regulón, la cual teóricamente debe mantenerse en ausencia del sistema PTS para garantizar una eficiente transcripción del gen *galP*. La inducción del regulón es evidente en la cepa PB11 Δ *rppH*; ya que presenta valores elevados de transcrito para el regulón *gal* a pesar de mantener el represor GalR, lo cual aunado a los valores observados en la cepa PB11, sugieren un efecto del sistema PTS sobre la expresión del regulón, donde el sistema IIA de este sistema podría estar implicado.

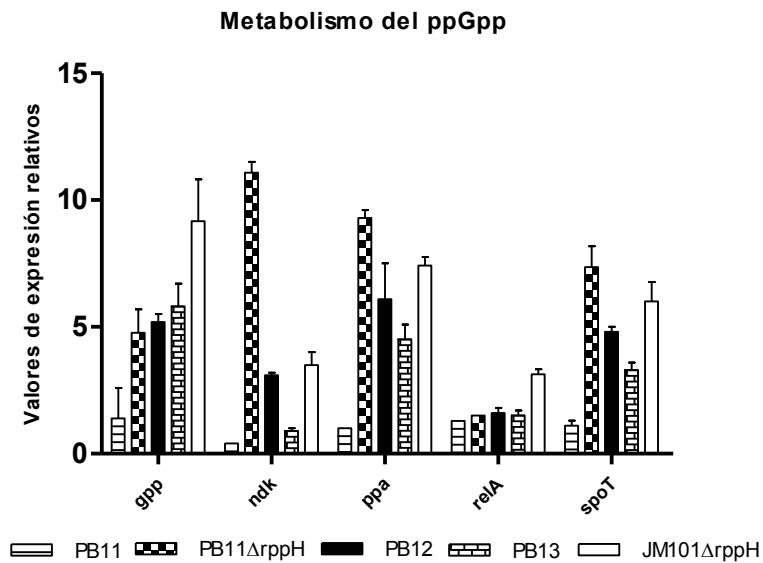


Gráfica 4.- Niveles de expresión relativos de genes del regulón *gal*.

e) **Metabolismo del tetrafosfato de guanosina (ppGpp):**

Una de las diferencias entre las cepas PB11 y PB12 que más llamaban la atención, eran los altos niveles de transcrito observados para los genes involucrados en el metabolismo del ppGpp en la cepa PB12 (la cual presentó incrementos de entre 3 y 6 veces), los cuales suponían una mayor concentración de este metabolito. La cepa PB13 tuvo valores elevados para este grupo de genes y de igual forma que en casos anteriores, muy similares a los observados para la cepa PB12. Por su parte, la cepa PB11 no tiene cambios con respecto a los de la cepa control (Gráfica 5). La inactivación del gen *rppH* en las cepas JM101 Δ *rppH* y PB11 Δ *rppH* ocasionó incrementos en los niveles de transcrito, superando incluso a los encontrados en la cepa PB12, lo que significa que el incremento en los niveles de transcrito de este grupo de genes en esta cepa se debe a la ausencia de RppH. Por el

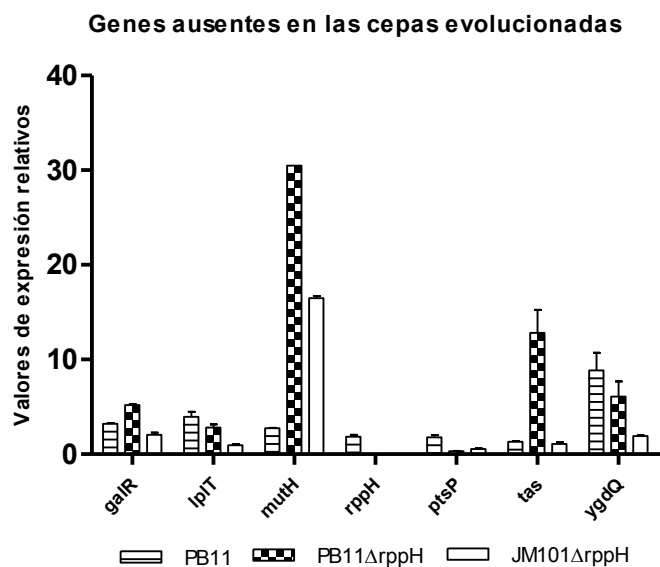
momento no se han determinado las concentraciones de ppGpp en las cepas evolucionadas; sin embargo, se sabe que al inactivar el factor sigma S la mayoría de genes dependientes de este factor de transcripción disminuyen drásticamente sus niveles de transcrito [117]. Adicionalmente, niveles elevados de ppGpp explicarían porque a pesar de poseer menores concentraciones de RpoS debido a la mutación puntual, éste sigue influyendo fuertemente sobre sus genes blanco. Bajo este escenario, una pregunta obligada, es qué efectos se generan en esta cepa, debidos a la transcripción basal de genes implicados en respuesta a estrés, y en consecuencia, si es posible establecer los elevados niveles de transcrito de genes involucrados con el metabolismo del ppGpp, como un mecanismo adaptativo producto del proceso de evolución.



Gráfica 5.- Niveles de expresión relativos de genes del metabolismo del ppGpp.

f) Genes ausentes en el cromosoma de la cepa PB12 y PB13:

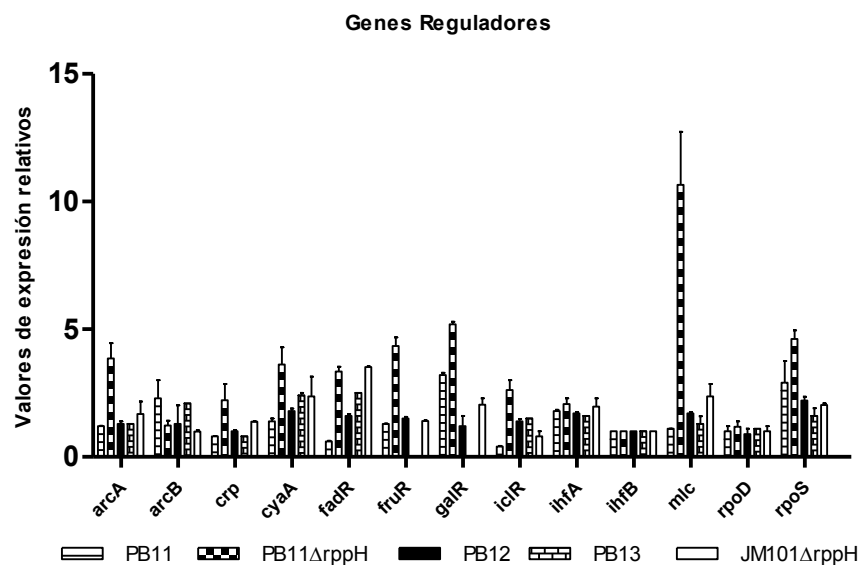
Con el objetivo de generar información que contextualise la ausencia del fragmento cromosomal en las cepas evolucionadas, se determinaron los niveles de transcrito de los genes que allí residen en las cepas PB11, PB11ΔrppH y JM101ΔrppH. Como resultado se observó una mayor expresión de todos los genes en la cepa PB11 (Gráfica 6), lo cual sugiere que podrían estar obedeciendo a una respuesta de estrés debido a la imposibilidad de metabolizar glucosa. Llama la atención los niveles de transcrito del gen *mutH* que se encontraron en las cepas ΔrppH; de 16 veces para la cepa JM101ΔrppH y de 30 veces para la cepa PB11ΔrppH, con respecto al control; lo cual significa que la ausencia de *rppH* tiene especial injerencia sobre el gen *mutH*.



Gráfica 6.- Niveles de expresión relativos de genes en el fragmento cromosomal ausente.

g) Genes reguladores:

En el estudio se incluyeron varios genes implicados en la regulación global (Tabla S10); los cuales en general, no varían mucho su nivel de transcrito con respecto al control, a excepción del gen *rpoS*, el cual tiene valores elevados 4.6 veces en la cepa PB11ΔrppH, 3 veces en la cepa PB11 y 2 veces en la cepa PB12 y JM101ΔrppH (Gráfica 7).



Gráfica 7.- Niveles de expresión relativos de genes reguladores

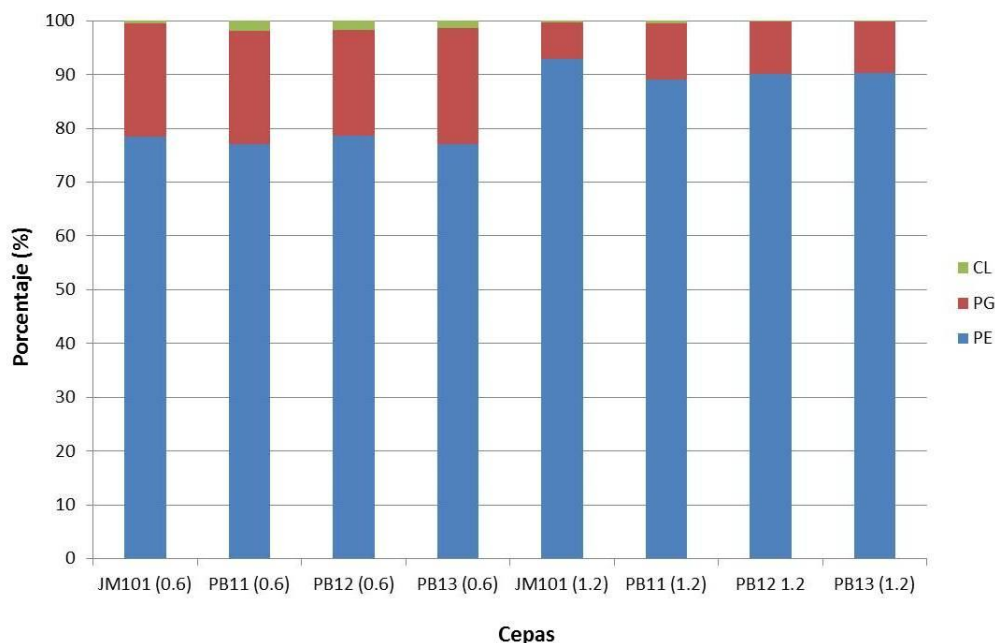
Estos datos, demuestran que en las cepas PB11 y PB12 se induce una condición de estrés, la cual se sugiere, más acentuada en la cepa PB11, de acuerdo a los niveles de transcrito observados. Por su parte, la cepa PB13, en la que no se observan niveles elevados de transcrito, hace suponer que ha logrado superar la ausencia del sistema PTS, lo cual concuerda con la elevada qs^{Glc} que presenta. Por otro lado, la cepa PB11 $\Delta rppH$ presentó los valores más elevados de transcrito, alcanzando niveles de entre 2 y 5 veces para la mayoría de los genes, a excepción del gen *mlc*, el cual se incrementó 10 veces. Para esta cepa resulta difícil tratar de establecer los efectos que producen los altos niveles de transcrito observados, ya que reguladores tales como RpoS, Mlc, GalR, IclR, Crp, y ArcA entre otros, tienen como blanco varios genes, produciendo efectos pleiotrópicos cada uno, los cuales de manera conjunta ofrecen un panorama más complejo. A pesar de ello, la cepa PB11 $\Delta rppH$ representa un punto metabólico intermedio entre la cepa PB11 y las cepas evolucionadas, que pone de manifiesto que las mutaciones puntuales y la ausencia de otros genes dentro de la región cromosomal, son esenciales para estabilizar metabólica y transcripcionalmente a las cepas PB12 y PB13. Esto a su vez constata que si bien, la ausencia del gen *rppH* por sí solo es suficiente para incrementar el consumo de glucosa así como su catabolismo, no basta para generar una cepa metabólicamente estable que pueda ser empleada en IVM como cepa de producción así como tampoco probablemente sería competitiva en términos evolutivos.

5. Las cepas PB12 y PB13 carecen del ciclo de la 2-acil-glicerofosfatidiletanolamina:

Como se estableció previamente, las cepas PB12 y PB13 carecen de un fragmento cromosomal en donde residía el operón *aas-lpT*, cuyos genes están implicados en el recambio de PE (Fig. 17), el cual es el principal fosfolípido de membrana. Existen reportes que indican que las cepas desprovistas tanto del gen *aas* como del gen *lpT*, incrementan alrededor del 30% sus niveles de 2-acil-GPE [143]. La acumulación de 2-acil-GPE es superior en cepas carentes de los genes *aas* y *pldB*, lo que sugiere un papel importante de las enzimas con actividad de fosfolipasa tipo 2 (PLA₂) en la atenuación de la acumulación de 2-acil-GPE [136]. Se ha reportado también que cepas carentes del gen *aas* presentan una composición alterada de su perfil de fosfolípidos de membrana, ya que tienen una mayor proporción de 2-acil-GPE y de 2-acil-fosfatidilglicerol [143]. Es importante señalar que concentraciones altas de 2-acil-GPE tienen efectos negativos en el crecimiento, ya que ocasionan lisis celular.

Con el objetivo de definir si las cepas evolucionadas presentan una redistribución de su perfil de fosfolípidos de membrana, debido a la ausencia del operón *aas-lpT*, se determinó su composición mediante ensayos de cromatografía de capa fina (TLC). Los resultados mostraron que ambas cepas evolucionadas tienen el mismo patrón en su composición de fosfolípidos de membrana que el de la

cepa silvestre (Gráfica 8), ya que no se observa una mayor cantidad de 2-acil-GPE ni de 2-acil-fosfatidilglicerol durante la fase exponencial de crecimiento.



Gráfica 8.- comparación de fosfolípidos de membrana en las cepas de este estudio a DO de 0.6 y 1.2. PE, Fosfatidiletanolamina, PG, Fosfatidilglicerol, CL, Cardiolipina.

Estos datos indican que las cepas evolucionadas, a pesar de carecer del ciclo que se encarga de regenerar la fosfatidiletanolamina, impiden la formación y/o acumulación de otras especies de fosfolípidos, manteniendo en niveles promedio las concentraciones de PE, PG y CL.

Una de las alternativas para poder contender contra la ausencia del ciclo de la liso-PE, es la degradación de la PE. La vía de degradación de fosfolípidos comprende la acción de algunas lisofosfolipasas A1 y A2 (PLA_1 y PLA_2), que generan ácidos grasos libres y glicerofosfodiésteres (Fig. 17 A y B).

Posteriormente, gracias a la acción de la enzima glicerofosforil-diéster-fosfodiesterasa (GDP), se obtiene una molécula de glicerol-3-fosfato (G3P) y un alcohol a partir de la molécula de glicerofosfodiéster. Los ácidos grasos libres y el G3P pueden reincorporarse a la célula a través de transportadores específicos (FadD y GlpT respectivamente), donde pueden canalizarse hacia la glucólisis en el caso del G3P y hacia la β -oxidación en el caso de los ácidos grasos libres (Fig 17 C). Cabe señalar que ambas cepas evolucionadas comparten una mutación en el gen *glpT*, el cual codifica para el transportador de G3P, lo que sugiere algún efecto en el rendimiento de dicho transportador.

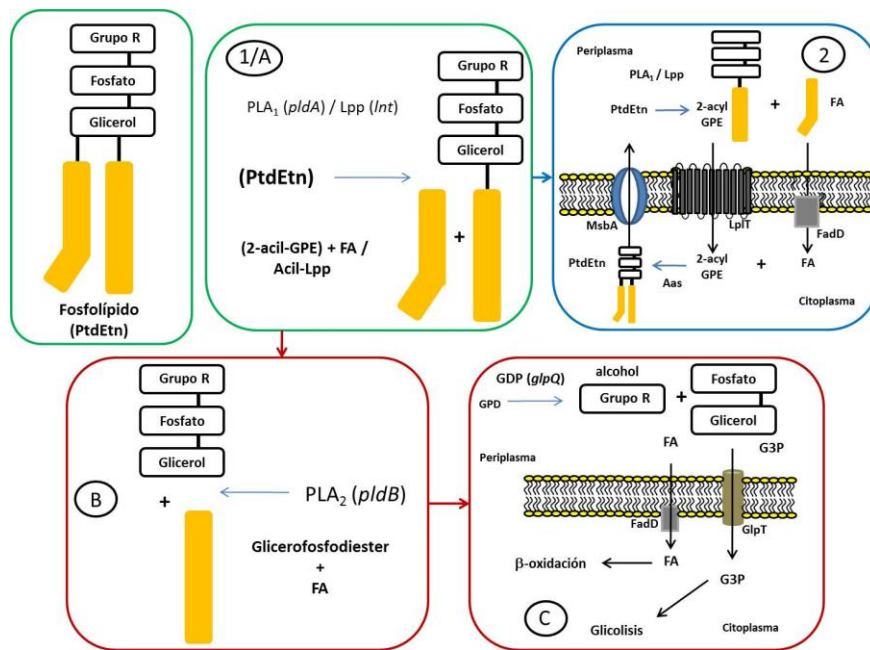


Figura 17.- Ciclo de la liso-PE. **1/A.** La PE es desacilada por acción de fosfolipasas de tipo 1 (PLA_1) o por lipoproteínas (Lpp), generando una molécula de ácido graso libre (FA) y una molécula de 2-acil-GPE. **2.** La molécula de 2-acil-GPE es internalizada por el transportador LpIT, mientras que la molécula de ácido graso se internaliza por conducto de FadD. Posteriormente, con ambas moléculas como sustrato, la enzima 2-acil-glicerofosfatidiletanolamina acil-transferasa / acil-ACP sintasa regenera PE mediante la acilación de la molécula de 2-acil-GPE. **B.** Mediante la acción de fosfolipasas A2 se genera otra molécula de ácido graso y glicerosfodiéster. **C.** Posteriormente gracias a la acción de la enzima glicerofosforil-diéster-fosfodiesterasa (GDP), el glicerosfodiéster es degradado para formar una molécula de G3P y un alcohol. Los ácidos grasos libres y el G3P son transportados al interior celular mediante transportadores específicos para ser aprovechados en la glucólisis para el caso del G3P, y para el caso de los ácidos grasos, para ser degradados en la β -oxidación.

La utilización de la vía de degradación de fosfolípidos podría explicar el incremento en el crecimiento de la cepa $PB11\Delta aas-lpIT$, ya que se elimina un ciclo de carbono donde las moléculas solo se interconvierten sin llegar a ser realmente aprovechadas. De hecho, se ha observado que la cepa PB11 responde favorablemente con la eliminación de tales ciclos, los cuales le resultan inútiles [112]. En el caso de las cepas evolucionadas, la utilización de la vía de degradación le permitiría por un lado, contener con la formación de especies nocivas de fosfolípidos, y por otro lado podría estar generando un balance metabólico favorable, ya que los productos de la degradación pueden ser reaprovechados de acuerdo a las necesidades celulares, lo que optimiza la utilización de las moléculas de carbono.

Con el objetivo de determinar en primera instancia si es la vía de degradación previamente descrita la que se está utilizando en ausencia del ciclo de de la liso PE, se inactivó el gen *glpQ*, el cual codifica para la enzima GDP. Posteriormente, para tratar de establecer si la reutilización de los productos de degradación es importante para el crecimiento de las cepas, se inactivaron los genes que codifican para los transportadores de G3P (*glpT*) y de ácidos grasos (*fadD*). Cabe señalar que el transportador FadD, además de internalizar las moléculas de ácidos grasos, les añade una molécula de CoA para

su “activación”, ya que sin esta reacción de esterificación las moléculas de ácidos grasos no pueden metabolizarse [144]. En consecuencia se generaron derivadas de las cepas JM101, PB11, PB12 y PB13 carentes de los genes *glpQ*, *glpT* y *fadD* de manera individual.

Cepa	μ	Porcentaje
JM101	0.71	100
JM101 Δ <i>glpT</i>	0.74	104
JM101 Δ <i>glpQ</i>	0.72	101
JM101 Δ <i>fadD</i>	0.69	97
PB11	0.13	100
PB11 Δ <i>glpT</i>	0.13	100
PB11 Δ <i>glpQ</i>	0.13	100
PB11 Δ <i>fadD</i>	0.12	92
PB12	0.44	100
PB12 Δ <i>glpT</i>	0.25	57
PB12 Δ <i>glpQ</i>	0.26	60
PB12 Δ <i>fadD</i>	0.32	73
PB13	0.47	100
PB13 Δ <i>glpT</i>	0.25	53
PB13 Δ <i>glpQ</i>	0.26	55
PB13 Δ <i>fadD</i>	0.29	62

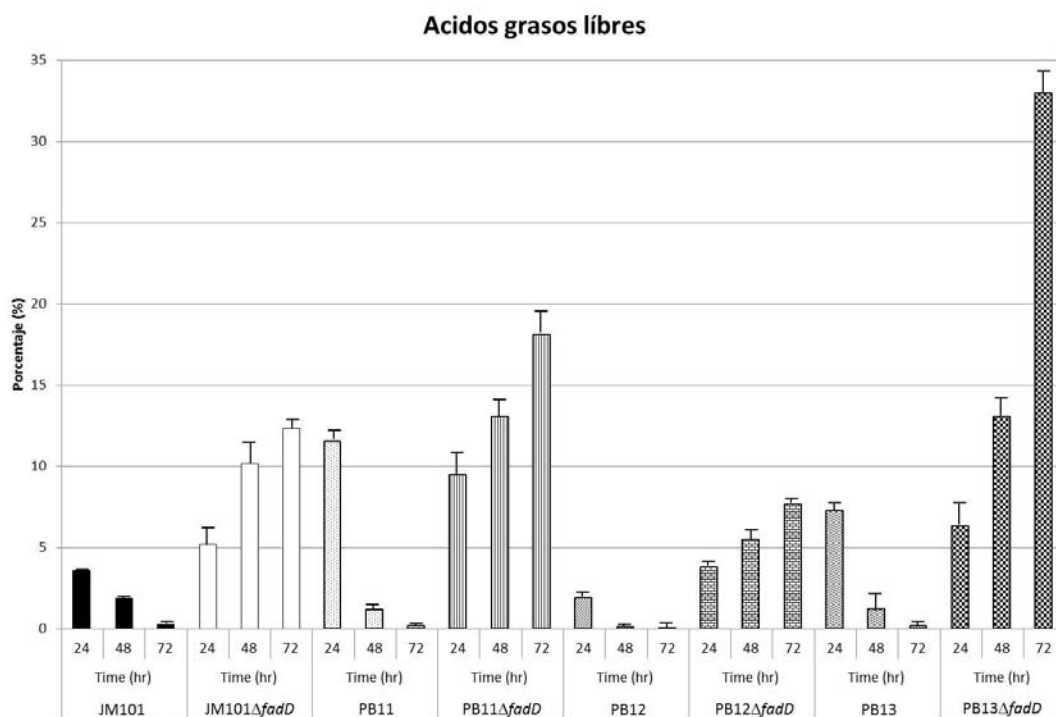
Tabla 3.- Velocidades específicas de crecimiento en derivadas de las cepas JM101, PB11, PB12 y PB13 carentes de los genes *glpQ*, *glpT* y *fadD*, utilizando medio mínimo M9-Glc 2g/L

Como resultado del bloqueo en la vía de degradación de la PE de las cepas evolucionadas, se observó una disminución del crecimiento del 40 y del 45% en las cepas PB12 Δ *glpQ* y PB13 Δ *glpQ* respectivamente, mientras que las cepas JM101 Δ *glpQ* y PB11 Δ *glpQ* no se vieron afectadas significativamente (Tabla 3). Esto demuestra que las cepas PB12 y PB13 utilizan la vía de degradación de fosfolípidos, cuyo bloqueo resulta inhibitorio para su crecimiento, debido probablemente a la formación de especies nocivas de fosfolípidos y a la disminución en la síntesis de “*novo*” de fosfolípidos. Adicionalmente la incapacidad para beneficiarse con los productos de degradación podría tener algún impacto sobre el crecimiento. En ese sentido, las derivadas de las cepas PB12 y PB13 mutantes en los genes *fadD* y *glpT*, disminuyeron significativamente su crecimiento (Tabla 3).

Las cepas PB12 Δ *glpT* y PB13 Δ *glpT*, disminuyeron su crecimiento un 43 y 47% respectivamente mientras que las cepas PB12 Δ *fadD* y PB13 Δ *fadD* disminuyeron su crecimiento un 27 y 38%, indicando que la asimilación de los productos de degradación (G3P y ácidos grasos libres) es importante para el crecimiento, al igual que lo observado en la cepa PB11 Δ *aas-lpIT*.

Adicional a la posible contribución de los ácidos grasos y del G3P como fuentes de carbono en las cepas evolucionadas, es probable que se genere de forma simultánea, una baja del flujo de carbono en la vía de degradación, derivado del bloqueo en la entrada de G3P y de los ácidos grasos, ya que las disminuciones del crecimiento observadas, sobre todo en el caso de las cepas carentes del gen *glpT* (las cuales son mayores al 40%), suponen un efecto similar al observado cuando se bloquea la vía de degradación. Esto sugiere que la eliminación del transportador GlpT, influye sobre el flujo de la vía de degradación; debido a la acumulación de G3P, lo que a su vez ocasionaría la acumulación de intermediarios previos, ocasionando una disminución en el flujo de carbono y por consiguiente también en el crecimiento.

Para el caso de las cepas evolucionadas carentes de FadD, al parecer el flujo de carbono por la vía de degradación no está tan afectado como en los dos casos anteriores, debido a que la eliminación de los grupos acilo de los fosfolípidos es independiente de la degradación de las moléculas de glicerofosfodiéster que se generan. El glicerofosfodiéster aún tiene que metabolizarse por la GDP, lo que significa que la eliminación de GlpT ocasiona la acumulación de G3P, alcoholes, glicerofosfodiéster y probablemente de 2-acil-GPE, mientras que la eliminación de FadD únicamente ocasiona la acumulación de ácidos grasos, que al parecer no incide tanto en el flujo de carbono de la vía de degradación.



Gráfica 9.- Acumulación de ácidos grasos libres en las diferentes cepas de este estudio y sus derivadas carentes del gen *fadD*, después de 24, 48 y 72 horas de cultivo. Los porcentajes se calculan a partir del total de especies lipídicas presentes en membrana y periplasma.

En ese sentido y considerando que existen diferencias en cuanto a crecimiento en las cepas PB12 Δ fadD y PB13 Δ fadD, es probable que éstas puedan deberse a variaciones en los flujos de carbono de ambas cepas para esta vía, generando por tanto, un diferencial en el aporte que representa la reutilización de los ácidos grasos generados, debido a su vez a la acumulación diferencial de éstos. Para corroborar si en las cepas PB12 y PB13 existe una acumulación diferencial de ácidos grasos libres, se realizaron determinaciones de su concentración mediante ensayos de cromatografía de capa fina (TLC). Los resultados se muestran en la gráfica 9 y representan concentraciones relativas dadas en porcentaje con respecto al total de especies lipídicas presentes en la membrana y el periplasma.

Como es posible observar, la cepa silvestre acumula alrededor del 4% de ácidos grasos durante las primeras 24 horas de cultivo, relaciones muy parecidas a las observadas en la cepa PB12. Por otra parte, las cepas PB11 y PB13 alcanzan una acumulación de alrededor del 10 y 8% respectivamente durante las primeras 24 horas de cultivo. Dichas concentraciones se incrementan a lo largo del cultivo en las cepas carentes del gen *fadD*, y es la cepa PB13 la que alcanza concentraciones cercanas al 35% después de 72 horas de cultivo. La cepa PB11 también acumula una cantidad considerable de ácidos grasos, alcanzando concentraciones cercanas al 20% después de 72 horas. Por su parte, en este mismo periodo de tiempo, las cepas JM101 y PB12 alcanzan concentraciones del 12 y del 8% respectivamente. Estos comportamientos nos indican que las cepas PB11 y PB13 son capaces de acumular más ácidos grasos, producto de una vía de degradación de fosfolípidos más activa. Así mismo, los datos sugieren que la cepa PB12 canaliza un flujo de carbono menor, y produce concentraciones de ácidos grasos muy pequeñas, similar a lo que ocurre en la cepa silvestre, no obstante en esta cepa las bajas concentraciones de ácidos grasos pueden deberse en parte a la utilización del ciclo de la liso PE, el cual se encuentra íntegro en la cepa JM101. De esta forma resulta comprensible que las cepas evolucionadas estén considerablemente afectadas con la inactivación de *fadD* y *glpT*, particularmente la cepa PB13, la cual destina un mayor flujo hacia la degradación de fosfolípidos y que contribuye en la optimización de su metabolismo. Por su parte, la cepa PB12, a pesar de destinar un flujo menor hacia la vía de degradación, se ve afectada únicamente debido a la imposibilidad de contender contra el bloqueo del metabolismo de fosfolípidos, lo cual incide sobre el crecimiento celular. Cabe señalar que en la cepa PB12, se generaron dos mutaciones puntuales en los genes *yjjU* y *rssA*, los cuales como se verá más adelante, de acuerdo a un análisis de secuencia, presenta un 15 y 18% de similitud respectivamente con la fosfolipasa de tipo 1 VipD de *Legionella pneumophila*, así como un 17% para ambos casos con la fosfolipasa de tipo 2 FabD de *E. coli*, lo cual podría explicar una menor acumulación de ácidos grasos en esta cepa.

Resulta interesante que la cepa PB11 no disminuya su μ , ya que a pesar de contar con el ciclo de la liso-PE el cual le ayuda a contender contra el bloqueo en la vía de degradación, la acumulación del 20% en las concentraciones de ácidos grasos que no puede reutilizar teóricamente debería incidir sobre el crecimiento.

6. Mutaciones puntuales en los genes *dhaM* y *glpT*:

Resulta interesante que las mutaciones puntuales en los genes *dhaM* y *glpT* (relacionados con el metabolismo del glicerol), se hayan fijado en el genoma en etapas tempranas y se hayan mantenido en todas las clonas aisladas. A continuación se describen brevemente las funciones y características de dichos genes:

glpT: Este gen codifica para el principal sistema de transporte de G3P en *E. coli*. Este transportador cuenta con 12 alfa hélices transmembranales, 7 regiones citoplásmicas y 6 periplásmicas [145,146,147]. La asimilación de G3P mediante GlpT ocasiona un contraflujo simultáneo de fosfato desde el interior celular. La mutación ocasiona el cambio de un codón de prolina por una leucina en la posición 139, dentro de la cuarta alfa hélice transmembranal, que se localiza embebida en la membrana. Debido al funcionamiento y la vía metabólica en la cual se sitúa, es posible que la mutación en este gen se haya seleccionado como un evento adaptativo, derivado de la ausencia del ciclo de la liso-PE, la cual, bajo estas condiciones podría optimizar su funcionamiento, lo que permitiría una mejor reutilización del carbono en forma de G3P. Sin embargo, dada la posición y la naturaleza de la mutación (ambos aminoácidos son hidrofóbicos), difícilmente podría haber generado algún cambio en el funcionamiento de la proteína. Hasta este momento, no se han realizado ninguna clase de aproximaciones para sustentar o desechar estas hipótesis.

dhaM: Codifica para la proteína DhaM, la cual es la subunidad M de una dihidroxiacetona fosfato. Esta subunidad consta de 3 dominios homólogos a ciertos componentes del sistema PTS (EI, Hpr y el componente IIA de manosa). Al parecer DhaM funciona como homodímero, y puede fosforilarse en múltiples sitios (al menos *in vitro*) utilizando PEP, aunque únicamente en presencia de EI y Hpr, y sirve como proteína donadora de grupos fosfato en las reacciones de dihidroxiacetona cinasa, llevando a cabo la reacción de fosforilación de la dihidroxiacetona (DHA), molécula formada a partir de glicerol en cepas carentes de la enzima glicerol cinasa (GlpK) en condiciones anaerobias. Una vez fosforilada, la molécula de dihidroxiacetona fosfato (DHAP), es entonces metabolizada a través de la glucólisis [148]. La mutación en esta proteína se generó en la posición 346, ocasionando el cambio de un codón de triptófano por un codón de término, lo que significa que el dominio homólogo a EI está incompleto (Fig. 18).

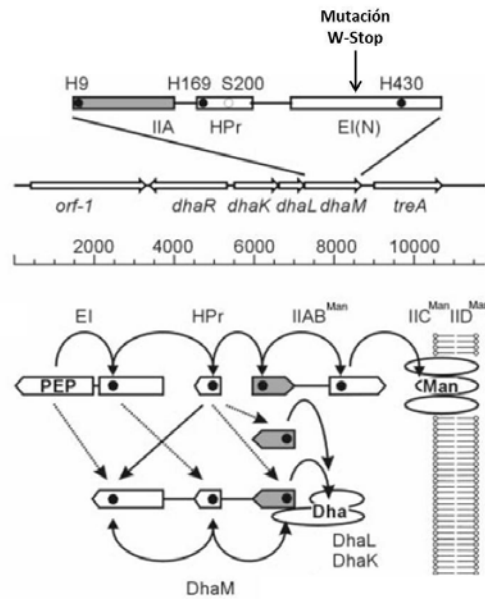


Figura 18.- Arriba. Operón *dhaKLM*, haciendo énfasis en los dominios homólogos de DhaM con el sistema de fosfotransferasa (PTS). También se señala la posición de la mutación puntual presente en las cepas evolucionadas. Abajo. Esquema del mecanismo de fosforilación de la dihidroxiacetona por conducto de la dihidroxiacetona cinasa (DhaK) a partir de PEP. Las líneas indican las diferentes interacciones posibles durante la cascada de fosforilación. Modificado de Gutknecht et al. 2001 [148].

De acuerdo con lo anterior, la mutación generaría una subunidad M incompleta, lo cual podría incidir sobre su capacidad de formar homodímeros y por tanto sobre su capacidad de metabolizar glicerol, en condiciones anaerobias. No obstante la mutación podría resultar irrelevante, en las cepas PB12 y PB13, tomando en cuenta que la funcionalidad de DhaM depende de la presencia de los componentes EI y Hpr del sistema PTS, y es indispensable únicamente en condiciones anaerobias. Debido a lo anterior, es probable que estemos ante un fenómeno que se presenta frecuentemente en experimentos de evolución conocido como pérdida de funciones no utilizadas [149]. Para corroborar el fenómeno de la pérdida de funciones no utilizadas como un evento adaptativo, se generaron derivadas de las cepas JM101, PB11, PB12 y PB13 carentes del gen *dhaM* y se procedió a su caracterización. Los resultados se presentan en la Tabla 4.

Cepa	μ	Porcentaje
JM101	0.71	100
JM101 Δ <i>dhaM</i>	0.71	100
PB11	0.13	100
PB11 Δ <i>dhaM</i>	0.11	85
PB12	0.44	100
PB12 Δ <i>dhaM</i>	0.37	84
PB13	0.47	100
PB13 Δ <i>dhaM</i>	0.28	60

Tabla 4.- Crecimiento en derivadas de las cepas JM101, PB11, PB12 y PB13 carentes del gen *dhaM* utilizando medio mínimo M9-Glc 2 g/L

Como es posible observar, la eliminación del gen *dhaM* trajo disminuciones del crecimiento del 15% en la cepa PB11, 16% en la cepa PB12 y del 40% en la cepa PB13, lo que descarta el fenómeno de pérdida de funciones no utilizadas como evento adaptativo. En este caso particular la ausencia total del gen ocasiona deficiencias en cuanto a crecimiento, lo que sugiere 3 elementos a considerar: 1) Que la mutación en el gen *dhaM* es adaptativa, 2) Que la proteína DhaM continúa siendo funcional a pesar de la mutación y 3) Que la versión mutante es funcional aún en condiciones aerobias en las cepas evolucionadas.

Se debe tener en consideración, según lo observado por Gutknecht [148], que la posibilidad de conservar la función de esta proteína es viable, ya que a pesar de la ausencia del sistema PTS y carecer del dominio homólogo a EI, los dominios homólogos a Hpr y IIA de manosa, aún podrían estarse fosforilando por conducto de otras proteínas homólogas a las del sistema PTS, aunque se deben tener en cuenta varios aspectos que se detallan a continuación.

E. coli posee cinco enzimas EI, seis proteínas Hpr y alrededor de 22 complejos de enzimas EII [150], cuyos elementos EI y/o Hpr podrían estar llevando a cabo la fosforilación de DhaM en su dominio Hpr y/o IIA respectivamente. Uno de estos sistemas homólogos al sistema de transporte de carbohidratos (PTS^{Glc}), es el sistema de fosfotransferasa de nitrógeno (PTS^{Ntr}), cuyo componente EI es codificado por el gen *ptsP*. Es importante recordar que las cepas evolucionadas, carecen de este gen, el cual se encontraba dentro de la región cromosomal que se perdió durante el proceso de evolución. Por su parte, de las cuatro proteínas homólogas a Hpr restantes solo tres de ellas (funcionales en condiciones aerobias) podrían llevar a cabo la fosforilación de DhaM en su dominio IIA [150]. En consecuencia, la funcionalidad de DhaM es posible a través de la fosforilación de otros componentes EI y Hpr homólogos.

Resulta interesante que las cepas PB12 y PB13 en ausencia del sistema PTS^{Glc} hayan evolucionado hacia la pérdida de otros dos componentes EI correspondientes al sistema de transporte de carbohidratos, al sistema de regulación relacionado al nitrógeno y el concerniente a la fosforilación de DHA, no obstante con los datos obtenidos hasta el momento, no es posible establecer algún sentido adaptativo.

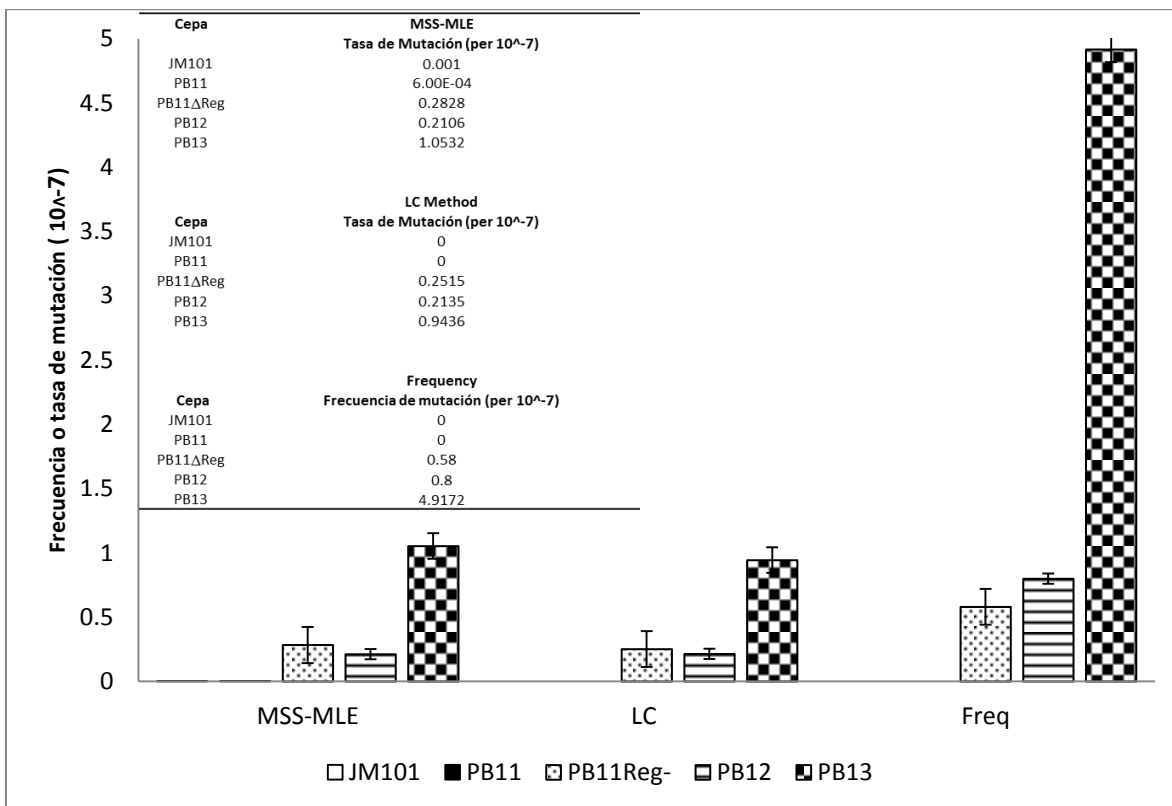
En lo concerniente a su funcionamiento en condiciones aerobias se debe señalar que, si bien es cierto que DhaK se utiliza en condiciones anaerobias, esto se debe a que su sustrato (DHA) se genera en estas condiciones. No obstante, todas las subunidades que forman el complejo de fosforilación (DhaK, L y M) se transcriben en ambas condiciones, por tanto, su funcionamiento no

estaría limitado a anaerobiosis. Por tal motivo, tomando en cuenta que la completa inactivación de DhaM afecta el crecimiento de las cepas evolucionadas, es posible plantearse la hipótesis de que esta proteína este llevando a cabo una función distinta, sobre todo porque su tarea principal no podría llevarse a cabo, ya que difícilmente se formaría DHA utilizando glucosa como única fuente de carbono. Del mismo modo, sería poco probable la formación de DHA durante la utilización de la vía de degradación de fosfolípidos, ya que como producto final se genera G3P, el cual es transformado directamente a DHAP por conducto de GlpD. Una posibilidad es su participación directa en la fosforilación de la glucosa entrante, o mediante la cascada de fosforilación DhaM-DhaK; sin embargo, esta posibilidad no se ha explorado en forma experimental.

7. Las cepas evolucionadas carecen del gen *mutH*:

Uno de los genes que se encontraban dentro de la región cromosomal es el gen *mutH*, el cual codifica para la endonucleasa MutH. Esta proteína funciona como monómero y es parte del sistema de reparación MMR (Mismatch Repair) de DNA, el cual interviene en vías de reparación generales, los cuales involucran apareamientos erróneos, inserciones y deleciones. Células deficientes en MMR exhiben fenotipos conocidos como mutables, en donde las tasas de mutación espontánea son muy elevadas; estas cepas también son hiper-recombinogénicas, debido a que MMR naturalmente impone una barrera a la recombinación entre especies divergentes [151]. Los reportes indican que cepas de *E. coli* deficientes en MutH incrementan su frecuencia de mutación hasta en 200 veces [152].

Con el objetivo de estimar la tasa de mutación de las cepas evolucionadas, se empleó el test de fluctuación de Luria-Delbruck, el cual se basó en la capacidad de generar resistencia a estreptomicina debido a una mutación puntual en el gen *rpsL* [153]. Los análisis se realizaron empleando los métodos de Lea-Coulson (LC) y Ma-Sandri-Sarkar (MSS-MLE) [154]. Asimismo se estimó la frecuencia de mutación.



Gráfica 10.- Tasas y frecuencias de mutación de las diferentes cepas empleadas en este estudio estimadas con el test de fluctuación de Luria-Delbruck, el cual se basó en la capacidad de generar resistencia a estreptomicina debido a una mutación puntual en el gen *rpsL*. Los análisis para las tasas de mutación se realizaron empleando los métodos de Lea-Coulson (LC) y Ma-Sandri-Sarkar (MSS-MLE).

Los resultados mostrados en la gráfica 10 y en la tabla S9 muestran una tasa de mutación de 0.2106×10^{-7} y de 1.0532×10^{-7} para las cepas PB12 y PB13 respectivamente al emplear el método MSS-MLE. Por otra parte al emplear el método LC, las cepas PB12 y PB13 tuvieron valores de 0.2135×10^{-7} y de 0.9436×10^{-7} respectivamente. Esto significa que la cepa PB12 posee una tasa de mutación alrededor de 200 veces superior con respecto a la cepa JM101 y PB11; mientras que la cepa PB13 presenta una tasa alrededor de 5 veces mayor con respecto a la cepa PB12 y 1000 veces mayor con respecto a la cepa silvestre. Adicionalmente se estimó la tasa de mutación de la cepa PB11ΔReg (Gráfica 10, Tabla S9) la cual tuvo una tasa de mutación muy similar a la de la cepa PB12 (0.2828×10^{-7} y 0.2515×10^{-7} empleando los métodos MSS-MLE y LC), de lo cual se infiere que la cepa PB13 posee una o varias mutaciones puntuales que le confieren una tasa de mutación más elevada.

Entre las mutaciones que pudieran ser las responsables del incremento en la tasa de mutación de la cepa PB13 se encuentra la mutación en el gen *polA*, el cual codifica para la DNA polimerasa I, la cual está implicada en varias vías de reparación de DNA [155]. La mutación se genera en la posición 695^{Ala-Val}, donde se sitúa el fragmento Klenow, el cual se encarga de corregir errores (proofreading) [156]. Con respecto a la frecuencia de mutación, la cepa PB13 presenta el valor más alto, que es de

alrededor de seis veces superior al de la cepa PB12, mientras que las cepas JM101 y PB11 presentaron un valor de cero (Gráfica 10, Tabla S9).

a) Implicaciones de una tasa de mutación elevada:

Durante los eventos de evolución adaptativa, donde el factor de estrés es mayúsculo, es común que se generen subpoblaciones que incrementan sus tasas de mutación (mutator populations) [157]. Estas poblaciones se seleccionan debido a su potencial para generar mayor variabilidad genética en menor tiempo, aumentando en consecuencia las probabilidades de supervivencia ante la presión selectiva. No obstante, la integridad genética de estas poblaciones se compromete en función de su tasa de mutación, ya que entre más elevada, mayor es el riesgo de generar una mutación deletérea, por lo que el mantenimiento de estas poblaciones a lo largo del tiempo no es deseable [16,158]. Se ha observado, tanto en poblaciones experimentales, como en poblaciones de vida libre, que posterior a la generación de subpoblaciones con elevada tasa de mutación y su incremento en "fitness", se tienden a seleccionar individuos dentro de esa subpoblación que restauran o disminuyen su tasa de mutación incluso a niveles normales, cuyos genotipos se propagan en diversas proporciones y velocidades a lo largo del tiempo, de acuerdo a la optimización en sus capacidades de reproducción y a las condiciones del medio en el que se desarrollan estos eventos [159,160]. En el caso de la evolución adaptativa a la que fue sometida la cepa PB11, podemos identificar varios factores que fueron determinantes al respecto. El primero de ellos es la fuerte presión de selección a la que fue sometida la cepa, lo cual a su vez se derivó de dos aspectos: las modificaciones genéticas que se le realizaron y el medio ambiente en el que se llevó a cabo el proceso. La ausencia del principal sistema de transporte de carbohidratos fue fundamental en el camino evolutivo de la bacteria, ya que además de haber eliminado las capacidades de internalizar glucosa vía PTS, también se eliminó la capacidad de utilizar otros carbohidratos que dependen de este sistema debido a la ausencia de sus componentes generales. Así mismo, se debe señalar la influencia que tuvo la eliminación de las capacidades reguladoras que se originaron con la inactivación del componente IIA. Otro de los factores que mediaron este proceso evolutivo, se centró en el medio ambiente, en este caso, la glucosa fue el carbohidrato para el cual tuvo que adaptarse, lo cual terminó por restringir el número de caminos posibles. De este modo el proceso de evolución fue dirigido en gran medida hacia la optimización de otro sistema de transporte y fosforilación de glucosa (*E. coli* posee tres sistemas de este tipo) aunque sin las restricciones de regulación transcripcional propias de una cepa normal. En ese sentido, la arquitectura genética y el medio ambiente en conjunto ejercieron un elevado grado de estrés sobre la cepa PB11, lo que explica la aparición de una subpoblación con elevada tasa de mutación. No obstante, el fenómeno de la pérdida del fragmento cromosomal en las cepas evolucionadas, representó dos eventos: 1) La generación de una población con elevada tasa de

mutación originada por la ausencia del gen *mutH*, que le permitió explorar un mayor número de soluciones en un menor lapso de tiempo; y 2) La ausencia de la región cromosomal también fue la solución para contender con la incapacidad de poder metabolizar la glucosa, debido a los efectos de la ausencia de los genes *rppH*, *aas*, *lpIT* y *galR*.

En consecuencia un solo evento de delección incrementó su capacidad de generar variabilidad para encontrar solución al estrés a enfrentar, y al mismo tiempo solucionó el problema. En estas circunstancias, de acuerdo a lo reportado [161,162], las cepas PB12 y PB13 tendrían que evolucionar hacia la selección de individuos que disminuyan su tasa de mutación, no obstante es evidente que al momento de su aislamiento, esto aún no sucedía, o al menos las clonas aisladas aún no presentan esta característica, lo que sugiere que este paso evolutivo aún no se presentaba al momento de finalizar el experimento.

8. Otras mutaciones relevantes en la cepa PB12:

Debido a que la cepa PB12 se ha utilizado en la sobreproducción de compuestos aromáticos, resulta muy importante para nuestro grupo, abordar también aquellas mutaciones que posee esta cepa y que en principio parecen tener relación con las características que ésta presenta y sugieren ser por tanto relevantes.

Para fines prácticos, las mutaciones puntuales no sinónimas se dividieron en tres grupos; aquellas ocurridas en genes implicados en regulación, en genes involucrados con el metabolismo y las ocurridas en genes de función desconocida (Tabla S4). Las mutaciones sinónimas y las originadas en regiones no codificantes no se abordaron en este trabajo.

a) Mutaciones puntuales no sinónimas en genes implicados en funciones de regulación:

En este grupo se encuentran los genes *arcB*, *barA*, *rpoD*, y *rna*. A continuación se hace una breve descripción general de ellos.

(1) Gen *arcB*:

Codifica para una proteína sensora cinasa miembro del sistema de dos componentes ArcA/B (anoxic redox control). Este sistema de dos componentes actúa durante condiciones anaerobias, principalmente como represor transcripcional de operones involucrados en la respiración, en el TCA, en el *shunt* del glioxilato, así como genes implicados en catabolismo de ácidos grasos [163,164,165]. *E. coli* sensa las condiciones redox de acuerdo al nivel de oxidación de las quinonas en la membrana, las cuales actúan como inhibidores del funcionamiento. Las quinonas por tanto determinan el nivel de

actividad de la sensora cinasa, induciendo cambios conformacionales en el homodímero cuyo mecanismo está relacionado con la formación de puentes disulfuro en las cisteínas 180 y 241 del dominio Pert-Arnt-Sim (PAS) (Fig. 19) [166]. La mutación en el gen *arcB*, cambia un residuo de tirosina por una cisteína en la posición 71 de la proteína, que corresponde a la segunda región transmembranal.

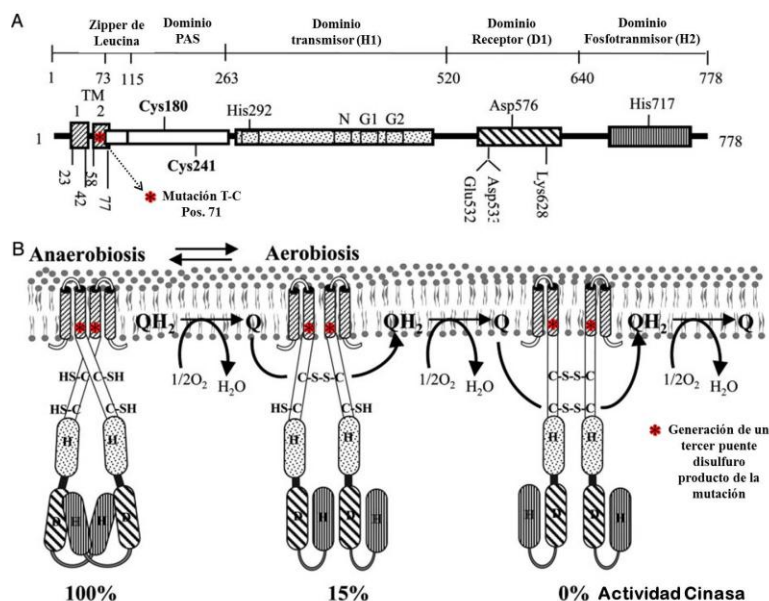


Figura 19.- A. Esquema de la estructura primaria de la proteína ArcB, donde se muestran los dominios presentes, así como los aminoácidos más importantes durante la cascada de fosforilaciones de la que depende su funcionamiento. Con el símbolo de un asterisco en color rojo se muestra la posición de la mutación en la cepa PB12, la cual cambió un residuo de tirosina por una cisteína en la posición 71 de la proteína, que corresponde a la segunda región transmembranal. B. Esquema de las diferentes configuraciones que adopta el homodímero y que determinan su nivel de actividad. Se señala también la posición donde se podría estar generando un tercer puente disulfuro, producto de la mutación (Modificado de Malpica et al. 2004) [166].

Anteriormente se había especulado que la mutación en este gen podría estar afectando de manera negativa el funcionamiento del sistema ArcA/B [8,109], lo que explicaría los niveles elevados de transcrito encontrados en la cepa PB12 cuando se le comparaba con la cepa PB11. Sin embargo, el hallazgo de la ausencia del gen *rppH* en la cepa PB12 y la caracterización de los perfiles de transcripción en las cepas $\Delta rppH$, establecen que en la cepa evolucionada, los niveles de transcrito para estos genes en realidad disminuyen. Aunque no se han realizado experimentos para determinar el efecto de la mutación en el gen *arcB*, la posición, el tipo de mutación y el mecanismo funcional del dímero apuntan a un posible efecto de semi-activación permanente del sistema, ya que el cambio hacia una cisteína en la región transmembranal sugiere la formación de un puente disulfuro adicional

en esta región, que modificaría la distancia de estas regiones en ambos monómeros, y que no necesariamente respondería a los niveles de oxidación, dado que el mecanismo no contempla un tercer puente disulfuro (Fig. 19). Por tanto, se presume que esta mutación congelaría al dímero en una configuración semi-activa, que explicaría en esta cepa la represión parcial de sus genes blanco aún en condiciones aerobias.

De confirmarse el efecto propuesto, significaría que existe una regulación diferencial entre ambas cepas evolucionadas, lo que también podría ayudar a explicar los comportamientos tan diferentes observados en el metabolismo de ácidos grasos de estas cepas, ya que este sistema de regulación, también influye sobre estos compuestos.

(2) Gen *barA*:

El gen *barA* codifica para la proteína sensora cinasa BarA (Bacterial Adaptive Response), la cual es miembro del sistema de dos componentes. BarA trabaja en conjunto con el regulador transcripcional de respuesta UvrY de forma muy específica y eficiente [167,168]. Está reportado que la proteína BarA regula la expresión de RpoS, un factor sigma alternativo de la RNA polimerasa que se activa durante la fase estacionaria de crecimiento [169]. Así mismo, se ha comprobado que este sistema de dos componentes actúa sobre la expresión de genes del metabolismo del carbono mediante la interacción con el sistema de regulación global Csr [170], el cual está involucrado en la represión de muchos genes de fase estacionaria (gluconeogénesis) así como en la activación de algunos genes de fase exponencial (Glucólisis) [171,172]. La mutación en este gen ocasiona el cambio de un residuo de fenilalanina por una leucina en la posición 366 de la proteína. Esta mutación afecta el primer dominio o dominio transmisor, relativamente cerca de dos regiones importantes para el funcionamiento de la proteína. La primera es el sitio donde se encuentra el residuo conservado de histidina esencial en este dominio, y la segunda región involucra los sitios necesarios para unir a la molécula de ATP, de la cual se toma el fosfato necesario para iniciar la cascada de señalizaciones (Fig. 20). Por esta razón se llevaron a cabo aproximaciones experimentales, para determinar los posibles efectos de la mutación en el sistema de dos componentes BarA/UvrY sobre el sistema de regulación global Csr. Desafortunadamente, los resultados de dichos experimentos no fueron concluyentes, lo que impide determinar si la mutación tuvo algún efecto [173].

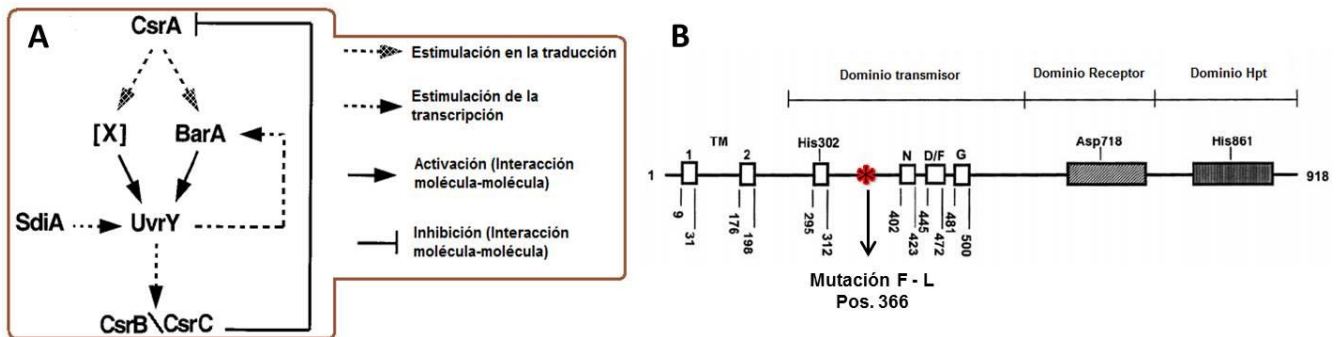


Fig. 20. A. Esquema de regulación entre los sistemas BarA/UvrY y Csr (Tomado y modificado de Suzuki, K., et. al. 2002). B. Esquema de la estructura primaria de la proteína BarA, donde se muestran los dominios presentes, así como los aminoácidos más importantes durante la cascada de fosforilaciones de la que depende su funcionamiento. Con el símbolo de un asterisco en color rojo se muestra la posición de la mutación en la cepa PB12, la cual cambió un residuo de fenilalanina por una leucina en la posición 366 de la proteína, que corresponde al primer dominio transmisor de la proteína (Tomado y modificado de Pernesting, et. al. 2001) [167].

(3) Gen *rpoD*:

Este gen codifica para el factor sigma 70, el cual es el principal factor sigma en *E. coli*, encargado del reconocimiento de promotores de genes esenciales. El cambio que se identificó en el gen *rpoD*, consistió en el reemplazo de una valina por una isoleucina en la posición 582, la cual se localiza en el motivo hélice-vuelta-hélice (HTH) de este factor de transcripción. De acuerdo a los reportes, y en particular a la determinación de la estructura de este factor sigma en la bacteria *Thermus aquaticus* [174], este motivo está involucrado en el reconocimiento de la región promotora -35; sin embargo, dada la naturaleza de los aminoácidos implicados (ambos son neutros no polares), así como el hecho de que dichos residuos no hacen contacto directo con el DNA, es poco probable que esta mutación ocasione efectos en el reconocimiento de las secuencias blanco. No obstante, no se han realizado experimentos de ningún tipo para determinar su efecto.

(4) Gen *rna*:

Este gen codifica para una proteína de 268 aminoácidos denominada RNAsa I, la cual se encarga de romper enlaces fosfodiéster entre 2 nucleótidos. Actúa principalmente sobre tRNAs, aunque no está limitada a ellos [175]. La mutación afectó la posición 90 de la proteína, originando la sustitución de un residuo de alanina por una treonina. Esta RNAsa está bastante estudiada, ya se ha cristalizado y se conoce su estructura [175]; por lo que se ha determinado que la mutación se dio fuera del sitio catalítico en una región no conservada de la proteína, y que difícilmente pudo haber ocasionado un efecto sobre su funcionamiento. Adicionalmente, con el objetivo de determinar la importancia de este gen dentro de la fisiología de la cepa PB12 así como de las evolucionadas se realizaron inactivaciones de este gen en las cepas previamente mencionadas, para su caracterización. Como

resultado, no se observaron modificaciones en cuanto al crecimiento de ninguna cepa, lo que indica que este gen no es determinante para un crecimiento óptimo, y sugiere aunado a las características de la mutación, que se trata de una mutación neutra (Fig. 21).

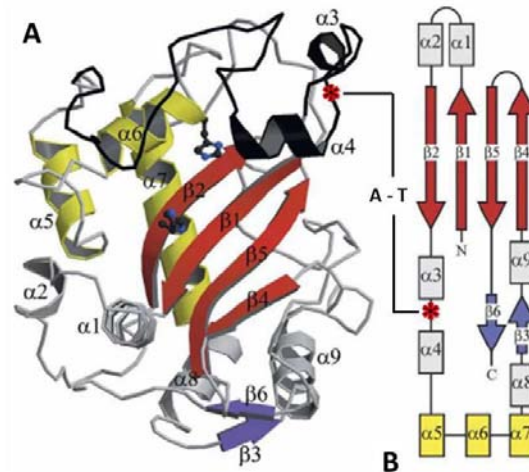


Fig. 21. A. Modelo de la estructura terciaria de la proteína RnaI. B. Esquema de la estructura secundaria de la proteína RnaI. En ambos esquemas se identifican los dominios estructurales de la proteína, así como la mutación puntual originada en la posición 90 de la proteína que corresponde a una estructura de unión (loop) entre dos α -hélices y que ocasiona el cambio de una alanina por una treonina en la cepa PB12 (Tomado y modificado de Rodríguez, *et. al.* 2008) [175].

b) Mutaciones puntuales no sinónimas en genes implicados en funciones metabólicas:

En este grupo se encuentran los genes *actP*, *arnT*, *chbC*, *csgF*, *dgoT*, *dppF*, *fdhD*, *fimH*, *rfe*, y *sucA* (Tabla S4). Hasta el momento no se han realizado aproximaciones experimentales para determinar los efectos de las mutaciones puntuales en estos genes. Llamen la atención las mutaciones en los genes *actP*, *chbC* y *dgoT* cuyos productos están implicados en el transporte de acetato, chitobiosa y galactonato respectivamente (Tabla S4), lo que sugiere que durante el proceso de evolución también se buscaron mutaciones que permitieran la entrada de glucosa mediante otros transportadores. Es probable que ninguna mutación tenga algún efecto que influya sobre el transporte de glucosa, y en caso de tenerlo, se estima que la internalización de este carbohidrato sea mínima, ya que la estereoespecificidad de tales transportadores es muy diferente.

c) Mutaciones puntuales no sinónimas en genes con función desconocida:

En este grupo se encuentran los genes *gfcD*, *yafV*, *ytfR*, *ydiQ*, *yjjU*, *rssA* y *ypdA*. A continuación se hace una breve descripción de los que se consideraron como los más importantes así como los

estudios realizados encaminados a la comprensión de sus funciones y los posibles efectos de las mutaciones.

(1) Los genes *yjjU* y *rssA* poseen un dominio “patatin-like”

El gen *yjjU* codifica para una presunta proteína de 357 aminoácidos. De acuerdo a un estudio de anotación de genes en *E. coli* [176], el gen *yjjU* se propuso como un presunto factor de transcripción; por esta razón se incluyó inicialmente como candidato para actuar de manera importante sobre la fisiología de la cepa PB12. La mutación en este gen se originó en la posición 338, que corresponde al aminoácido 113 de la proteína, donde se sustituyó un residuo de treonina por una isoleucina.

Por su parte, el gen *rssA* codifica para una proteína hipotética de 314 aminoácidos. Este gen llamó la atención debido a que se encuentra situado en el operón *rssAB* donde se localiza el gen que codifica para la proteína RssB, el cual está implicado en la regulación de RpoS [177]. Es posible que el producto del gen *rssA* pudiera estar involucrado de algún modo en la degradación del factor de transcripción RpoS, situación que se repite frecuentemente en distintos operones, donde los genes que allí residen, se encuentran involucrados en la realización de un fin común. La mutación se originó en la posición 773 del gen, que corresponde al aminoácido 258 de la proteína, donde se sustituyó un residuo de arginina por una histidina.

Con el objetivo de establecer alguna similitud en cuanto a secuencia con alguna proteína ya descrita, se realizaron alineamientos locales básicos (BLAST), utilizando las bases de datos de NCBI, Ecocyc y Ecogene [www.ncbi.nlm.nih.gov/blast/; <http://ecocyc.org/ECOLI/blast.html>; <http://www.ecogene.org/>]. Derivado de los análisis, llamaba la atención la ausencia de identidad con factores de transcripción, sistemas de dos componentes, o alguna otra proteína reguladora. No obstante fue evidente lo ampliamente distribuidos que se encuentran ambos genes en el género *Escherichia*, ya que únicamente se obtenían similitudes con el mismo gen pero de las diferentes especies y cepas reportadas de éste organismo.

Cabe señalar que en la base de datos de Ecogene, los genes *yjjU* and *rssA* (*yckK*) se consideran como genes homólogos distantes. Según esta base de datos, *rssA* (familia COG1752) parece ser el gen “original”, mientras que *yjjU* (familia COG4667) se obtuvo mediante un evento de transferencia horizontal. De acuerdo a esta base de datos, el gen *yjjU* estaría bajo el control del factor de transcripción CueR, el cual se encarga de regular genes involucrados con el mantenimiento de la homeostasis de cobre, mientras que *rssA* se encontraría bajo el control de GadE, el cual se encarga de regular genes involucrados en el mantenimiento de la homeostasis de pH [<http://www.ecogene.org/gene/EG12120>].

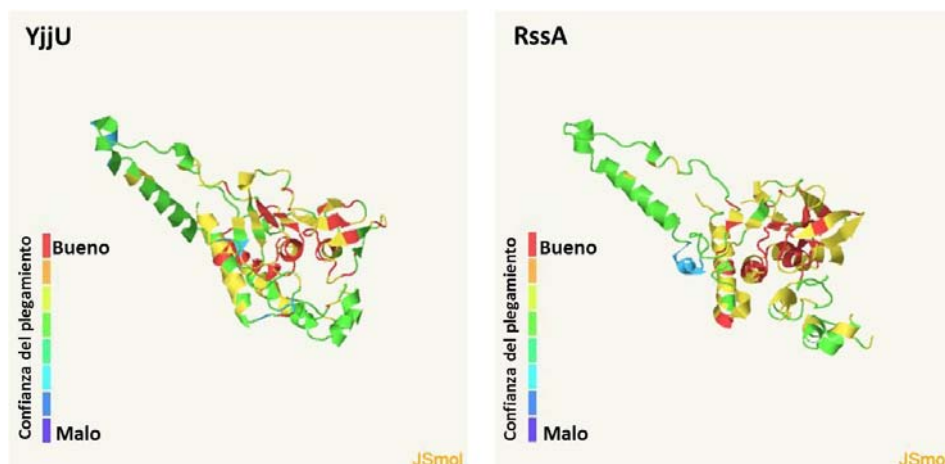


Fig. 22. Modelo de la estructura terciaria de las proteínas YjjU y RssA predicho mediante Phyre², el cual se basó en la identidad que poseen con la proteína VipD de *Legionella pneumophila*. Los diferentes colores indican ocho diferentes niveles de confianza que denotan la seguridad del plegamiento predicho a lo largo de la proteína.

Con el objetivo de tratar de establecer la posible función de ambos genes, se realizó un análisis de predicción de estructura y función de ambos genes. Para ello se utilizó el programa Protein Homology/ Analogy Recognition Engine V 2.0 (Phyre²), el cual predice y modela la estructura de la proteína codificada por el gen en cuestión, haciendo análisis de comparación con secuencias de proteínas ya cristalizadas y depositadas en la base de datos Protein data Bank (PDB). Ambos genes arrojaron resultados muy parecidos, encontrándose varios plegamientos posibles; de entre los cuales, el plegamiento más probable posee un 18% de identidad con la proteína VipD de *Legionella pneumophila* para ambas proteínas. Este plegamiento obtuvo altos niveles de confianza con una cobertura del 78% (del aminoácido 20 al 300) para el caso de la proteína YjjU y del 88% (del aminoácido 15 al 310) para el caso de la proteína RssA.

En la figura 22 se muestran los resultados del modelamiento de ambas proteínas utilizando como molde la estructura de VipD (Fig. 22). Durante el análisis se identificó un dominio denominado “patatin”, el cual es característico de algunas esterasas, lipasas y acil-transferasas, ya que se encarga de hidrolizar los grupos acilo de diversos compuestos lipídicos, como es el caso de VipD.

La proteína VipD de *Legionella pneumophila* es una fosfolipasa de tipo 1 la cual está implicada en el bloqueo de la formación de endosomas por parte de los macrófagos durante procesos de infección [178]. Cabe señalar que entre la lista de plegamientos posibles arrojada por Phyre², los primeros puestos eran ocupados por plegamientos basados en fosfolipasas, y fue una fosfolipasa de tipo 2 la que ocupaba el segundo lugar. El dominio patatin en la proteína YjjU abarca del aminoácido 27 al 196 y en la proteína RssA del 8 al 168.

Resulta interesante por tanto, que se hayan generado mutaciones en genes implicados con el metabolismo de fosfolípidos, tomando en cuenta que esta cepa carece del ciclo de la liso-PE, lo que

sugiere algún sentido adaptativo de tales mutaciones. Por tal motivo se especula que dichas mutaciones pudieran estar disminuyendo la funcionalidad de las proteínas, dado que la derivada de la cepa PB12 carente del gen *fadD* presenta una menor acumulación de ácidos grasos a pesar de carecer del ciclo de la liso PE.

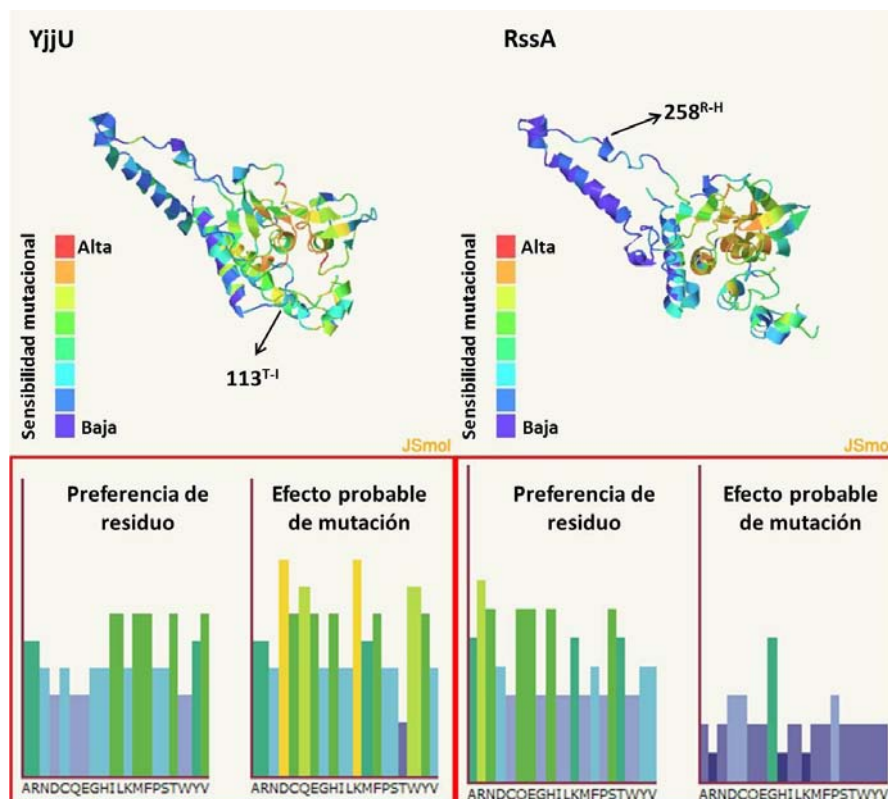


Fig. 23. Modelo de la estructura terciaria de las proteínas YjjU y RssA predicho mediante Phyre², el cual se basó en la identidad que poseen con la proteína VipD de *Legionella pneumophila*. Los diferentes colores en la estructura indican el grado de sensibilidad a una mutación a lo largo de la proteína. Dentro de las estructuras se señala la posición y la naturaleza de la mutación encontrada en estas proteínas en la cepa PB12. En las gráficas, tomando en cuenta los 20 aminoácidos y siguiendo el mismo patrón de colores, se señala la preferencia del residuo señalado en cada proteína, así como el efecto probable de una mutación.

Para averiguar un posible efecto a tales mutaciones, adicionalmente se realizó un estudio de mutagénesis, empleando el algoritmo SuSPect, el cual utiliza diversas características en cuanto a secuencia y estructura para predecir posibles efectos sobre el plegamiento de mutaciones puntuales no sinónimas. Dicho algoritmo utiliza un total de 77 características para discriminar entre mutaciones neutrales y aquellas que ocasionen algún efecto sobre la estructura [179]. Como resultado predice y asigna susceptibilidad de mutaciones a los aminoácidos que conforman la proteína de acuerdo a la estabilidad del plegamiento predicho. De acuerdo a estos análisis, la mutación en ambas proteínas no tendría un efecto considerable sobre su funcionamiento. En el caso de la proteína YjjU, debido principalmente a que la mutación originó el cambio de una treonina por una isoleucina, lo que el programa asignó como de bajo efecto (azul), a pesar de encontrarse en una región de sensibilidad

media (Fig. 23). Para el caso de la mutación en *RssA* la mutación en cuestión no generaría un cambio considerable debido a que se originó en una región a la que el programa designó como de baja sensibilidad (azul) aunado a la naturaleza de la mutación, la cual también designó como de bajo efecto (Fig. 23).

De esta forma, los resultados sugieren que las proteínas codificadas por los genes *yjjU* y *rssA* podrían desenvolverse como fosfolipasas, aunque resulta aventurado establecer su especificidad. En cuanto a las mutaciones ocurridas en estos genes, los resultados sugieren que podrían ser neutras, especialmente aquella que se origino en el gen *rssA*. Por tal motivo, no es posible explicar con los datos con los que se cuentan hasta este momento, la razón de las bajas concentraciones de ácidos grasos encontradas en la cepa PB12 y sus derivadas. No obstante debido a que los resultados obtenidos son basados únicamente en predicciones, deben tomarse con precaución.

(2) Gen *ypdA*:

Codifica para una proteína sensora cinasa (YpdA), cuyo regulador de respuesta es YpdB. Este sistema de dos componentes no ha sido muy estudiado, y se conoce poco acerca de sus efectores así como de los genes que regula, sin embargo se ha especulado que tenga influencia sobre el metabolismo central de carbono [180,181]. La mutación afectó la posición 200 de la proteína, cambiando un residuo de alanina por una histidina. Tomando en cuenta proteínas ya caracterizadas tales como YehU de *E. coli* y AgrA de *Staphylococcus aureus*, es posible especular que la mutación se haya generado en una región citoplásmica conectora del dominio transmembranal y el dominio GAF, que teóricamente no tendría un efecto drástico sobre su funcionamiento.

d) Inactivaciones de genes relevantes de función desconocida que sufrieron mutaciones en la cepa PB12:

Para finalizar con el análisis de las mutaciones puntuales en genes de función desconocida en la cepa PB12, se procedió como última aproximación experimental, a explorar la importancia de algunos genes sobre la capacidad de crecimiento de la cepa PB12. Para tal efecto, se procedió a la generación y caracterización de derivadas de la cepa PB12 carentes de los genes *yjjU*, *rssA* y *ypdA*, debido a que se consideraron como genes relevantes para esta cepa, tomando en cuenta la supuesta función que desempeñan. Los resultados de dichas caracterizaciones se presentan en la Tabla 5.

Strain	μ	%
JM101	0.71	100
JM101 Δ yjjU	0.74	104
JM101 Δ rssA	0.72	101
JM101 Δ ypdA	0.76	107
PB11	0.13	100
PB11 Δ yjjU	0.16	123
PB11 Δ rssA	0.17	131
PB11 Δ ypdA	0.13	100
PB12	0.43	100
PB12 Δ yjjU	0.34	79
PB12 Δ rssA	0.43	100
PB12 Δ ypdA	0.44	102

Tabla 5.- Velocidades específicas de crecimiento de diferentes derivadas de las cepas JM101, PB11 y PB12 carentes de los genes *yjjU*, *rssA* y *ypdA* utilizando medio mínimo M9-Glc 2 g/L.

Como resultado de tales experimentos se observó que la inactivación del gen *ypdA* no tiene efecto significativo en la velocidad de crecimiento en las distintas cepas. No obstante la inactivación de ambos genes que codifican para las presuntas fosfolipasas, ocasionaron incrementos del 23% para el caso de *yjjU* y del 31% para el caso de *rssA* en la cepa PB11. Por su parte la cepa PB12 únicamente modificó su μ cuando se inactivó el gen *yjjU*, observándose una disminución del 21%. Estos datos sugieren una participación importante para los genes *yjjU* y *rssA* sobre la fisiología de las cepas PB11 y PB12, aunque en el contexto genético y metabólico de ambas, no es posible establecer porque se genera un efecto diferencial en estas cepas, especialmente en el incremento del crecimiento observado en la cepa PB11, ya que si *YjjU* y *RssA* están implicadas en el metabolismo de fosfolípidos, su inactivación ocasionaría un bloqueo en las vías de recambio o degradación de estos compuestos. Por su parte resulta lógica la respuesta observada en la cepa PB12, donde podrían estarse dando dichos efectos, afectando por consiguiente su crecimiento. Lo anterior refleja por tanto, lo mucho que aún se desconoce sobre las adaptaciones metabólicas en la cepa PB11 y las adaptaciones genéticas en la cepa PB12, las cuales derivan en diversos efectos y que ocasionarían las diferencias observadas en este y varios aspectos a lo largo del proyecto.

C. Diseño de nuevas estrategias de producción en cepas PTS-

Este proyecto nació de la necesidad de responder una pregunta que surgió hace 20 años: ¿Que cambios genéticos le permitieron a la cepa PB12 recuperar su capacidad de asimilación de glucosa?

Actualmente es posible asegurar que la respuesta a tal cuestionamiento fue respondida, no obstante las implicaciones científicas intrínsecas que conllevaba dar solución a dicho problema derivó en la formulación de más preguntas, muchas de ellas anticipadas desde el momento en el que fue concebido el proyecto. En consecuencia, los aspectos a discutir en este trabajo fueron desarrollados en tres grandes temas: A) Cambios genéticos clave en la recuperación del crecimiento en glucosa en las cepas PTS⁻, B) Mecanismos adaptativos en cepas PTS⁻ creciendo en glucosa como única fuente de carbono y C) Empleo de ésta información para el diseño de nuevas estrategias de producción en cepas PTS⁻.

Los primeros dos aspectos se desarrollaron a lo largo de este escrito y tratan de dar respuesta a las preguntas que surgieron luego de la secuenciación del genoma de las cepas evolucionadas, tomando en cuenta por supuesto las inquietudes derivadas en los antecedentes y el contexto en el que se desarrolló el proyecto. Por otra parte, la necesidad de responder a la pregunta inicial surgió a su vez de la necesidad de desarrollar el tercer tema, el cual plantea una segunda pregunta:

1. ¿Por qué la cepa PB12 es tan buena productora de compuestos aromáticos?

La obtención de tan buenos resultados en la producción de compuestos aromáticos en la cepa PB12 [11,118,182,183,184], es resultado de múltiples aspectos, los cuales se han conjugado para hacer de esta cepa un microorganismo único.

Como aspecto principal a considerar, se encuentra su capacidad de consumo de glucosa, la cual representa un 60% del valor presente en la cepa silvestre JM101 y un 250% de aquel observado para la cepa PB11 (Sección: Caracterización de las cepas utilizando glucosa como única fuente de carbono), y es resultado esencialmente de la ausencia de los genes *rppH* y *galR*, los cuales le permitieron internalizar y metabolizar la glucosa en forma más eficiente. Así mismo, las mutaciones puntuales generadas podrían haber incidido sobre esta capacidad, logrando en conjunto situar el nivel del catabolismo de la glucosa en un punto favorable para realizar IVM sobre la cepa. Esta capacidad, por tanto, es responsable de que la cepa PB12 no genere un sobreflujo de carbono en la vía glicolítica, impidiendo la acumulación de grandes concentraciones de acetato como producto de desecho (del total de carbono consumido, la cepa PB12 acumula un 36% de acetato, mientras las cepas JM101 y PB13 generan más del 50% –porcentaje molar-) [109]. La canalización del flujo de carbono mediante ingeniería genética hacia la síntesis de compuestos aromáticos así como la optimización de la vía mediante la construcción de operones sintéticos logran disminuir aún más la producción de acetato, acumulando concentraciones debajo del 10%, a pesar de utilizar hasta 100 g/L de glucosa en cultivos de alta densidad celular [11].

Quizá el aspecto más importante de la cepa PB12 es su arquitectura genética, de la cual evidentemente derivan una gran cantidad de características que desembocaron después del proceso de evolución adaptativa en una cepa favorable para la producción de compuestos aromáticos (aunque probablemente no esté limitada a ello). Esta observación se desprende de algunos esfuerzos realizados en el grupo de trabajo para tratar de reproducir las capacidades más atractivas de la cepa PB12 en otras cepas de *E. coli*; específicamente en la cepa W3110. En ella se realizó ingeniería genética reversa con los datos obtenidos de este análisis. Sin embargo, los resultados no fueron los esperados, ya que a pesar de reproducir las mutaciones más importantes en esta cepa (*rppH* y *galR*) sobre un fondo PTS⁻ y haber recuperado su capacidad de asimilación de glucosa, los niveles de producción se encontraban muy por debajo de la cepa PB12, debido principalmente a la gran cantidad de acetato generado durante los procesos de producción (datos no publicados).

Dentro de los aspectos únicos dentro de la arquitectura genética de la cepa PB12 y que podrían influir en sus capacidades de producción de aromáticos se encuentra la mutación puntual en el gen *rpoS*. Está ampliamente reportado que cepas bajo condiciones de estrés nutricional desarrollan esta mutación como respuesta [185,186,187]. En consecuencia, el organismo deja de producir este factor de transcripción global, lo cual resulta favorable en dichas circunstancias aumentando sus probabilidades de supervivencia en fase estacionaria (GASP) y le permite desarrollar nuevos mecanismos de regulación en ausencia del factor sigma S, ya que la competencia por la RNA polimerasa desaparece [188]. El caso de la cepa PB12 es un tanto distinto, ya que posee una mutación en el gen *glnX* que ocasiona la supresión de codones de término del tipo ambar (Sección: Determinación del genoma de las cepa JM101 y PB11), lo que ocasionaría la producción de RpoS, aunque ciertamente disminuida. De este modo en la cepa PB12 habría una regulación diferencial con respecto a otras cepas, ya que no depende tanto de la respuesta mediada por RpoS, lo cual ya ha sido demostrado [117]. Adicionalmente, la cepa PB12 carece del componente IIA del sistema PTS, lo cual ocasiona una serie de adaptaciones de regulación que son únicas en éste microorganismo, resultado de la arquitectura genética de la cepa, y que resultó favorable en cuestiones de producción, y generó distintas respuestas con respecto a otras cepas, incluso bajo las mismas condiciones y con modificaciones genéticas idénticas.

No obstante las ventajas que ha demostrado esta cepa, cabe señalar aquellas posibles desventajas, ya que la cepa PB12 presenta una tasa de mutación alrededor de 200 veces superior a la cepa silvestre JM101, lo que supondría inestabilidad sobre características genéticas previamente establecidas; sin embargo a lo largo de todos los proyectos realizados con esta cepa, no se han

presentado eventos que supongan inestabilidad. De hecho ha llamado la atención la capacidad de adaptación que presentan las cepas productoras, ya que han pasado por un buen número de modificaciones genéticas y siguen presentando una gran plasticidad metabólica, logrando contener eficientemente contra tales alteraciones, lo cual hace pensar que su capacidad de adaptación provenga en parte de su capacidad para seguir generando diversidad.

Pese a los beneficios obtenidos con la cepa PB12, es posible encontrar alguna otra cepa con características genéticas distintas, que logre superar incluso los niveles en cuanto a producción que esta cepa ha conseguido; no obstante para ello sería necesario, realizar análisis de la influencia de la arquitectura genética sobre las modificaciones metabólicas y de regulación que se generaron en esta cepa producto del proceso de evolución adaptativa y como resultado de las modificaciones encaminadas al incremento en la producción de aromáticos.

2. Nuevas estrategias de producción utilizando cepas de E. coli PTS

A raíz de lo anterior, es importante preguntarse si hemos alcanzado el máximo nivel de producción de compuestos aromáticos empleando a la cepa PB12, ya que sobre ella, se han implementado una gran cantidad de estrategias sencillas y combinadas, logrando disminuir bastante el margen de estrategias restantes. Como resultado de este y otros trabajos de investigación realizados en el laboratorio, nos hemos convencido de la calidad de dichas estrategias. A pesar de ello valdría la pena tomar algunas de las estrategias de IVM más importantes, y conjugarlas con los conocimientos derivados de este trabajo para desarrollar sobre otros fondos genéticos, estas nuevas estrategias. El empleo de nuevos fondos genéticos incluso llevando a cabo las mismas estrategias de producción, refrescarían el campo de estudio, generando nuevos datos y nuevas cepas, las cuales pudieran incluso tener una arquitectura genética más favorable y que rindieran aún más que lo obtenido con la cepa PB12.

Se ha hecho énfasis a lo largo de este escrito, sobre la importancia que representa poseer una capacidad de consumo de sustrato lo suficientemente elevado como para impedir la limitación de biomasa, pero lo suficientemente bajo para impedir la formación de compuestos indeseados como producto de desecho. Sin embargo, este nivel de qs óptimo difiere entre las cepas de acuerdo a las características intrínsecas dadas por su arquitectura genética.

La utilización de nuevos fondos genéticos implicaría por tanto, encontrar este balance, ya sea mediante sobreexpresión directa de los genes implicados en sistemas de transporte de glucosa alternos, o incluso mediante el empleo de estrategias de evolución ya sea en cultivos continuos o

mediante transferencias seriales en matraz, lo cual generaría nuevas cepas con diferentes capacidades de metabolizar la glucosa. Una vez hallado el balance óptimo de consumo de sustrato, dependerá de cada genotipo, la plasticidad metabólica que de él se derive, para poder llevar a buen puerto las modificaciones genéticas subsecuentes que redireccionen el flujo de carbono hacia la síntesis de compuestos aromáticos.

VII. Conclusiones:

- Las cepas evolucionadas PB12 y PB13 son cepas divergentes que incrementaron su μ utilizando glucosa como única fuente de carbono debido al incremento en su capacidad de transporte, fosforilación y optimización del metabolismo de la glucosa. Las bases genéticas a dichas adaptaciones se identificaron gracias a la secuenciación de su genoma y comprenden esencialmente la ausencia de un fragmento cromosomal de 10,328 pb, en donde residían genes tales como *rppH*, *mutH*, *aas*, *lpIT* y *galR* entre otros.
- La pérdida del fragmento cromosomal ocurrió en etapas tempranas del proceso de evolución y se fijó en la población debido al carácter adaptativo para crecer mejor utilizando la glucosa del medio en ausencia del principal sistema de transporte de glucosa.
- La ausencia del gen *rppH* originó un incremento generalizado en los niveles de transcrito de la cepa, lo cual ocasionó efectos opuestos en el crecimiento. Por un lado incremento los niveles de transcrito de los genes *galP*, *glk*, y aquellos involucrados en la vía glucolítica, cuyos productos son responsables del transporte, la fosforilación y el metabolismo de la glucosa respectivamente. Por otra parte, la ausencia de RppH, ocasiona un efecto de carga metabólica que afecta las capacidades de crecimiento; sin embargo, en las cepas evolucionadas se ha atenuado dicho efecto, derivado probablemente de una o varias mutaciones puntuales que logran disminuir los niveles de transcrito.
- La ausencia del gen *galR* ocasionó la desrepresión del regulón *gal*, lo cual aunado al incremento en la vida media de los mensajeros y la inducción intrínseca generada en cepas carentes del sistema PTS, originó una sobreexpresión del gen *galP*, especialmente en la cepa PB13, lo que podría explicar la diferencia de crecimiento entre ambas cepas evolucionadas.

- La ausencia del operón *aas-lp/T* ocasionó la eliminación del ciclo de la liso-PE, el cual representa el 80% de los fosfolípidos de membrana y constituye un ciclo fútil bajo un fondo PTS⁻. En consecuencia, la ausencia del operón resultó en la optimización del metabolismo de carbono, ya que forzó a la célula a la utilización de la vía de degradación de fosfolípidos como único mecanismo para contender contra los efectos nocivos de la pérdida del ciclo de recambio. Como resultado las cepas evolucionadas no reciclan PE, sino que lo degradan hasta la formación de G3P y ácidos grasos libres, metabolitos que la célula logra reaprovechar, canalizándolos hacia glucólisis y β -oxidación respectivamente; siendo la cepa PB13 la que más activa tiene esta vía.
- La ausencia del gen *mutH* ocasionó una disminución en las capacidades de reparación de las cepas evolucionadas, lo que incrementó su tasa de mutación en alrededor de 200 veces con respecto a la cepa silvestre JM101; no obstante, la cepa PB13 presenta una tasa de mutación 1000 veces mayor, producto probablemente de una mutación en el fragmento Klenow de la DNA polimerasa I, el cual se encarga de corregir errores durante la replicación.
- La elevada tasa de mutación en la cepa PB12, hace obligatorio la reinsertión del gen *mutH* en las cepas utilizadas para la producción de compuestos aromáticos, ya que esta característica podría ir en detrimento de las capacidades de producción.
- Además del fragmento cromosomal, las cepas PB12 y PB13 comparten dos mutaciones puntuales, situadas en los genes *glpT* y *dhaM*, no obstante el carácter de ambas mutaciones es incierto.
- Contrario a lo que se había especulado con anterioridad, la mutación en el gen *arcB* originada únicamente en la cepa PB12, podría ser responsable de una disminución en los niveles de transcrito de aquellos genes controlados por el sistema de dos componentes ArcA/B, los cuales están involucrados esencialmente en respiración, TCA y ciclo del glioxilato.
- La cepa silvestre JM101 posee una mutación en la posición 98 del gen *rpoS*, lo que ocasiona el cambio de un residuo de glutamina por un codón de término (CAG – UAG). Como consecuencia de esta mutación se produce una proteína incompleta de 32 aminoácidos en lugar de la proteína normal que cuenta con 330 residuos. No obstante, la cepa JM101 tiene la capacidad de suprimir codones de término (mutaciones ámbar) originados por errores en la replicación. Por tal motivo es probable que se genere una menor cantidad de RpoS funcional, pero lo suficiente para activar genes altamente dependientes de este factor sigma. Esta

mutación por tanto no es resultado del proceso de evolución adaptativa en las cepas PB12 y PB13 como anteriormente se pensaba.

- De acuerdo a los análisis realizados, es muy probable que los genes *yjjU* y *rssA* codifiquen para lipasas A1 o A2. El efecto de las mutaciones originadas en estos genes en la cepa PB12 parece ser neutro.
- La efectividad de la cepa PB12 para la producción de compuestos aromáticos reside en su capacidad de consumir glucosa (qs^{Glc}) lo suficientemente rápido como para generar biomasa, sin ocasionar un sobreflujo de carbono que permita la formación de productos de desecho. Esta capacidad se deriva de la arquitectura genética de la cepa y constituye por tanto una característica única que es posible replicar en otros fondos genéticos aunque no bajo la misma estrategia.
- La eficiencia de otros fondos genéticos para la síntesis de compuestos aromáticos dependerá en gran medida de estudios previos que permitan establecer la qs^{Glc} adecuada, así como de la plasticidad metabólica, intrínseca a su arquitectura genética para realizar IVM y determinar su viabilidad.

VIII. Perspectivas:

- Determinación del papel del complejo DhaK en las cepas evolucionadas: Queda pendiente la determinación del papel que desempeña la dihidroxiacetona cinasa en las cepas evolucionadas, ya que parece ser importante para el crecimiento en estas cepas. Se especula que pudiera estar implicada en la fosforilación de la glucosa internalizada vía GalP.
- Caracterización de las proteínas YjjU y RssA: La caracterización de los productos de los genes *yjjU* y *rssA*, permitirá la identificación exacta de las funciones de estas proteínas, las cuales se proponen en este trabajo como presuntas fosfolipasas. Para ello se requerirán análisis enzimáticos para determinar su capacidad de hidrólisis de grupos acilo de diversos compuestos lipídicos.

IX. Metodología experimental

A. Cepas bacterianas:

Todas las cepas empleadas en este estudio se enlistan en la siguiente tabla A:

Cepas	Características	Referencias
JM101	F' [traD36 proAB+ lacI _q lacZΔM15] supE thi Δ(lac-proAB) rpoS(33am)	Messing J. 1979
JM101Δaas-lplT	JM101aas-lplT::Cm	Este trabajo
JM101ΔdhaM	JM101dhaM::Cm	Este trabajo
JM101ΔfadD	JM101fadD::Cm	Este trabajo
JM101ΔglpQ	JM101glpQ::Cm	Este trabajo
JM101ΔglpT	JM101glpT::Cm	Este trabajo
JM101ΔReg	JM101ptsP-rppH-ygdT-mutH-ygdQ-ygdR-tas-lplT-aas-omrA-omrB-galR::Cm	Este trabajo
JM101ΔrppH	JM101rppH::Cm	Este trabajo
JM101ΔrssA	JM101rssA::Cm	Este trabajo
JM101ΔyjjU	JM101yjjU::Cm	Este trabajo
JM101ΔypdA	JM101ypdA::Cm	Este trabajo
PB11	JM101ptsH-ptsI-crr::kan	Flores N. et al. 2005
PB11Δaas-lplT	PB11aas-lplT::Cm	Este trabajo
PB11Δaas-lplT-rppH	PB11aas-lplT::FRT, rppH::Cm	Este trabajo
PB11ΔdhaM	PB11dhaM::Cm	Este trabajo
PB11ΔfadD	PB11fadD::Cm	Este trabajo
PB11ΔglpQ	PB11glpQ::Cm	Este trabajo
PB11ΔglpT	PB11glpT::Cm	Este trabajo
PB11ΔReg	PB11ptsP-rppH-ygdT-mutH-ygdQ-ygdR-tas-lplT-aas-omrA-omrB-galR::Cm	Este trabajo
PB11ΔrppH	PB11rppH::Cm	Este trabajo
PB11ΔrssA	PB11rssA::Cm	Este trabajo
PB11ΔyjjU	PB11yjjU::Cm	Este trabajo
PB11ΔypdA	PB11ypdA::Cm	Este trabajo
PB12	PB11ev. (D=0.4h ⁻¹)	Flores N. et al. 2005
PB12ΔdhaM	PB12dhaM::Cm	Este trabajo
PB12ΔfadD	PB12fadD::Cm	Este trabajo
PB12ΔglpQ	PB12glpQ::Cm	Este trabajo
PB12ΔglpT	PB12glpT::Cm	Este trabajo

PB12ΔrssA	PB12rssA::Cm	Este trabajo
PB12ΔyjjU	PB12yjjU::Cm	Este trabajo
PB12ΔypdA	PB12ypdA::Cm	Este trabajo
PB13	PB11ev.(D=0.8h ⁻¹)	Flores S. et al. 2002
PB13ΔdhaM	PB13dhaM::Cm	Este trabajo
PB13ΔfadD	PB13fadD::Cm	Este trabajo
PB13ΔglpQ	PB13glpQ::Cm	Este trabajo
PB13ΔglpT	PB13glpT::Cm	Este trabajo
10Glc	PB11ev.(D=0.4h ⁻¹)	Flores N. 1996
32Glc	PB11ev.(D=0.6h ⁻¹)	Flores N. 1996
40Glc	PB11ev.(D=0.6h ⁻¹)	Flores N. 1996
99Glc	PB11ev.(D=0.8h ⁻¹)	Flores N. 1996

B. Medios de cultivo y condiciones para todas las determinaciones del crecimiento y parámetros cinéticos:

Las cepas se crecieron a 37°C con una agitación de 300 rpm en 25 mL de medio mínimo M9 con glucosa 2 g/L, utilizando matraces Klett bafleados de 125 mL y una incubadora con agitación orbital Brinkmann Orbimix1010.

C. Componentes del medio mínimo:

Cada 25 mL de medio mínimo M9-Glc 2 g/L contiene:

2.5 mL de sales M9 10x; 22.2 mL de H₂O; 50 μ L de MgSO₄ 1M.; 2.5 μ L de Vitamina B1 20 mg/mL; y 2.5 μ L de CaCl₂ 1M.; y 0.25 mL de glucosa al 20%.

Cada 100 mL de sales M9 10x contienen:

6 g de NaHPO₄; 3 g de K₂HPO₄; 0.5 g de NaCl y 1 g de NH₄Cl

D. Caracterización de todas las cepas empleadas en este estudio:

A partir del glicerol de respaldo, las células se crecieron en medio LB con sus respectivos antibióticos; las cepas PB11, PB12 y PB13 tienen resistencia a kanamicina, mientras que las mutantes generadas, tienen resistencia a kanamicina y cloranfenicol; ya crecidas en medio LB, se inocularon en medio mínimo M9 con glucosa 2 g/L y se incubaron a 37°C a 300 rpm.

Durante su crecimiento en medio mínimo M9 con glucosa se esperó a que los cultivos llegaran a la fase exponencial (entre 100 y 150 unidades Klett aproximadamente), para poder inocular a medio mínimo M9 con glucosa 2 g/L.

Nota: Para todas las cinéticas realizadas en este estudio la caracterización se realizó por lo menos dos veces de forma independiente y en cada caso por duplicado.

E. Extracción de DNA de las cepas JM101, PB11, PB12 y PB13 para la secuenciación de su genoma:

Las cepas JM101, PB11, PB12 y PB13 se crecieron a partir del glicerol original, en medio LB líquido. El DNA se obtuvo mediante CsCl maxiprep [120]. La calidad y la cantidad del DNA se verificó de acuerdo a las recomendaciones de Roche Nimblegen Inc. y de la Unidad Universitaria de Secuenciación Masiva de DNA del IBt, UNAM.

F. Secuenciación comparativa (Roche NimbleGen Inc.):

Las muestras de DNA de las cepas JM101 y PB12 se enviaron a la compañía Roche NimbleGen (RN) para un análisis de secuenciación comparativa (CGS) el cual emplea a la cepa *E. coli* K-12 MG1655 (ATC #47076) como cepa de referencia [189]. Los resultados de esta secuenciación se muestran en las tablas S1-S6.

G. Secuenciación Masiva (Illumina GAllx & Ion Torrent PGM):

Para su secuenciación, las muestras de DNA de las cepas JM101, PB11, PB12 y PB13 se enviaron a la unidad universitaria de secuenciación masiva. Los resultados de esta secuenciación se muestran en las tablas S1-S8.

1. Secuenciación mediante Illumina GAllx:

Para la secuenciación, primero se construyó una librería “pair ending” (PE) de las cepas JM101, PB11 y PB12. La librería PE se construyó siguiendo las recomendaciones de Illumina Inc. Brevemente, 5 µg de DNA cromosomal de cada cepa se fragmentó por nebulización de nitrógeno durante 6 min. A una presión de 32 psi. El DNA fragmentado se purificó utilizando el kit de purificación QIAquick PCR Purification, y se resuspendió en 30 µl de buffer de elución (EB: 10 mM Tris / HCl, pH 8.5). Los extremos de los fragmentos de DNA se repararon con una mezcla de polimerasas T4 y Klenow, así

como una T4 polinucleótido cinasa para extremos 5'. Para facilitar la ligación de los adaptadores de doble cadena, se incorporó a cada extremo 3' un residuo de adenina, utilizando la polimerasa Klenow exo minus (exo-). Los adaptadores se ligaron a cada fragmento de DNA utilizando una T4 DNA ligasa durante 15 min a temperatura ambiente. El DNA ligado se purificó utilizando el kit de purificación Qiagen MinElute (Qiagen, USA) y se resuspendió en 15 µL de EB. Las muestras de DNA se corrieron en un gel de agarosa al 2%; los fragmentos de 500 pb se seleccionaron y se purificaron utilizando el kit de extracción QIAquick gel extraction. Para enriquecer las muestras purificadas, se realizaron reacciones de PCR (12 ciclos por reacción: 98, 65 y 72°C) usando los primers PE 1.0 y 2.0 y la polimerasa Phusion DNA. Los productos de PCR se purificaron con el kit QIAquick PCR purification (Qiagen, USA) y se eluyeron en 50 µL de EB. La validación y cuantificación de la librería se realizó con el equipo Agilent Bioanalyzer 2100. Finalmente se utilizaron 18 pM de muestra para llevar a cabo la secuenciación PE empleando 2x36 ciclos con el equipo GAllx, el cual emplea un método basado en terminadores fluorescentes reversibles.

2. Secuenciación mediante Ion Torrent PGM:

Para la secuenciación, se construyó una librería de las cepas JM101 y PB13. La librería se construyó siguiendo las recomendaciones de Illumina Inc. Brevemente, se fragmentó 1 µg de DNA y posteriormente se repararon los extremos 5' en una sola reacción utilizando la mezcla de enzimas NEBNext DNA Fragmentation Master Mix. Posteriormente se ligaron los adaptadores utilizando la enzima T4 DNA ligasa. Se seleccionaron aquellos fragmentos de al menos 100 pb y se purificaron utilizando el kit AMPure XP Beads. Para enriquecer las muestras purificadas, se realizaron reacciones de PCR (14 ciclos por reacción: 98, 58 y 72°C) usando los primers 1.0 y 2.0 y la polimerasa NEBNext High-Fidelity. Los productos de PCR se purificaron con el kit QIAquick PCR purification (Qiagen, USA) y se eluyeron en 25 µL de EB. Finalmente se utilizaron 26 pM de muestra para llevar a cabo la secuenciación con el equipo Ion Personal Genome Machine (PGM™) System, el cual se basa en la detección de protones liberados durante la polimerización del DNA.

H. Ensamble del genoma e identificación de variantes por Winter Genomics Inc. (WG):

1. Datos obtenidos con el equipo Illumina GAllx:

Antes del ensamble, las lecturas de baja calidad se eliminaron usando el paquete ShortRead 1.8.0 [190]. El ensamble para cada cepa se realizó con el programa PE-Assembler 1.1 [191] y para cerrar los "gaps" se utilizó el programa IMAGE 2.1 [192]. Posteriormente, para alinear las lecturas a los contigs resultantes se utilizó el programa Bowtie 0.12.5 short read aligner [193]; las bases no

soportadas se removieron con el programa Biostrings 2.18.0 package. Los “contigs” se reordenaron a lo largo del genoma de la cepa *E. coli* K12 MG1655 [189] usando el programa Mauve 2.3.1 [195,195]. Finalmente, los contigs se compararon con el genoma de referencia utilizando los programas Mauve and Murasaki 1.68.6 [196]. Utilizando la delección del operón PTS como marcador, fue posible identificar correctamente cada cepa. Para los alineamientos de las cepas PB11 y PB12 contra la cepa JM101 se utilizó el programa BLAT v34 [197] y posteriormente, para identificar las variantes se utilizó el programa VarScan 1.2 software [198]. Las variantes ambiguas se filtraron utilizando un script de Perl. Para la mayoría de los análisis se utilizó el “cluster” del Instituto de Biotecnología, UNAM.

2. Datos obtenidos con el equipo Ion Torrent PGM:

Las lecturas de las cepas JM101 y PB13 obtenidas, se filtraron basándose en su calidad, utilizando el programa FASTX-Toolkit (Hannon Lab. FASTX-Toolkit.). Posteriormente las lecturas provenientes de la cepa PB13 se alinearon contra la secuencia de la cepa JM101 utilizando el programa bwa 0.6.2 [199]; para buscar variantes se utilizó el programa Samtools 0.1.18 [199]. Para la mayoría de los análisis se utilizó el “cluster” del Instituto de Biotecnología, UNAM.

3. Confirmación de las mutaciones puntuales:

Para confirmar los SNP’s reportados por WG, se resecuenciaron aquellos genes portadores. Para ello se amplificó mediante PCR los genes en cuestión y se enviaron a la unidad de síntesis y secuenciación del IBT, UNAM donde se realizó la resecuenciación por el método de Sanger.

I. Resecuenciación de la región cromosomal perdida:

Para la re-secuenciación se amplificó mediante PCR en las cepas JM101, PB12 y PB13 el fragmento cromosomal perdido durante el proceso de evolución adaptativa. Para ello se emplearon oligonucleótidos que empalman en los genes flanqueantes a dicha región (Tabla S11). El producto de PCR se envió a la unidad de síntesis y secuenciación del IBT, UNAM; de éste modo se determinó con exactitud la secuencia nucleotídica del re-arreglo cromosomal (Fig. S1).

J. Generación de todas las mutantes en las distintas cepas:

La generación de todas las mutantes se realizó mediante el método Datsenko-Wanner (Datsenko & Wanner 2000) [200]. Brevemente, se diseñaron y sintetizaron oligos de interrupción para el gen a interrumpir (Tabla S11), con los cuales se amplificó el gen que confiere resistencia a cloranfenicol, utilizando como templado el plásmido pKD3 y los oligos correspondientes. Los productos de PCR se

purificaron por medio de un kit para extracción de DNA en gel (GenJet PCR Purification kit / Thermo Scientific). Después de su purificación, se digiere el producto de PCR con la enzima HindIII a 37°C durante 45 minutos para eliminar completamente el plásmido pKD3. Los productos de PCR se purificaron con el kit (GenJet PCR Purification kit / Thermo Scientific). Al final el producto fue suspendido en agua, para favorecer la electroporación.

Las células transformantes JM101, PB11, PB12 y PB13 con el plásmido pKD46, se crecieron en medio LB agar con Cb a una concentración final de 100 µg/mL (Cb¹⁰⁰) a 30°C durante 24 horas. Se tomó una asada de las colonias y se crecieron en 25 ml de medio SOB con Cb¹⁰⁰ y 2.5 mL de arabinosa 1 M a una temperatura de 30°C hasta alcanzar una OD⁶⁰⁰ equivalente a 0.6. Las células se centrifugaron durante 12 minutos a 4600 rpm y 4°C para posteriormente ser lavadas con agua fría tres veces. Se toman 50 µL de células transformantes y de entre 5 y 10µL de producto de PCR. Las células se electroporan y se incuban a 30°C y 300 rpm durante 3 horas en 1mL de medio SOC. Posteriormente, se siembran 200 µL del cultivo en medio agar LB con Cm²⁵ y se incuban a 37°C para seleccionar las células transformadas. Cuando no exista crecimiento de colonias en el medio con antibiótico en las 24 horas siguientes de haberlas sembrado, se siembra el cultivo restante en varias cajas. Las células seleccionadas se crecen en medio LB sin antibiótico a 42°C durante por lo menos 14 horas para eliminar el plásmido pKD46. Para comprobar la pérdida del plásmido, las células se crecieron en medio LB agar con Cb¹⁰⁰. De las células resultantes se hace un respaldo en glicerol a -70°C. La comprobación de las mutantes se realiza mediante PCR utilizando oligonucleótidos específicos para sitios flanqueantes del gen interrumpido y comparando el tamaño de los productos de PCR con un control.

K. Extracción de RNA y síntesis de cDNA en las cepas JM101, JM101Δ*rppH*, PB11, PB11Δ*rppH*, PB12 y PB13:

Se realizaron cultivos por triplicado para la extracción de RNA en fermentadores de 1 litro, utilizando un volumen de trabajo de 750 mL de medio mínimo M9 con 2 g/L de glucosa (para el caso de las cepas Δ*rppH* se adicionó Cm 15 µg/mL; para el caso de la cepa PB11Δ*rppH*, se adiciono Km 30 µg/mL y Cm 15 µg/mL) a 37°C, 600 rpm y un flujo de aire de 1vvm (un volumen de aire / volumen de medio). El cultivo se inició a una OD₆₀₀ de 0.1 utilizando un preinóculo crecido en el mismo medio y con sus respectivos antibióticos, y se colectó 50 mL de cultivo, cuando este llegó a una OD₆₀₀ de 1.

El RNA total se aisló y se purificó utilizando una versión modificada del método de fenol caliente [8,15,108,117] y el cDNA se sintetizó mediante el kit RevertAid™ H Minus de síntesis de primera cadena de cDNA de acuerdo a las instrucciones del fabricante (Fermentas, Burlington, Canadá). Para cada reacción, se utilizaron 5 µg aproximadamente de RNA y 10 pmol/µL de una mezcla de oligonucleótidos reversos (Tabla S11) por cada gen determinado. Las condiciones de amplificación

fueron 10 min a 95°C, y un ciclo de dos pasos a 95°C por 15 s y 60°C por 60 s para un total de 40 ciclos. La secuencia de estos oligonucleótidos han sido previamente publicadas [8,15,108,117]. El tamaño de todos los amplicones fue de 101 pb. La concentración final de los primers en un volumen final de 15 µl fue de 0.2 µM. Se agregaron 5 ng de cDNA para cada gen. El RT-qPCR se realizó con el equipo ABI Prism 7300 (Perkin Elmer/Applied Biosystems, Foster City, CA) utilizando el kit MaximaR SYBR Green/ROX qPCR Master Mix 2X (Fermentas LifeSciences) y condiciones de reacción previamente descritas [15]. Los experimentos se realizaron por triplicado para cada gen y por cada cepa, obteniendo valores muy similares (la desviación estándar para todos los casos no fue mayor a 0.3). Se incluyó rutinariamente una reacción control sin templado para cada gen. Para evaluar la eficiencia del PCR, se realizaron curvas estándares y todos los ensayos de qPCR mostraron eficiencias altas de amplificación (90-100 %), los genes valores de R² > 0.9976 y pendientes entre -3.4 y -3.7. Todos los experimentos de RT-qPCR cumplen con los requisitos de MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) [201]. La técnica de cuantificación utilizada para analizar los datos fue el método 2-ΔΔC_Q descrito por Livak y Schmittgen (2001) [202]. Todos los datos se normalizaron utilizando el gene *ihfB* como control interno (Figura S2). Para cada valor de transcripción de cada gen analizado en la cepa silvestre JM101, se le asigno el valor de 1 y se utilizó como control para normalizar los datos; por lo tanto los datos se reportan como niveles de expresión relativos.

L. Determinación de las tasas de mutación en las diferentes cepas de *E. coli*.

El estimado de las tasas de mutación se basó en la determinación de mutantes espontáneas en el locus *rpsL* (mutantes resistentes a estreptomicina). Para ello se empleó el test de Luria-Delbrück con algunas modificaciones, el cual ya ha sido reportado ampliamente en la literatura [203]. Brevemente, 100 µL de 30 cultivos en LB saturados fueron mezclados y posteriormente se sembraron en medio LB agar con 200 µg/mL de estreptomicina. 2 mL de una dilución 1 x 10⁻⁶ de la mezcla de cultivos fue utilizada para estimar el número total de células. Para estimar la tasa de mutación se emplearon los métodos Ma-Sandri-Sarkar (MSS) y Lea Coulson [154]. La frecuencia de mutación también fue calculada.

M. Marcaje “in vivo” de las cepas JM101, PB11, PB12, PB13 y sus derivadas carentes del gen *fadD* con [¹⁴C] acetato y análisis de composición lipídica de membrana.

La composición lipídica de membrana de las diferentes cepas fueron determinadas por triplicado, utilizando [1-¹⁴C] acetato (60 mCi / mmol; Perkin Elmer). Para ello las distintas cepas fueron cultivadas a 37°C y 300 rpm en 1.5 mL de medio LB a una DO₆₀₀ inicial de 0.1 a partir de células

precultivadas en el mismo medio. Se adicionó 1 $\mu\text{Ci/mL}$ de $[1-^{14}\text{C}]$ acetato cuando las células alcanzaron una $\text{DO}_{600}=0.3$. Se cosecharon las células a las 24, 48 y 72 hr. y fueron resuspendidas en 100 μL de agua. Las especies lipídicas fueron extraídas de acuerdo al método de Bligh & Dyer (1959) [204]. La fase de cloroformo fue utilizada para el análisis lipídico por cromatografía en capa fina de alta resolución (high-performance TLC silica gel 60 plates - Merck), y acetato de etilo – hexano – ácido acético (60:40:5 [vol/vol/vol]) como fase móvil. Las cromatografías de capa fina en dos dimensiones fueron realizadas como se describe en Rudder et al. 1997 [205]. La radiactividad fue detectada utilizando el equipo Storm 820 PhosphorImager (Amersham Biosciences), y el análisis de las imágenes se llevó a cabo utilizando el programa ImageQuant TL (Amersham Biosciences).

X. Material suplementario

A. Tablas:

Tabla S1: Diferencias en la cepa JM101 con respecto a la cepa MG1655 (Mutaciones puntuales no sinónimas en la cepa JM101).

JM101 vs MG1655					
Mutaciones puntuales no sinónimas (WG)					
Gen	Descripción Básica	Mutaciones			
	Genes de función desconocida / Hipotéticos	MG1655	JM101	Nucleótido	Pos/Cambio
<i>yjdN</i>	Presunta proteína.	AGC	AaC	98	33 S-N
<i>ycbS</i>	Presunta proteína de membrana externa.	CAG	CcG	2036	679 Q-P
<i>paaA</i>	Presunta subunidad de la fenilacetato-CoA oxigenasa.	TAT	aAT	232	78 Y-N
<i>yfaS_1</i>	Presunta proteína.	GTC	aTC	1135	379 V-I
		CAG	Cat	2907	969 Q-H
<i>yjgl</i>	Presunta proteína.	CCT	tCT	154	52 P-S
<i>frwC</i>	Presunta permeasa de un sistema de transporte dependiente de PTS	TCA	TtA	848	283 S-L
	Metabolismo y Transporte	MG1655	JM101	Nucleótido	Pos/Cambio
<i>metC</i>	Cistationin β -liasa.	GCG	tCG	259	87 A-S
<i>fabH</i>	β -ketoacil-ACP sintasa. Enzima clave en el inicio de la síntesis de ácidos grasos.	CAG	tAG	214	72 Q-stop
<i>ycfS</i>	L,D-transpeptidasa.	CTG	CcG	539	180 L-P
<i>aaeB</i>	Proteína de un sistema de exporte de ácidos aromáticos.	ACG	cCG	148	50 T-P
<i>pabB</i>	Para-aminobenzoato sintasa.	CTC	CcC	35	12 L-P
<i>yedY</i>	Molibdoproteína periplásmica con actividad reductasa.	GCC	GaC	956	319 A-D
<i>murC</i>	UDP-N-acetilmuramato-alanin ligasa.	ATG	AcG	182	61 M-T
<i>oppA</i>	Proteína de unión a periplasma miembro de un sistema de transporte tipo ABC.	AAT	tAT	811	271 N-Y
<i>oppF</i>	Proteína miembro de un sistema de transporte de mureína del tipo ABC.	TCC	gCC	973	325 S-A
<i>acnA</i>	Aconitato hidratasa 1.	GGG	aGG	844	282 G-R
		AGC	gGC	1564	522 S-G
<i>puuP</i>	Transportador de putrecina.	TAT	TgT	329	110 Y-C

<i>nrfG</i>	Proteína que forma parte de una hemo-liasa.	GGC	Gat	239 / 240	80 G-D
<i>ppa</i>	Pirofosfatasa inorgánica. Parte del metabolismo del (p)ppGpp	CCG	tCG	67	23 P-S
<i>dacB</i>	D-alanil-D-alanin endopeptidasa / D-alanil-D-alanin carboxipeptidasa	GAT	tAT	781	261 D-Y
Genes de fagos		MG1655	JM101	Nucleótido	Pos/Cambio
<i>ydfU</i>	Proteína de profago (Qin).	CTG	CaG	626	209 L-Q
<i>intQ</i>	Presunta integrasa de profago (Qin).	CTT	CcT	820	274 L-P
Genes codificantes para moléculas de RNA		MG1655	JM101	Nucleótido	Pos/Cambio
<i>glnX</i>	tRNA de Glutamina.	G	a	35	NA
Genes con función de regulación		MG1655	JM101	Nucleótido	Pos/Cambio
<i>hha</i>	Proteína moduladora de la expresión de la hemolisina.	CCT	ttT	13 / 14	5 P-F
<i>rpoS</i>	Factor de transcripción, subunidad Sigma S (sigma 38) de la RNA polimerasa.	CAG	tAG	97	33 Q-stop
<i>glyQ</i>	Subunidad α de la glicil-tRNA sintetasa	GAG	GcG	143	48 E-A
<i>cytR</i>	Factor de transcripción implicado en el transporte y utilización de ribonucleósidos y desoxiribonucleosidos.	ACA	AtA	881	294 T-I
<i>tdcA</i>	Participa en el control de genes involucrados en el transporte y metabolismo de la treonina y de la serina (operón <i>tdc</i>)	TTG	TTt	597	199 L-F
		TGT	gGT	907	C-G
<i>glpR</i>	Factor de transcripción, represor del regulón de glicerol 3 fosfato.	GGT	GcT	164	55 G-A
<i>creC</i>	Proteína sensora cinasa miembro del sistema de dos componentes CreCB.	CGC	CcC	230	77 R-P
<i>narQ</i>	Proteína sensora cinasa miembro del sistema de dos componentes NarQX.	GCA	GtA	1382	461 A-V
Deleciones		MG1655	JM101	Nucleótido	Pos/Cambio
<i>lacZ</i>	Enzima β -galactosidasa; esta cepa carece del péptido alfa (aminoácidos 11-42).	Presente	Ausente	32-126	Péptido Omega

Tabla S2: Diferencias en la cepa JM101 con respecto a la cepa MG1655 (Mutaciones puntuales sinónimas en la cepa JM101).

JM101 vs MG1655					
Mutaciones puntuales sinónimas (WG)					
Gen	Descripción básica	Mutaciones			
	Genes de función desconocida / Hipotéticos	MG1655	JM101	Nucleótido	Pos/Cambio
<i>yafL</i>	Presunta lipoproteína.	TCG	TCt	36	12 S-S
<i>ylbE_1</i>	Presunta proteína.	GAA	GAg	114	38 E-E
<i>ygdH</i>	Proteína conservada.	GGC	GGa	465	155 G-G
<i>yjgL</i>	Presunta proteína.	TTT	TTc	57	19 F-F
<i>nrfG</i>	Proteína no esencial de función desconocida.	GGC	GGt	240	80 G-G
<i>creD</i>	Proteína de membrana interna de función desconocida.	TTG	TTa	63	21 L-L
<i>yphB</i>	Proteína conservada.	CCG	CCt	147	49 P-P
	Metabolismo y Transporte	MG1655	JM101	Nucleótido	Pos/Cambio
<i>gatC</i>	Permeasa de galactitol dependiente de PTS.	CAG	CAa	831	277 Q-Q
<i>astA</i>	Arginin succiniltransferasa.	GCC	GCt	879	293 A-A
<i>mgIC</i>	Proteína integral de membrana miembro del sistema de transporte de galactosa tipo ABC.	CAG	CAa	519	173 Q-Q
<i>ulaE</i>	Xilulosa 5-fosfato 3-epimerasa.	TCC	TCt	492	164 S-S
	Constituyentes celulares	MG1655	JM101	Nucleótido	Pos/Cambio
<i>fimD</i>	Proteína de membrana externa involucrada en el exporte y ensamble de fimbrias tipo 1.	ATC	ATt	1506	502 I-I
	Genes con función de regulación	MG1655	JM101	Nucleótido	Pos/Cambio
<i>glrK</i>	Proteína sensora cinasa miembro de un sistema de dos componentes (GlrKR).	ATC	ATt	1248	416 I-I
<i>phoP</i>	Proteína reguladora de respuesta miembro del sistema de dos componentes (PhoPQ).	ACT	ACc	468	156 T-T

Tabla S3: Diferencias en la cepa JM101 con respecto a la cepa MG1655 (Mutaciones en regiones no codificantes en la cepa JM101).

JM101 vs MG1655			
Regiones no codificantes (WG)			
Descripción	MG1655	JM101	Pos K12 MG1655
Región promotora del gen <i>lacI</i> (<i>lacIq</i>)	G	a	366796
Entre los genes <i>iap</i> y <i>cas2</i> .	C	g	2875935
Entre los genes <i>dppA</i> y <i>proK</i> .	G	a	3705905
Entre los genes <i>dppA</i> y <i>proK</i> .	C	a	3705969
Entre los genes <i>zinT</i> y <i>yodB</i> .	G	a	2040194
Entre los genes <i>asnT</i> y <i>yeeJ</i> .	T	c	2042665
Entre los genes <i>ppiC</i> y <i>yifO</i> .	C	t	3957956
Entre los genes <i>ykgR</i> y <i>eaeH</i> .	G	a	312766
Entre los genes <i>ychS</i> y <i>rttR</i> .	G	a	1286262
Región promotora del gen <i>pgpB</i> (fosfatidilglicerofosfatasa)	C	t	1337338
Entre los genes <i>yjhV</i> y <i>fecE</i> .	T	g	4508260
Entre los genes <i>sgcX</i> y <i>yjhP</i> .	T	g	4530431
Entre los genes <i>glyY</i> y <i>yjeV</i> .	C	t	4390747
Región promotora del gen <i>hfq</i> (RNA con funciones de regulación)	C	t	4398276
Región promotora del gen <i>iclR</i> (Represor de genes involucrados en el ciclo del glioxilato)	T	c	4221660

Tabla S4: Diferencias entre la cepa PB12 y JM101. A) Se muestran las 23 mutaciones puntuales no sinónimas en genes estructurales presentes en la cepa PB12 al compararla con la cepa silvestre JM101. Los asteriscos indican las 21 mutaciones detectadas por ambos métodos/compañías, Roche NimbleGen Inc. (RN) y Winter Genomics (WG). Adicionalmente 16 mutaciones sinónimas en diferentes genes también fueron detectadas por ambos métodos (Tabla S3). Mutaciones en 14 genes fueron corroborados mediante secuenciación por Sanger, los cuales se indican con el símbolo +. B) Se muestran los genes ausentes en la cepa PB12, producto de la deleción chromosomal ocurrida durante el proceso de evolución adaptativa.

JM101 vs PB12					
A) Mutaciones Puntuales no sinónimas detectadas por RN y WG					
Gen	Descripción básica	Mutaciones			
	Genes de función desconocida / Hipotéticos	JM101	PB12	Nucleótido	Pos/Cambio
<i>gfcD</i> *	Lipoproteína (Hipotético).	GTA	cTA	916	306 V-L
<i>yafV</i> *	Presunta hidrolasa de unión a NAD(P) de la familia de las amidasas.	CGC	CaC	452	151 R-H
+ <i>ytfR</i> ^(WG)	Presunto componente del transportador de galactosa	GTC	GcC	602	201 V-A
+ <i>rssA</i> *	Proteína hipotética (Presunta fosfolipasa)	CGC	CaC	773	258 R-H
+ <i>yjjU</i> *	Proteína hipotética (Presunta Fosfolipasa)	ACT	gCT	338	113 T-A
<i>ydiQ</i> *	Presunta subunidad de una flavo proteína.	GTG	GcG	218	73 V-A
+ <i>ypdA</i> *	Presunta proteína sensora cinasa miembro de un sistema de dos componentes	GCG	tCC	598	200 A-S
	Metabolismo y transporte	JM101	PB12	Nucleótido	Pos/Cambio
+ <i>actP</i> *	Permeasa de Acetato/glicolato.	GTA	GcA	731	244 V-A
<i>arnT</i> *	L-Ara4N transferasa; cataliza la adición de este grupo a lípidos A.	TAC	TgC	1193	398 Y-C
+ <i>chbC</i> *	Proteína de membrana miembro del sistema de transporte de quitobiosa dependiente de PTS.	GCG	GtG	347	116 A-V
+ <i>csgF</i> ^(WG)	Proteína miembro del complejo de ensamble y secreción de fibras "curli"	AAU	AgU	89	30 N-S
+ <i>dgoT</i> *	Probable transportador de galactonato.	GAT	GgT	653	218 D-G
<i>dhaM</i> *	Subunidad M de la enzima dihidroxiacetona cinasa.	TGG	TGa	1038	346 W-stop

<i>+dppF*</i>	Componente de unión a ATP de un transportador de dipeptidos.	CAC	CgC	680	227 H-R
<i>fdhD*</i>	Proteína involucrada en la actividad formato deshidrogenasa.	AGT	gGT	652	218 S-G
<i>fimH*</i>	Subunidad menor de la adhesina del complejo de fimbrias específica a la D-manosa.	ACT	gCT	535	179 T-A
<i>glpT*</i>	Principal sistema de transporte de glicerol 3 fosfato.	CCG	CtG	416	139 P-L
<i>rfe*</i>	Undecaprenil fosfato α -N-acetilglucosaminil transferasa .	GCC	aCC	238	80 A-T
<i>+sucA*</i>	Componente E1 del complejo de la oxoglutarato deshidrogenasa.	GAT	GAa	372	124 D-E
Genes con función de regulación		JM101	PB12	Nucleótido	Pos/Cambio
<i>+arcB*</i>	Proteína de sensora cinasa miembro de un sistema de dos componentes.	TAC	TgC	212	71 Y-C
<i>+barA*</i>	Proteína de sensora cinasa miembro de un sistema de dos componentes.	TTC	cTC	1096	366 F-L
<i>+rna*</i>	Fosfodiesterasa de RNA.	GCC	aCC	268	90 A-T
<i>+rpoD*</i>	Subunidad sigma 70 de la RNA polimerasa.	GTT	aTT	1744	582 V-I
B) Delección cromosomal (RN y WG)					
Gen	Descripción básica	Mutaciones			
<i>ptsP</i>	Proteína de un complejo PTS involucrado en el metabolismo del nitrógeno.	Presente	Ausente	---	---
<i>rppH</i>	Pirofosfohidrolasa de RNA implicada en la degradación de los mRNA's	Presente	Ausente	---	---
<i>ygdT</i>	Proteína hipotética.	Presente	Ausente	---	---
<i>mutH</i>	Endonucleasa miembro del sistema MMR de reparación de DNA.	Presente	Ausente	---	---
<i>ygdQ</i>	Presunto transportador	Presente	Ausente	---	---
<i>ygdR</i>	Presunta proteína	Presente	Ausente	---	---
<i>tas</i>	Presunta reductasa que actúa en condiciones de estrés por falta de carbohidratos.	Presente	Ausente	---	---
<i>lplT</i>	Transportador de lisofosfatidiletanolamina	Presente	Ausente	---	---
<i>aas</i>	2-acilglicerofosfoetanolamin aciltransferasa / acil-ACP sintetasa.	Presente	Ausente	---	---
<i>omrA</i>	sRNA involucrado en la regulación de la composición de proteínas de membrana externa.	Presente	Ausente	---	---
<i>omrB</i>	sRNA involucrado en la regulación de la composición de proteínas de membrana externa.	Presente	Ausente	---	---
<i>galR</i>	Factor de transcripción del regulón Gal.	Presente	Ausente	---	---

Tabla S5: Diferencias entre la cepa JM101 y PB12. Se muestran las 16 mutaciones puntuales sinónimas en la cepa PB12 al compararla con la cepa silvestre JM101.

JM101 vs PB12					
Mutaciones puntuales sinónimas (RN y WG)					
Gen	Descripción	Mutaciones			
	Genes de función desconocida / Hipotéticos	JM101	PB12	Nucleótido	Pos/Cambio
<i>yddW</i>	Presunta lipoproteína.	CAA	CAg	150	50 Q-Q
<i>yjgI</i>	Presunta proteína.	GTC	GTt	261	87 V-V
<i>sgcX</i>	Presunta endoglucanasa dependiente de Zinc.	ACA	ACg	912	304 T-T
	Metabolismo y transporte	JM101	PB12	Nucleótido	Pos/Cambio
<i>yaaH</i>	Transportador de Acetato.	GGC	GGt	138	46 G-G
<i>zitB</i>	Transportador de Zinc.	AAG	AAa	750	250 K-K
<i>pgl</i>	6-Fosfogluconolactonasa-	GGC	GGt	531	177 G-G
<i>moeA</i>	Molibdenotransferasa.	CAG	CAa	570	190 Q-Q
<i>rspB</i>	Oxidoreductasa dependiente de Zinc.	GCC	GCt	507	169 A-A
<i>chbF</i>	Diacetilquitobiosa 6 fosfato hidrolasa.	CTG	CTa	1308	436 L-L
<i>chbC</i>	Proteína de membrana miembro del sistema de transporte de quitobiosa dependiente de PTS.	GTA	GTg	33	11 V-V
<i>lptC</i>	Componente del sistema de transporte de lipopolisacáridos.	ACG	ACa	393	131 T-T
<i>yjeS</i>	Epoxiqueuosin reductasa.	ATT	ATc	36	28 I-I
<i>hflK</i>	Proteína de membrana interna involucrada con la proteólisis de la proteína cII del fago lambda.	AGC	AGt	345	115 S-S
<i>nanM</i>	N-acetilneuraminil mutarotasa.	GCT	GCc	762	224 A-A
	Genes con función de regulación	JM101	PB12	Nucleótido	Pos/Cambio
<i>nemR</i>	Factor de transcripción involucrado en la degradación de N-etilmaleimida.	CCG	CCA	123	41 P-P
<i>atoS</i>	Regulador de respuesta de un sistema de dos componentes implicado en la regulación de genes del metabolismo de ácidos grasos de cadena corta.	GAA	GAg	975	325 E-E

Tabla S6: Diferencias entre la cepa JM101 y PB12. Se muestran las mutaciones en regiones no codificantes en la cepa PB12 al compararla con la cepa silvestre JM101.

JM101 vs PB12			
Regiones no codificantes (WG)			
Descripción	JM101	PB12	Pos K12 MG1655
Entre los genes <i>tomB</i> <i>acrB</i> .	A	g	480476
Entre los genes <i>ychE</i> y G0-10697.	T	c	1298686
Entre los genes <i>rluB</i> y <i>REPt109/btuR</i> .	C	t	1325761
Entre los genes <i>ydeT</i> y <i>yneL</i> .	T	g	1588198
Región promotora del gen <i>hydN</i> , el cual codifica para una presunta proteína transportadora de electrones.	C	a	2836155
Presunta región promotora del gen <i>mocA</i> el cual codifica para una molibdopterin citidiltransferasa específica.	C	t	3013976
Región promotora del gen <i>yjhl</i> , el cual codifica para el elemento de fago KpLE2.	T	g	4523835

Tabla S7: Diferencias entre la cepa JM101 y PB13. Se muestran las mutaciones puntuales no sinónimas en la cepa PB13 al compararla con la cepa silvestre JM101.

JM101 vs PB13					
A) Mutaciones puntuales no sinónimas (WG)					
Gen	Descripción Básica	Mutaciones			
	Genes de función desconocida / Hipotéticos	JM101	PB13	Nucleótido	Pos/Cambio
<i>wzyE</i>	Presunta polimerasa encargada de sintetizar cadenas de polisacáridos lineales (ECA).	GCG	GtG	221	74 A-V
<i>ydaF</i>	Presunta proteína.	ACT	AaT	116	39 T-N
<i>ydcT</i>	Presunto transportador de espermidina / putrecina	CCG	CtG	113	38 P-L
<i>ydiM</i>	Presunto transportador.	TTC	cTC	982	328 F-L
	Metabolismo y Transporte	JM101	PB13	Nucleótido	Pos/Cambio
<i>bglA</i>	6-fosfo-β-glucosidasas in <i>E. coli</i> .	CCG	CtG	839	280 P-L
<i>deoB</i>	Fosfodesoxiribomutasa.	GAC	GgC	848	283 D-G
<i>dhaM</i>	Subunidad M de la enzima dihidroxiacetona cinasa.	TGg	TGa	1038	346 W-stop
<i>glpT</i>	Principal sistema de transporte de glicerol 3 fosfato.	CCG	CtG	416	139 P-L
<i>ispG</i>	1-hidroxi-2-metil-2-(E)-butenil 4-difosfato sintasa.	GTC	GcC	749	250 V-A
<i>mepA</i>	Murein DD-endopeptidasa.	CGT	CaT	197	66 R-H
<i>msbA</i>	Transportador de lipopolisacáridos de tipo ABC.	ACG	gCG	367	123 T-A
<i>xapB</i>	Transportador de xantosina.	ACC	gCC	1051	351 T-A
	Constituyentes celulares	JM101	PB13	Nucleótido	Pos/Cambio
<i>fliN</i>	Proteína implicada en el movimiento flagelar	GAT	GgT	371	124 D-G
	Genes con función de regulación	JM101	PB13	Nucleótido	Pos/Cambio
<i>polA</i>	DNA Polimerasa I (Pol I)	GCG	GtG	2084	695 A-V
B) Delección cromosomal (WG)					
Gen	Descripción Básica	Mutaciones			
<i>ptsP</i>	Proteína de un complejo PTS involucrado en el metabolismo del nitrógeno.	Presente	Ausente	---	---
<i>rppH</i>	Pirofosfohidrolasa de RNA implicada en la degradación de los mRNA's	Presente	Ausente	---	---
<i>ygdT</i>	Proteína hipotética.	Presente	Ausente	---	---
<i>muth</i>	Endonucleasa miembro del sistema MMR de reparación de DNA.	Presente	Ausente	---	---

<i>ygdQ</i>	Presunto transportador	Presente	Ausente	---	---
<i>ygdR</i>	Presunta proteína	Presente	Ausente	---	---
<i>tas</i>	Presunta reductasa que actúa en condiciones de estrés por falta de carbohidratos.	Presente	Ausente	---	---
<i>lplT</i>	Transportador de lisofosfatidiletanolamina	Presente	Ausente	---	---
<i>aas</i>	2-acilglicerofosfoetanolamin aciltransferasa / acil-ACP sintetasa.	Presente	Ausente	---	---
<i>omrA</i>	sRNA involucrado en la regulación de la composición de proteínas de membrana externa.	Presente	Ausente	---	---
<i>omrB</i>	sRNA involucrado en la regulación de la composición de proteínas de membrana externa.	Presente	Ausente	---	---
<i>galR</i>	Factor de transcripción del regulón Gal.	Presente	Ausente	---	---

Tabla S8: Diferencias entre la cepa JM101 y PB13. Se muestran las mutaciones puntuales sinónimas en la cepa PB13 al compararla con la cepa silvestre JM101.

JM101 vs PB13					
Mutaciones puntuales sinónimas (WG)					
Gen	Descripción Básica	Mutaciones			
Genes de función desconocida / Hipotéticos		JM101	PB13	Nucleótido	Pos/Cambio
<i>ygeH</i>	Presunto regulador transcripcional	GCT	GCc	732	244 A-A
<i>ygfF</i>	Presunta oxidoreductasa	GTC	GTt	522	174 V-V
Metabolismo y Transporte		JM101	PB13	Nucleótido	Pos/Cambio
<i>aroH</i>	DHAP sintasa.	CGT	CGc	24	8 R-R
<i>fadE</i>	Acil-CoA deshidrogenasa.	GCG	GCa	270	90 A-A
<i>fliH</i>	Proteína citoplásmica implicada en el exporte de proteínas flagelares.	TTG	cTG	448	150 L-L
<i>trpC</i>	Enzima bifuncional, fosforibosil-antranilato isomerasa / indol-3-glicerol fosfato sintasa.	GGT	GGc	1005	335 G-G
<i>ilvI</i>	Subunidad de la enzima bifuncional aceto-hidroxi-butanoato sintasa /acetolactato sintasa.	AAC	AAt	816	272 N-N
Reparación de DNA		JM101	PB13	Nucleótido	Pos/Cambio
<i>recN</i>	Proteína implicada en la reparación de DNA de doble hebra.	GCG	GCa	1641	547 A-A

Tabla S9: Resultados del análisis estadístico para determinar las tasas y frecuencias de mutación en las cepas JM101, PB11, PB11ΔReg, PB12 y PB13.

Cepas	MSS-MLE				
	Tasa de Mutación (10^7)	Límite superior	Límite inferior	Diferencia Superior	Diferencia Inferior
JM101	0.001	0.0021	2.00E-04	0.0011	8.00E-04
PB11	6.00E-04	0.0014	1.00E-04	8.00E-04	5.00E-04
PB11ΔReg	0.2828	0.3712	0.2036	0.0885	0.0792
PB12	0.2106	0.2527	0.1715	0.0421	0.0391
PB13	1.0532	1.1897	0.923	0.1365	0.1302

Cepas	Método Lea Coulson				
	Tasa de Mutación (Medianas)	Límite superior	Límite inferior	Diferencia Superior	Diferencia Inferior
JM101	0	0	0	0	0
PB11	0	0	0	0	0
PB11ΔReg	0.2515	0.2924	0.2045	0.0409	0.047
PB12	0.2135	0.231	0.1909	0.0175	0.0226
PB13	0.9436	1.1025	0.8878	0.1589	0.0558

Cepas	Frecuencia de Mutación				
	Frecuencia de Mutación (10^7)	Límite superior	Límite inferior	Diferencia Superior	Diferencia Inferior
JM101	0	0	0	0	0
PB11	0	0	0	0	0
PB11ΔReg	0.58	0.7576	0.4545	0.1776	0.1255
PB12	0.8	0.8854	0.6875	0.0854	0.1125
PB13	4.9172	5.5952	4.5469	0.678	0.3703

Tabla S10: Valores relativos de transcrito de las cepas PB11, PB12, JM101 Δ rppH y PB11 Δ rppH comparados con la cepa JM101. Para cada gen en todas las cepas, el nivel de transcripción del gen de la cepa silvestre JM101 se empleó como control para normalizar los datos, teniendo JM101 siempre un valor de RT-qPCR de uno. De esta forma los datos de éste estudio se reportan como valores de expresión relativa, comparados al nivel de expresión de JM101. Los resultados presentados son el promedio de al menos tres mediciones independientes de los valores de expresión para cada gen. Los valores se obtuvieron a partir de diferentes cDNA's generados al menos de dos muestras de biorreactores independientes. Los valores de expresión por RT-qPCR obtenidos para cada gen difieren menos del 30%. La desviación estándar para cada gen se muestra como: \pm (valor). * Indica valores previamente reportados [8,108,117].

Gene	PB11*	PB11 Δ rppH	PB12*	PB13	JM101 Δ rppH
Glucólisis					
<i>aceE</i>	0.4 \pm 0.02	1.32 \pm 0.21	1.4 \pm 0.02	2.3 \pm 0.1	1.57 \pm 0.23
<i>aceF</i>	0.6 \pm 0.0	1.00 \pm 0.12	1.2 \pm 0.0	2.3 \pm 0.2	2.66 \pm 0.51
<i>eno</i>	0.5 \pm 0.1	1.28 \pm 0.41	0.5 \pm 0.14	0.5 \pm 0.1	0.89 \pm 0.01
<i>fbaA</i>	0.9 \pm 0.0	1.46 \pm 0.18	1.1 \pm 0.05	1.3 \pm 0.0	0.82 \pm 0.08
<i>gapA</i>	0.4 \pm 0.1	3.76 \pm 0.67	1.7 \pm 0.2	2.3 \pm 0.1	6.98 \pm 1.17
<i>glk</i>	1.0 \pm 0.0	6.04 \pm 0.03	2.2 \pm 0.1	2.8 \pm 0.2	3.43 \pm 0.72
<i>gpmA</i>	0.9 \pm 0.0	2.46 \pm 0.13	1.8 \pm 0.0	1.6 \pm 0.0	3.07 \pm 0.21
<i>lpdA</i>	1.1 \pm 0.1	3.16 \pm 0.36	2.0 \pm 0.1	2.0 \pm 0.1	2.33 \pm 0.00
<i>pfkA</i>	0.3 \pm 0.0	1.48 \pm 0.11	0.5 \pm 0.0	1.0 \pm 0.0	1.49 \pm 0.11
<i>pgi</i>	1.0 \pm 0.1	7.71 \pm 0.91	6.6 \pm 0.3	5.9 \pm 0.2	3.47 \pm 0.10
<i>pgk</i>	0.7 \pm 0.0	1.08 \pm 0.11	1.2 \pm 0.0	1.4 \pm 0.0	1.24 \pm 0.20
<i>ppc</i>	0.6 \pm 0.05	1.97 \pm 0.40	0.8 \pm 0.0	1.0 \pm 0.1	3.81 \pm 0.27
<i>pykA</i>	0.4 \pm 0.0	0.83 \pm 0.03	1.0 \pm 0.0	1.3 \pm 0.1	1.16 \pm 0.12
<i>pykF</i>	0.8 \pm 0.0	2.00 \pm 0.26	0.9 \pm 0.0	1.2 \pm 0.0	0.89 \pm 0.31
<i>tpiA</i>	0.5 \pm 0.0	8.83 \pm 2.61	1.8 \pm 0.0	1.7 \pm 0.0	4.84 \pm 0.12
Gene	PB11*	PB11 Δ rppH	PB12*	PB13	JM101 Δ rppH
TCA					
<i>acnA</i>	5.2 \pm 0.6	10.80 \pm 1.63	5.7 \pm 0.2	3.3 \pm 0.3	2.88 \pm 0.31
<i>acnB</i>	1.7 \pm 0.36	24.80 \pm 2.18	1.7 \pm 0.42	1.2 \pm 0.0	1.76 \pm 0.06
<i>fumA</i>	3.3 \pm 0.67	24.07 \pm 3.41	3.6 \pm 0.67	2.1 \pm 0.5	2.62 \pm 0.72
<i>gltA</i>	0.8 \pm 0.1	6.68 \pm 0.13	1.3 \pm 0.2	0.8 \pm 0.0	1.73 \pm 0.31
<i>icdA</i>	1.0 \pm 0.1	4.66 \pm 0.78	1.9 \pm 0.0	1.5 \pm 0.0	1.03 \pm 0.11

<i>mdh</i>	1.2 ± 0.1	3.64 ± 0.32	1.8 ± 0.5	1.6 ± 0.4	1.93 ± 0.16
<i>sdhA</i>	0.9 ± 0.0	5.96 ± 0.67	2.0 ± 0.3	1.2 ± 0.1	4.75 ± 0.98
<i>sdhB</i>	0.8 ± 0.0	8.05 ± 0.86	1.7 ± 0.0	1.0 ± 0.1	4.40 ± 0.86
<i>sucA</i>	1.4 ± 0.1	5.42 ± 1.39	1.7 ± 0.1	2.2 ± 0.1	3.28 ± 0.20
<i>sucB</i>	0.8 ± 0.0	8.28 ± 1.91	1.4 ± 0.1	1.4 ± 0.0	3.31 ± 0.62
<i>sucC</i>	0.9 ± 0.1	7.41 ± 1.21	1.6 ± 0.0	1.6 ± 0.1	1.99 ± 0.43
<i>sucD</i>	1.0 ± 0.2	7.16 ± 0.11	1.4 ± 0.3	1.3 ± 0.3	1.89 ± 0.32
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Gluconeogénesis, shunt del glioxilato y poxB					
<i>aceA</i>	12.0 ± 1.2	24.20 ± 3.16	1.9 ± 0.2	2.4 ± 0.5	1.66 ± 0.38
<i>aceB</i>	15.6 ± 2.5	28.65 ± 5.91	3.7 ± 0.3	1.6 ± 0.3	2.54 ± 0.09
<i>acs</i>	5.6 ± 0.745	21.82 ± 3.20	8.0 ± 0.136	2.5 ± 0.3	1.83 ± 0.19
<i>glcB</i>	11.7 ± 1.4	12.62 ± 0.93	3.7 ± 0.5	2.1 ± 0.3	1.04 ± 0.28
<i>maeB</i>	1.2 ± 0.2	10.06 ± 1.19	2.5 ± 0.03	1.2 ± 0.0	1.18 ± 0.26
<i>pckA</i>	0.8 ± 0.06	2.81 ± 0.89	2.3 ± 0.02	1.8 ± 0.0	0.83 ± 0.21
<i>poxB</i>	4.2 ± 0.35	5.30 ± 0.56	5.7 ± 0.48	2.0 ± 0.4	3.3 ± 0.48
<i>ppsA</i>	3.7 ± 0.21	5.57 ± 1.34	2.4 ± 0.09	0.7 ± 0.1	7.34 ± 2.08
<i>sfcA</i>	1.9 ± 0.33	3.72 ± 0.28	1.7 ± 0.1	2.0 ± 0.0	2.56 ± 0.85
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Regulón Gal					
<i>galE</i>	38.8 ± 6.8	40.25 ± 1.97	46.6 ± 6.5	124.8	0.53 ± 0.12
<i>galK</i>	39.0 ± 2.3	43.78 ± 3.43	48.21 ± 4.0	97.1 ± 26.9	2.12 ± 0.09
<i>galM</i>	8.1 ± 3.94	2.67 ± 0.85	3.3 ± 0.7	5.3 ± 3.1	0.96 ± 0.18
<i>galP</i>	12.4 ± 1.08	25.11 ± 0.98	13.1 ± 1.86	38.02 ± 14.8	1.98 ± 0.36
<i>galR</i>	3.2 ± 0.1	5.2 ± 0.10	1.2 ± 0.4	ND	2.04 ± 0.25
<i>galS</i>	4.9 ± 0.47	23.48 ± 0.07	3.2 ± 0.52	2.3 ± 0.7	4.14 ± 0.35
<i>galT</i>	35.6 ± 2.37	64.34 ± 5.20	42.5 ± 4.0	65.8 ± 26.4	1.10 ± 0.12
<i>galU</i>	1.7 ± 0.3	2.36 ± 0.22	2.1 ± 0.36	2.3 ± 0.8	1.27 ± 0.01
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Respiración					
<i>cyoA</i>	0.6 ± 0.01	4.33 ± 0.38	1.4 ± 0.07	2.9 ± 0.5	2.16 ± 0.07
<i>cyoB</i>	0.3 ± 0.0	3.11 ± 0.04	1.4 ± 0.1	1.5 ± 0.2	2.42 ± 0.58
<i>cyoC</i>	0.3 ± 0.0	2.77 ± 0.55	1.2 ± 0.2	1.6 ± 0.3	2.93 ± 0.19
<i>cyoD</i>	0.3 ± 0.02	3.05 ± 0.03	1.1 ± 0.1	1.4 ± 0.0	1.90 ± 0.43
<i>nuoA</i>	0.4 ± 0.0	9.67 ± 1.98	2.4 ± 0.1	2.2 ± 0.0	3.53 ± 0.54
<i>nuoF</i>	0.4 ± 0.1	6.84 ± 0.28	1.4 ± 0.1	1.2 ± 0.1	4.68 ± 0.11

<i>nuoN</i>	0.4 ± 0.02	1.55 ± 0.18	1.0 ± 0.08	0.8 ± 0.2	0.44 ± 0.02
<i>ubiE</i>	0.4 ± 0.0	12.01 ± 1.26	1.6 ± 0.0	1.5 ± 0.0	4.39 ± 1.31
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Metabolismo de ppGpp					
<i>gpp</i>	1.4 ± 1.2	4.77 ± 0.93	5.2 ± 0.3	5.8 ± 0.9	9.17 ± 1.65
<i>ndk</i>	0.4 ± 0.0	11.08 ± 0.43	3.1 ± 0.1	0.9 ± 0.1	3.49 ± 0.52
<i>ppa</i>	1.0 ± 0.0	9.29 ± 0.32	6.1 ± 1.4	4.5 ± 0.6	7.42 ± 0.33
<i>relA</i>	1.3 ± 0.0	1.51 ± 0.00	1.6 ± 0.2	1.5 ± 0.2	3.13 ± 0.20
<i>spoT</i>	1.1 ± 0.2	7.36 ± 0.82	4.8 ± 0.2	3.3 ± 0.3	6.00 ± 0.77
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Operón PTS					
<i>crr</i>	0	0	0	0	2.46 ± 0.37
<i>ptsH</i>	0	0	0	0	1.07 ± 0.12
<i>ptsI</i>	0	0	0	0	1.34 ± 0.14
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Genes dentro del fragmento cromosomal perdido					
<i>galR</i>	3.2 ± 0.1	5.2 ± 0.10	1.2 ± 0.4	ND	2.04 ± 0.25
<i>lplT</i>	3.95 ± 0.53	2.83 ± 0.32	0	0	0.95 ± 0.12
<i>mutH</i>	2.76 ± 0.02	30.48 ± 0	0	0	16.49 ± 0.22
<i>rppH</i>	1.84 ± 0.20	0	0	0	0
<i>ptsP</i>	1.8 ± 0.2	0.30 ± 0.06	0	0	0.55 ± 0.07
<i>tas</i>	1.31 ± 0.07	12.80 ± 2.44	0	0	1.07 ± 0.21
<i>ygdQ</i>	8.86 ± 1.86	6.10 ± 1.58	0	0	1.92 ± 0.06
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Genes con mutaciones puntuales en la cepa PB12					
<i>arcB</i>	2.3 ± 0.7	1.23 ± 0.19	1.3 ± 0.72	2.1 ± 0.0	0.99 ± 0.06
<i>barA</i>	4.14 ± 0.91	2.30 ± 0.53	3.46 ± 0.46	ND	3.25 ± 0.09
<i>rna</i>	0.89 ± 0.18	8.03 ± 0.00	3.40 ± 0.75	ND	2.06 ± 0.42
<i>rssA</i>	4.89 ± 1.27	6.07 ± 0.95	4.04 ± 0.43	ND	6.45 ± 0.31
<i>rpoD</i>	1.0 ± 0.2	1.18 ± 0.24	0.9 ± 0.2	ND	1.0 ± 0.2
<i>rpoS</i>	2.9 ± 0.86	4.62 ± 0.34	2.2 ± 0.2	1.6 ± 0.3	2.03 ± 0.07
<i>yjjU</i>	0.86 ± 0.26	4.31 ± 0.35	8.51	ND	1.44 ± 0.03
<i>ypdA</i>	1.59 ± 0.49	1.36 ± 0.11	1.38 ± 0.15	ND	1.20 ± 0.31
Gene	PB11*	PB11ΔrppH	PB12*	PB13	JM101ΔrppH
Reguladores					
<i>arcA</i>	1.2 ± 0.02	3.86 ± 0.60	1.3 ± 0.09	1.3 ± 0.0	1.68 ± 0.49
<i>arcB</i>	2.3 ± 0.7	1.23 ± 0.19	1.3 ± 0.72	ND	0.99 ± 0.06

<i>barA</i>	4.14 ± 0.91	2.30 ± 0.53	3.46 ± 0.46	ND	3.25 ± 0.09
<i>crp</i>	0.8 ± 0.02	2.22 ± 0.64	1.0 ± 0.05	0.8 ± 0.0	1.38 ± 0.01
<i>cyaA</i>	1.4 ± 0.1	3.62 ± 0.68	1.8 ± 0.1	2.4 ± 0.1	2.37 ± 0.78
<i>fadR</i>	0.6 ± 0.031	3.34 ± 0.19	1.6 ± 0.082	2.4 ± 0.0	3.52 ± 0.03
<i>fruR</i>	1.3 ± 0.002	4.34 ± 0.34	1.5 ± 0.07	ND	1.41 ± 0.02
<i>galR</i>	3.2 ± 0.1	5.2 ± 0.10	1.2 ± 0.4	ND	2.04 ± 0.25
<i>iclR</i>	0.4 ± 0.03	2.62 ± 0.39	1.4 ± 0.075	1.5 ± 0.0	0.81 ± 0.20
<i>ihfA</i>	1.8 ± 0.04	2.07 ± 0.22	1.7 ± 0.06	1.6 ± 0.0	1.97 ± 0.32
<i>ihfB</i>	1.0 ± 0.00	1.0 ± 0.00	1.0 ± 0.00	1.0 ± 0.00	1.0 ± 0.00
<i>mlc</i>	1.1 ± 0.003	10.66 ± 2.07	1.7 ± 0.061	1.3 ± 0.3	2.37 ± 0.49
<i>rpoD</i>	1.0 ± 0.2	1.18 ± 0.22	0.9 ± 0.2	1.1 ± 0.0	1.0 ± 0.2
<i>rpoS</i>	2.9 ± 0.86	4.62 ± 0.34	2.2 ± 0.2	1.6 ± 0.3	2.03 ± 0.07

Tabla S11.- Se muestran las secuencias nucleotídicas de los oligos empleados en este estudio con diversos fines.

Oligonucleótidos empleados en este estudio		
Gen	Oligonucleótido	Sentido (5' – 3')
A. Oligonucleótidos usados para resecuenciar por el método de Sanger y confirmar inactivaciones y deleciones.		
<i>aas-lpIT</i>	aas-lpITa	TTA CTG TCG CTC GCG TTC TT
	aas-lpITb	TAT TAA TGA CGG CGC TGC CA
<i>actP</i>	actPa	CCT CCG GCT ATG ACG GCT TA
	actPb	GCA ACC ACC GCG AGG ATA GT
<i>arcA</i>	arcAa	CCA CGA CCA AGC TAA TGA TG
	arcAb	ACA ACG GAC GAT GAG TTA CG
<i>arcB</i>	arcBa	CAT CCG CTC AGA ATT ACG CCA A
	arcBb	CAT ACG CCA CCG GAA GTG ATC TA
	arcBc	GCA AGG AAG CTG GTG AAA TCA AC
	arcBd	GGT CCA GCT TGA TAA TCA ACC G
<i>barA</i>	barAa	AAT CCG GGG ATA AAT AGC GT
	barAb	GCA AAA TAG CCT GTT GTG CT
	barAc	TGT CGC TGG CTG CTT ATC AC
	barAd	CAG GCT TGT AGC GCA ACA AC
<i>chbC</i>	chbCa	ATT CGC CTC GAC GCC TCA AC
	chbCb	GCC GCT TCG ACG GAA CCA TA
<i>dgoT</i>	dgoTa	CGT AGG TTC TCG CGT GAC TT
	dgoTb	CAG CGC CAT CAG GCA CAT AA
<i>dhaM</i>	dhaMa	TCG GTG AAC GCA GTA TTG GT
	dhaMb	TTG CCG GAT GAC ATC AGA AC
<i>dppF</i>	dppFa	CAG ATG TCG CGC TGA AGA AC
	dppFb	CTG GCT GAT TCT GCT GTA CG
<i>fadD</i>	fadDa	GTT GCG ATG ACG ACG AAC AC
	fadDb	GGC GTC TGA CGA CTG ACT TA
<i>galR</i>	galRa	GTG GTC TGC ACG ACG ACT CT
	galRb	GGT AGC GCA GTC TGG CGT AT
<i>glpQ</i>	glpQa	GCA GCA TTC TGG CGG TTA TC
	glpQ b	CCG TTC CTC TCG AAC CAA CT
<i>glpT</i>	glpTa	TCT TAA CAT CGC GGC AAC TC
	glpT b	CAT GAT CGC CAT GCT AAG GT
<i>mdlB</i>	mdlBa	GAT GCA GCA CGG ACA TAT CG
	mdlBb	CGC ACT ATG ACA GTG CAC AG
<i>ptsP</i>	ptsPa	GGT CAG ACA GGT GGT GTC AT
	ptsPb	GAT ACA GCG GAT CGG CCA TA
<i>rna</i>	rna-a	CGC CAT GTG GAA CTA ATA CG
	rna-b	GTT GCT CCA GCA GAA TAT CG

<i>rpoD</i>	rpoDa	ACC TAC GCA ACC TGG TGG AT
	rpoDb	TCC AAC GGT GCA AGC CAA TG
<i>rpoS</i>	rpoSa	TAA CGG AAC CAG TTC AAC ACG CT
	rpoSb	ACG GTC TGA AAA CCT GGA CAG TGT
<i>rppH</i>	rppHa	AAC TCT GCA CAT AAC TGT G
	rppHb	ATT GGC GTG ACT TAA CCT C
<i>rssA</i>	rssAa	CCA TTC AGT ATG CGC AAA TCC GGC C
	rssAb	GCA GTT TAA GCC CGT TCA TTC GTG GC
<i>sucA</i>	sucAa	GTC GCA TCA GGC AAC CAG TG
	sucAb	GGT GGC ATC GGC TAC GGA TT
<i>yjjU</i>	yjjUa	GGC GCA GCT AAA ATT GTC TTT G
	yjjUb	CCA GGA TGC AAG CCC AAT GCG GC
<i>ypdA</i>	ypdAa	ACG ATG GTG GCT GTT TTG AA
	ypdAb	TTT GCA GCA TCC CGG TAA TA
<i>ytfR</i>	ytfRa	GAA GCC TGA TGC GAC GCT TA
	ytfRb	GAC CAG CAG CAC CAG CAA TA

B. Oligonucleótidos empleados para las inactivaciones realizadas por el método Datsenko-Wanner.

<i>arcA</i>	arcA-cm1 (Fw)	GCA GCA TGC ATC AGG CAG GTC AGG GAC TTT TGT ACT TCC TGT TTC GAT TTG TGT AGG CTG GAG CTG CTT C
	arcA-cm2 (Rv)	GAC GGT GGT AAA GCC GAT TAA TCT TCC AGA TCA CCG CAG AAG CGA TAA CCC ATA TGA ATA TCC TCC TTA G
<i>aas-lpIT</i>	aas-lpIT (Fwd)	TAC CTG TGC GTC CAG TGT TTG CCG TTT ACA CCT CAA TAA GCC AAC AGT GGG TGT AGG CTG GAG CTG CTT C
	aas-lpIT (Rv)	GGA TGC CTA TCA TTA CCG CCA GCG AGT AAA TGC CCA GCA TCA ACA ACA TGC ATA TGA ATA TCC TCC TTA G
<i>dhaM</i>	dhaM (Fwd)	CCT GGT CAT AGT TTC ACA TAG CAG CCG ACT GGG AGA AGG TGT CGG TGA ATG TGT AGG CTG GAG CTG CTT C
	dhaM (Rv)	TGC GTT TTA ACG TCC AGC GTT AGC GTT TCT TCT GGT TGT ATC GCA TAC AGC ATA TGA ATA TCC TCC TTA G
<i>fadD</i>	fadD (Fwd)	CGT TAT CAA TCT CTG GTA GAT ATG TTT GAG CAG TCG GTC GCG CGC TAC GCG TGT AGG CTG GAG CTG CTT C
	fadD (Rv)	TAC GCC AGG ATG CTG CAT GAC GAC ATC TTC AAT CTC GTT GGG ATA GAC GTC ATA TGA ATA TCC TCC TTA G
<i>glpQ</i>	glpQ (Fwd)	TAG TCA TCG CCC ATC GCG GTG CCA GTG GAT ATT TGC CGG AGC ATA CGC TGG TGT AGG CTG GAG CTG CTT C
	glpQ (Rv)	AAG GAT GCA CTA CCA GTT TAT TCT GCT GAG CAT CTT GCA CCA TGC CAG TGC ATA TGA ATA TCC TCC TTA G
<i>glpT</i>	glpT (Fwd)	GTC GAG CAT TGC GGT GAT GTT TGT ACT GTT GTT CCT CTG CGG TTG GTT CCG TGT AGG CTG GAG CTG CTT C
	glpT (Rv)	GGT GCC AGT TCC AGC GCA TGC AGA CCG ATC AGC ATC ACA GGA CCG TAG ATC ATA TGA ATA TCC TCC TTA G
<i>barA</i>	barA-cm1 (Fw)	GAG TCG AGG TCG TTG ACT GAA ACG GTT ATG ATC TGA CGC GTC GTC GTG CGG TGT AGG CTG GAG CTG CTT C
	barA-cm2 (Rv)	CCG GAT AAG GCG TTA ACG CCG CAT CCG GCA TAA ACA CAG GCA CTT TGT CAC ATA TGA ATA TCC TCC TTA G
<i>rna</i>	rna-cm1 (Fw)	ATG ATT TAA TAG ATA GTA TTC AGT GTA TGT ATG AGT TCC ACA CCC ATT

		ATG TGT AGG CTG GAG CTG CTT C
	rna-cm2 (Rv)	ATT TGC GTC ACA TTC AGG GAA TTC CTC AAT GCA GCG CAT TAT GTA TAA ATC ATA TGA ATA TCC TCC TTA G
<i>rppH</i>	rppH-cm1 (Fw)	TAT CCA CCC CTT CCT CTG TTT ATA ACT CTG CAC ATA ACT GTG AGT TAT TCG TGT AGG CTG GAG CTG CTT C
	rppH-cm2 (Rv)	AAC CAG TTA CGC GTT GAA GCA AGG ATT CGA ACG TCT TTG CGG CTT AAT CCC ATA TGA ATA TCC TCC TTA G
<i>rssA</i>	rssA-cm1 (Fw)	TGC GCT AAA AAA AGT GGG TAT TGA AAT TGA TAT CGT TGC AGG ATG TTG TGT AGG CTG GAG CTG CTT C
	rssA-cm2 (Rv)	AGC GTG CGC GCG ATG GAA ATC AAG CGT AGA TAT TTG CGG GCA AAC AGG TTC ATA TGA ATA TCC TCC TTA G
<i>yjjU</i>	yjjU-cm1 (Fw)	GAA AAT TTT ATC CAA AGC CAG TCC AGC GGA CTG GCT TTT GCG GTT TTA GCG TGT AGG CTG GAG CTG CTT C
	yjjU-cm2 (Rv)	ACA TTT GTG ACG GCA CAG TGC GAA TGA CGA CCA ACG TTT TAG CGC CCT GCC ATA TGA ATA TCC TCC TTA G
<i>ypdA</i>	ypdA-cm1 (Fw)	TCG ATC GGG CAG CGT TAA TGC TTA TCT GCC TGT TCT TTC TCA TCC GTA TCG TGT AGG CTG GAG CTG CTT C
	ypdA-cm2 (Rv)	CCG CTC AAT CAC CTT TGG ATC GAT GCC GTG CCC GGT ATC TCG CAC CGC AAC ATA TGA ATA TCC TCC TTA G
C. Oligonucleótidos empleados para los análisis por RT-qPCR*		
<i>barA</i>	barAa	CGGTGTGCCACGTATGAAGA
	barAb	CCAACAGTTCAGCAGCTCC
<i>gpp</i>	gpp-a	CAATAAGCCGCGATTTGCTC
	gpp-b	ACGCACCAGATAAGCAGCG
<i>lplT</i>	lplTa	TGCGGCAGCGAAGTTAGTTA
	lplTb	GCTCGTGTGCAGGGAAAAA
<i>mutH</i>	mutHa	AGCGCAAGCACAGCAACTTT
	mutHb	CAATCCAGCCTTTATCGCGT
<i>ndk</i>	ndk-a	AACGCGGTAGCAAAAAACGT
	ndk-b	TGTTCAACGGTCAGGTGCAG
<i>ppa</i>	ppa-a	ACGTCCTGCGGGTAAAGAT
	ppa-b	GCCGCTCTTTGTCGATTTT
<i>ptsP</i>	ptsPa	GACTGCCTTGTTGGGCAAT
	ptsPb	AAGGTAACGTGGCATCCTGC
<i>relA</i>	relAa	CGTAAACTGATTGCGTGGCA
	relAb	GGCGTAAAGACGTACACCCG
<i>rna</i>	rna-a	TTCTCTTCTGCCAACGCCTT
	rna-b	TGACTCTGGCAAATCCGGT
<i>rpoD</i>	rpoDa	GATTCTGCGACCACCGAAAG
	rpoDb	TCGATACCGAAACGCATACG
<i>rppH</i>	rppHa	GGCAATTTCCGCAAGGC
	rppHb	TCGAACGTCTTTGCGGCT
<i>rssA</i>	rssAa	AAATTGTTCCCGTCGCTTTG
	rssAb	ATGATGCGCGAATAGCAAGA
<i>spoT</i>	spoTa	ATCCGATCTCTTCCCGGATG

	spoTb	GCACTGCATAAGCGAAGTCG
<i>tas</i>	tas-a	ATAGCTGGACGGATTCTGCG
	tas-b	CGACACGCCGATATAACGAA
<i>ubiE</i>	ubiEa	GGCAGAATCCATCCGTATGC
	ubiEb	CCCCTGCCGTCAGATTGTAG
<i>ygdQ</i>	ygdQa	CATCAATCGCCTGGGTTACG
	ygdQb	GAACAAGCCACCCAGAAGCA
<i>yjjU</i>	yjjUa	GCCTGGACGAATAGCTCTGG
	yjjUb	CGAAAGGATTAAGCTGCGCG
<i>ypdA</i>	ypdAa	AAGGCGTTGTCACCATCAGC
	ypdAb	GACCCGCTCAATCACCTTG
*La secuencia de los oligonucleótidos restantes utilizados en este estudio ya han sido reportados [8, 108,117].		

B. Figuras:

Figura S1: Secuencia nucleotídica del rearreglo cromosomal después del proceso de evolución adaptativa en las cepas PB12 y PB13.

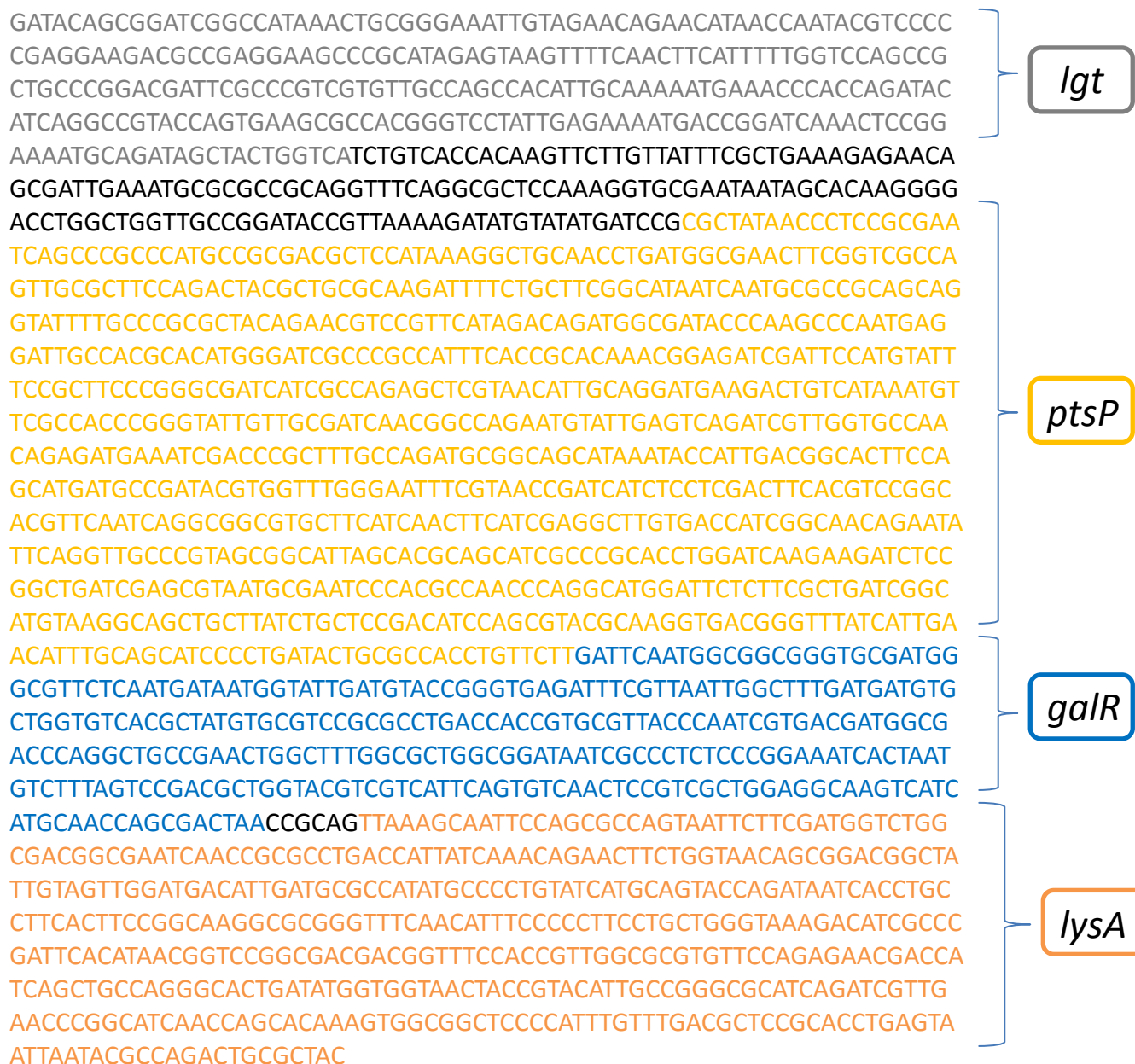
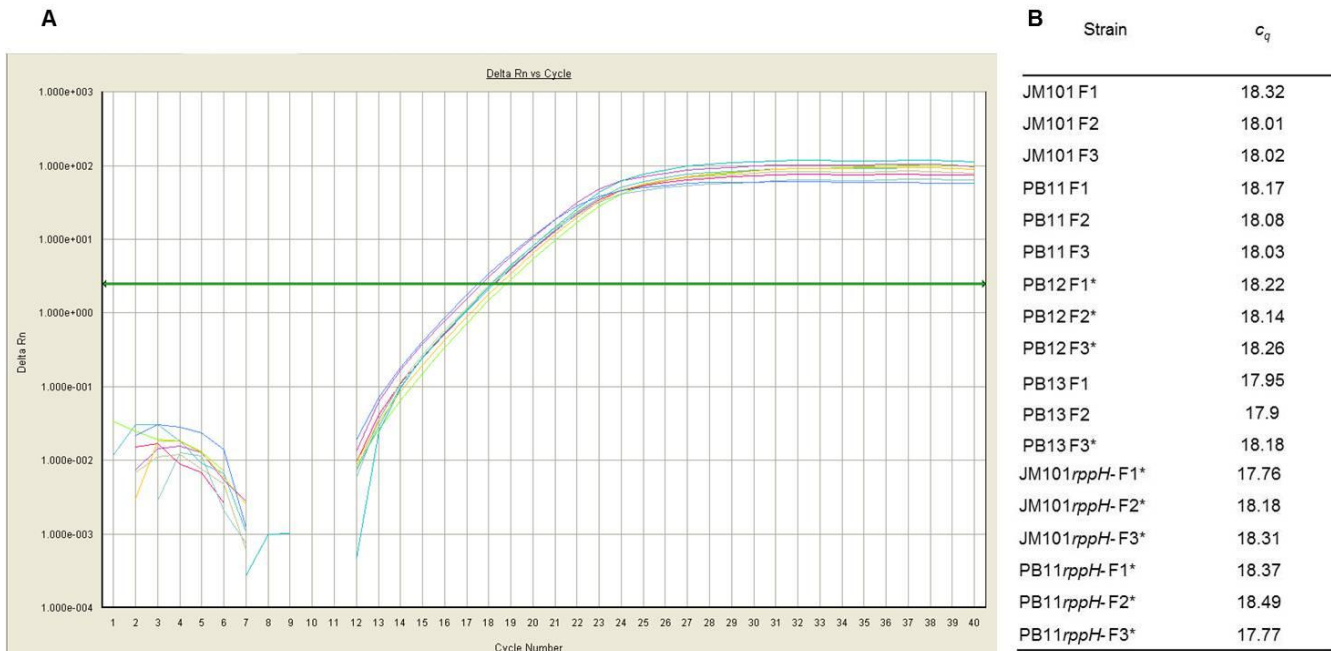


Figura S2: Curvas de amplificación para el gen *ihfB* en las diferentes cepas de este estudio. Esta figura muestra las posiciones de las curvas de amplificación del gen *ihfB* (Sección A) y los valores de Ct de éste gen (Sección B) en diferentes cepas empleadas en este estudio. Como es posible observar, todas las curvas de amplificación del gen *ihfB* son muy similares, y han sido usadas como referencia para la determinación de los niveles de RT-qPCR. Los valores presentados en la tabla se desprenden de tres diferentes fermentaciones (F1, F2 y F3) para todas las cepas del estudio. Ya que los valores incluidos en la sección B. son muy similares, solo el tercer valor es presentado en la gráfica. Estos resultados demuestran que los valores de expresión obtenidos para el gen *ihfB* son reproducibles en todas las cepas, lo cual representa la característica principal para ser utilizado como gen de referencia de acuerdo a lo reportado en el manual MIQE [201], y corroboran la estabilidad de la expresión de este gen de referencia bajo estas condiciones.



XI. Referencias:

1. Bozell JJ, Petersen GR. (2010). Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy s –Top 10” revisited. Green Chem. 12:539–554.
2. Bruschi F, Dunder M, Gahan P, Gartland K, Szente M, Viola-Magni M, Akbarova Y. (2011). Biotechnology worldwide and the European Biotechnology Thematic Network’ Association (EBTNA). Curr. Opin. Biotechnol. S22: S7–S14.
3. Chen X, Zhou L, Tian K, Kumar A, Singh S, Prior BA, Wang Z. (2013). Metabolic engineering of *Escherichia coli*: A sustainable industrial platform for bio-based chemical production. Biotechnol. Adv. 31(8):1200-23.
4. Sprenger GA. (2007). From scratch to value: engineering *Escherichia coli* wild type cells to the production of L-phenylalanine and other fine chemicals derived from chorismate. Appl. Microbiol. Biotechnol. 75(4):739-49.
5. Krämer M, Bongaerts J, Bovenberg R, Kremer S, Müller U, Orf S, Wubbolts M, Raeven L. (2003). Metabolic engineering for microbial production of shikimic acid. Metab. Eng. 5(4):277-83.
6. Escalante A, Salinas Cervantes A, Gosset G, Bolívar F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: peculiarities of regulation and impact on growth and product formation. Appl. Microbiol. Biotechnol. 94(6):1483-94.
7. Flores N, Xiao J, Berry A, Bolivar F, Valle F. (1996). Pathway engineering for the production of aromatic compounds in *Escherichia coli*. Nat. Biotechnol. 14: 620–23.
8. Flores N, Flores S, Escalante A, de Anda R, Leal L, Malpica R, Georgellis D., Gosset G, Bolivar F. (2005). Adaptation for fast growth on glucose by differential expression of central carbon metabolism and gal regulon genes in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. Metabol. Eng. 7: 70-87.
9. Gosset G, Yong-Xiao J, Berry A. (1996). A direct comparison of approaches for increasing carbon flow to aromatic biosynthesis in *Escherichia coli*. J. Ind. Microbiol. 17(1): 47-52.
10. Báez J, Bolívar F, Gosset G. (2001). Determination of 3-deoxy-D-arabino-heptulosonate-7-phosphate productivity and yield from glucosa in *Escherichia coli* devoid of the glucose phosphotransferase system. Biotechnol. Bioeng. 6: 530-35.
11. Rodriguez A, Martínez JA, Báez-Viveros JL, Flores N, Hernández-Chávez G, Ramírez OT, Gosset G, Bolivar F. (2013). Constitutive expression of selected genes from the pentose phosphate and aromatic pathways increases the shikimic acid yield in high-glucose batch cultures of an *Escherichia coli* strain lacking PTS and pykF. Microb. Cell Fact. 12(1):86.
12. Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK. (2006). Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. Nat. Genet. 38:1406–1412.
13. Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y, Palsson B. (2010). Genetic basis of growth adaptation of *Escherichia coli* after deletion of *pgi*, a major metabolic gene. PLoS Genet. 6:1001186.
14. Raghunathan A. & Palsson B. (2003). Scalable method to determine mutations that occur during adaptive evolution of *Escherichia coli*. Biotechnol. Lett. 25:435–441.
15. Aguilar C, Escalante A, Flores N, de Anda R, Riveros-McKay F, Gosset G, Morett E, Bolívar F. (2012). Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system. BMC Genomics. 10(13):385.

16. Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. (2000). The evolution of mutation rates: separating causes from consequences. *Bioessays*. 22(12):1057-66.
17. Tenaillon O, Denamur E, Matic I. (2004). Evolutionary significance of stress-induced mutagenesis in bacteria. *Trends Microbiol.* 12(6): 264-70.
18. Lee DH & Palsson BØ. (2010). Adaptive evolution of Escherichia coli K-12 MG1655 during growth on a Nonnative carbon source, L-1,2-propanediol. *Appl. Environ. Microbiol.* 76 (13):4158-68.
19. Minty JJ, Lesnefsky AA, Lin F, Chen Y, Zaroff TA, Veloso AB, Xie B, McConnell CA, Ward RJ, Schwartz DR, Rouillard JM, Gao Y, Gulari E, Lin XN. (2011). Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in Escherichia coli. *Microb. Cell Fact.* 25;10:18.
20. Lee SY. & Papoutsakis ET. (Ed.). (1999). Metabolic Engineering Marcel Dekker, New York, 423 pp.
21. Dressaire C, Picard F, Redon E, Loubière P, Queinnec I, Girbal L, Coccagn-Bousquet M. (2013). Role of mRNA stability during bacterial adaptation. *PLoS One.* 8(3):e59059.
22. Lackner DH & Bähler J. (2008). Translational control of gene expression from transcripts to transcriptomes. *Int. Rev. Cell. Mol. Biol.* 271:199-251.
23. Belasco JG. (2010). All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat. Rev. Mol. Cell Biol.* 11(7):467-78.
24. Lin PH, Singh D, Bernstein JA, Lin-Chao S. (2008). Genomic analysis of mRNA decay in E. coli with DNA microarrays. *Methods Enzymol.* 447:47-64.
25. McDowall KJ, Lin-Chao S, Cohen SN. (1994). A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.* 269(14):10790-6.
26. Mackie, G. A. (1998). Ribonuclease E is a 5'-end-dependent endonuclease. *Nature.* 442 395:720-723.
27. Bouvier M, Carpousis AJ. (2011). A tale of two mRNA degradation pathways mediated by RNase E. *Mol. Microbiol.* 82(6):1305-10.
28. Deana A, Celesnik H, Belasco JG. (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature.* 451(7176):355-8.
29. Papavera, N. (2001). Historia de la biología comparada; desde el génesis hasta el siglo de las luces. UNAM. ISBN: 9683690084.
30. Mayr, E. (2001). What Evolution Is? Basic Books. Perseus Books Group. New York. ISBN 0-465-04425-5.
31. Darwin, C. (2001). El origen de las especies. Editores Mexicanos Unidos. S.A. (2da. Ed.) México. ISBN 968-15-1044-5.
32. Mase J. (2011). Genetic drift. *Curr. Biol.* 21 pp. R837–R838.
33. Kimura M. (1968). Evolutionary rate at the molecular level. *Nature.* 217(5129):624-6.
34. King JL & Jukes TH. (1969). Non-Darwinian evolution. *Science.* 164(3881):788-98.
35. Cairns J, Overbaugh J, Miller S. (1988). The origin of mutants. *Nature.* 335(6186):142-5.
36. Foster PL. (2007). Stress-induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.* 42(5):373-97.
37. Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T. (2006). DNA Repair and Mutagenesis. Washington, D.C: ASM Press.
38. Kelley WL. (2006). Lex marks the spot: the virulent side of SOS and a closer look at the LexA regulon. *Molec. Microbiol.* 62:1228-1238.

39. Fernandez de Henestrosa AR, Ogi T, Aoyagi S, Chafin D, Hayes JJ, Ohmori H, Woodgate R. (2000). Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Molec. Microbiol.* 35:1560–1572.
40. Wade JT, Reppas NB, Church GM, Struhl K. (2005). Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites. *Genes Dev.* 19:2619–2630.
41. Fuchs RP, Fujii S, Wagner J. (2004). Properties and functions of *Escherichia coli*: Pol IV and Pol V. *Adv. Protein. Chem.* 69:229–264.
42. Lehmann AR. (2006). New functions for Y family polymerases. *Mol. Cell.* 24:493–495.
43. Nohmi T. (2006). Environmental stress and lesion-bypass DNA polymerases. *Annu. Rev. Microbiol.* 60:231–253.
44. Tippin B, Pham P, Goodman MF. (2004). Error-prone replication for better or worse. *Trends. Microbiol.* 12:288–295.
45. Kuban W, Banach-Orlowska M, Schaaper RM, Jonczyk P, Fijalkowska IJ. (2006). Role of DNA polymerase IV in *Escherichia coli* SOS mutator activity. *J. Bacteriol.* 188:7977–7980.
46. Kuban W, Jonczyk P, Gawel D, Malanowska K, Schaaper RM, Fijalkowska IJ. (2004). Role of *Escherichia coli* DNA polymerase IV in in vivo replication fidelity. *J. Bacteriol.* 186:4802–4807.
47. McKenzie GJ, Magner DB, Lee PL, Rosenberg SM. (2003). The *dinB* operon and spontaneous mutation in *Escherichia coli*. *J. Bacteriol.* 185:3972–3977.
48. Strauss BS, Roberts R, Francis L, Pouryazdanparast P. (2000). Role of the *dinB* gene product in spontaneous mutation in *Escherichia coli* with an impaired replicative polymerase. *J. Bacteriol.* 182:6742–6750.
49. Wolff E, Kim M, Hu K, Yang H, Miller JH. (2004). Polymerases leave fingerprints: analysis of the mutational spectrum in *Escherichia coli rpoB* to assess the role of polymerase IV in spontaneous mutation. *J. Bacteriol.*;186:2900–2905.
50. Kim SR, Maenhaut-Michel G, Yamada M, Yamamoto Y, Matsui K, Sofuni T, Nohmi T, Ohmori H. (1997). Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: an overexpression of *dinB/dinP* results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc. Natl. Acad. Sci. USA.* 94:13792–13797.
51. Wagner J, Nohmi T. (2000). *Escherichia coli* DNA polymerase IV mutator activity: genetic requirements and mutational specificity. *J. Bacteriol.* 182:4587–4595.
52. Hengge-Aronis R. (2002). Signal transduction and regulatory mechanisms involved in control of the σ S (RpoS) subunit of RNA polymerase. *Microbiol. Mol. Biol. Rev.* 66:373–395.
53. Patten CL, Kirchhof MG, Schertzberg MR, Morton RA, Schellhorn HE. (2004). Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. *Mol. Genet. Genomics.* 272:580–591.
54. Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R. (2005). Genome-wide analysis of the general stress response network in *Escherichia coli*: σ S-dependent genes, promoters, and sigma factor selectivity. *J. Bacteriol.* 187:1591–1603.
55. Layton JC, Foster PL. (2003). Error-prone DNA polymerase IV is controlled by the stress-response sigma factor, RpoS, in *Escherichia coli*. *Mol. Microbiol.* 50:549–561.
56. Kunkel TA, Erie DA. (2005). DNA mismatch repair. *Annu. Rev. Biochem.* 74:681–710.
57. Foster PL. (1999). Are adaptive mutations due to a decline in mismatch repair? The evidence is lacking. *Mutat. Res.* 436:179–184.
58. Reddy M. & Gowrishankar J. (1997). A genetic strategy to demonstrate the occurrence of spontaneous mutations in non-dividing cells within colonies of *Escherichia coli*. *Genetics.* 147:991–1001.

59. Feng G, Tsui HCT, Winkler ME. (1996). Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. J Bacteriol. 178:2388–2396.
60. Tsui HCT, Feng G, Winkler ME. (1997). Negative regulation of *mutS* and *mutH* repair gene expression by the Hfq and RpoS global regulators of *Escherichia coli* K-12. J. Bacteriol. 179:7476–7487.
61. Bjedov I, Tenaillon O, Gerard B, Souza V, Denamur E, Radman M, Taddei F, Matic I. (2003). Stress-induced mutagenesis in bacteria. Science. 300:1404–1409.
62. Foster PL, Gudmundsson G, Trimarchi JM, Cai H, Goodman MF. (1995). Proofreading-defective DNA polymerase II increases adaptive mutation in *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 92:7951–7955.
63. Harris RS, Feng G, Ross KJ, Sidhu R, Thulin C, Longerich S, Szigety SK, Hastings PJ, Winkler ME, Rosenberg SM. (1999). Mismatch repair is diminished during stationary-phase mutation. Mutat. Res. 437:51–60.
64. Zhao J. & Winkler ME. (2000). Reduction of GC to TA transversion mutation by overexpression of MutS in *Escherichia coli* K-12. J. Bacteriol. 182:5025–5028.
65. Braeken K, Moris M, Daniels R, Vanderleyden J, Michiels J. (2006). New horizons for (p)ppGpp in bacterial and plant physiology. Trends Microbiol.14:45–54.
66. Magnusson LU, Farewell A, Nyström T. (2005). ppGpp: a global regulator in *Escherichia coli*. Trends Microbiol. 13:236–242.
67. Gourse RL, Ross W, Rutherford ST. (2006). General pathway for turning on promoters transcribed by RNA polymerases containing alternative σ Factors. J. Bacteriol. 188:4589–4591.
68. Kvint K, Farewell A, Nystrom T. (2000). RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of σ_s . J. Biol. Chem. 275:14795–14798.
69. Chang DE, Smalley DJ, Conway T. (2002). Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. Molec. Microbiol. 45:289–306.
70. Traxler MF, Chang DE, Conway T. (2006). Guanosine 3', 5'-bispyrophosphate coordinates global gene expression during glucose-lactose diauxie in *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 103:2374–2379.
71. Mellon I. (2005). Transcription-coupled repair: a complex affair. Mutat. Res. 577:155–161.
72. Beletskii A, Bhagwat AS. (1996). Transcription-induced mutations: increases in C to T mutations in the nontranscribed strand during transcription in *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 93:13919–13924.
73. Klapacz J, Bhagwat AS. (2002). Transcription-dependent increase in multiple classes of base substitution mutations in *Escherichia coli*. J. Bacteriol. 184:6866–6872.
74. Klapacz J. & Bhagwat AS. (2005). Transcription promotes guanine to thymine mutations in the non-transcribed strand of an *Escherichia coli* gene. DNA Repair. 4:806–813.
75. Wright BE, Longacre A, Reimers JM. (1999). Hypermutation in derepressed operons of *E. coli* K12. Proc. Natl. Acad. Sci. USA. 96:5089–5094.
76. Rudner R, Murray A, Huda N. (1999). Is there a link between mutation rates and the stringent response in *Bacillus subtilis*? Ann. NY. Acad. Sci. 870:418–422.
77. Datta A. & Jinks-Robertson S. (1995). Association of increased spontaneous mutation rates with high levels of transcription in yeast. Science. 268:1616–1619.
78. Davis BD. (1989). Transcriptional bias: A non-Lamarckian mechanism for substrate-induced mutations. Proc. Natl. Acad. Sci. USA. 86:5005–5009.

79. Fitch WM. (1982). The challenges to Darwinism since the last centennial and the impact of molecular studies. *Evolution*. 36(6):1133–1143.
80. Wright BE. (2004). Stress-directed adaptive mutations and evolution. *Molec. Microbiol.* 52:643–650.
81. Clomburg JM. & Gonzalez R. (2010). Biofuel production in *Escherichia coli*: the role of metabolic engineering and synthetic biology. *Appl. Microbiol. Biotechnol.* 86:419–434.
82. Kalia VC, & Purohit HJ. (2008). Microbial diversity and genomics in aid of bioenergy. *J. Ind. Microbiol. Biotechnol.* 35:403-419.
83. Ren Q, Ruth K, Thöny-Meyer L, Zinn M. (2010). Enantiomerically pure hydroxycarboxylic acids: current approaches and future perspectives. *Appl. Microbiol. Biotechnol.* 87:41–52.
84. Lu T, Khalil A, Collins J. (2009). Next-generation synthetic gene networks. *Nat Biotechnol*, 27:1139–1150.
85. Purnick P. & Weiss R. (2009). The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10:410–422.
86. Kim T, Sohn S, Kim H, Lee S. (2008). Strategies for systems-level metabolic engineering. *Biotechnol J*, 3:612–623.
87. Mukhopadhyay A, Redding A, Rutherford B, Keasling J (2008). Importance of systems biology in engineering microbes for biofuel production. *Curr. Opin. Biotechnol.* 19:228–234.
88. Conrad TM, Lewis NE, Palsson B. (2011) Microbial laboratory evolution in the era of genome-scale science. *Mol. Syst. Biol.* 7:509
89. Portnoy VA, Bezdán D, Zengler K (2011) Adaptive laboratory evolution– harnessing the power of biology for metabolic engineering. *Curr. Opin. Biotechnol.* 22:590–594.
90. Nam H, Conrad TM, Lewis NE. (2011). The role of cellular objectives and selective pressures in metabolic pathway evolution. *Curr. Opin. Biotechnol.* 22:595–600.
91. Barrick JE. & Lenski RE. (2013). Genome dynamics during experimental evolution. *Nat. Rev. Genet.* 14(12):827-39.
92. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 1977. 74(12):5463-7.
93. Mardis ER. (2013). Next-generation sequencing platforms. *Annu. Rev. Anal. Chem.* 6:287-303.
94. Gosset G. (2005). Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate: sugar phosphotransferase system. *Microb. Cell Fact.* 4:14.
95. Gosset G. (2009). Production of aromatic compounds in bacteria. *Curr. Opin. Biotechnol.* 20:651–658.
96. Ferenci T. (2001). Hungry bacteria: definition and properties of a nutritional state. *Environ. Microbiol.* 3:605–611
97. Postma PW, Lengeler JW, Jacobson GR (1996). Phosphoenolpyruvate: Carbohydrate phosphotransferase systems. In: Neidhart FC (ed) *Escherichia coli* and *Salmonella*. Cellular and molecular biology. ASM Press, Washington, D. C., pp 1149–1174.
98. Saier MH Jr, Ramseier TM, Reizer J (1996). Regulation of carbon utilization. In: Neidhart FC (ed) *Escherichia coli* and *Salmonella*. Cellular and molecular biology. ASM Press, Washington, D. C., pp 1325–1343.
99. Novotny MJ, Frederickson WL, Waygood EB, Saier MH Jr. (1985). Allosteric regulation of glycerol kinase by enzyme III_{glc} of the phosphotransferase system in *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol.* 162(2):810-6.
100. Plumbridge J. (2002). Regulation of gene expression in the PTS in *Escherichia coli*: the role and interactions of Mlc. *Curr. Opin. Microbiol.* 5(2):187-93.

101. Körner H, Sofia HJ, Zumft WG. (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. FEMS Microbiol. Rev. 27(5):559-92.
102. Seok YJ, Koo BM, Sondej M, Peterkofsky A. (2001). Regulation of *E. coli* glycogen phosphorylase activity by HPr. J. Mol. Microbiol. Biotechnol. 3(3):385-93.
103. Görke B. & Rak B. (1999). Catabolite control of *Escherichia coli* regulatory protein BglG activity by antagonistically acting phosphorylations. EMBO J. 18(12):3370-9.
104. Park YH, Lee CR, Choe M, Seok YJ. (2013). HPr antagonizes the anti- σ^{70} activity of Rsd in *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 110(52):21142-7.
105. Lux R, Jahreis K, Bettenbrock K, Parkinson JS, Lengeler JW. (1995). Coupling the phosphotransferase system and the methyl-accepting chemotaxis protein-dependent chemotaxis signaling pathways of *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 92(25):11583-7.
106. Liao JC, Hou SY, Chao YP. (1996). Pathway analysis, engineering, and physiological considerations for redirecting central metabolism. Biotechnol. Bioeng. 52(1):129-40.
107. Flores N. (1995). Construcción y caracterización de cepas de *Escherichia coli* mutantes en el sistema de transporte de carbohidratos PTS. Instituto de Biotecnología, UNAM. Tesis de Maestría.
108. Flores N, Leal L, Sigala JC, de Anda R, Escalante A, Martínez A, Ramirez OT, Gosset G, Bolivar F. (2007). Growth Recovery on Glucose under Aerobic Conditions of an *Escherichia coli* Strain Carrying a Phosphoenolpyruvate: Carbohydrate Phosphotransferase System Deletion by Inactivating *arcA* and Overexpressing the Genes Coding for Glucokinase and Galactose Permease J. Mol. Microbiol. Biotechnol. 13: 105-16.
109. Flores N, Leal L, Sigala JC, de Anda R, Escalante A, Martínez A, Ramirez OT, Gosset G, Bolivar F. (2007). Growth Recovery on Glucose under Aerobic Conditions of an *Escherichia coli* Strain Carrying a Phosphoenolpyruvate: Carbohydrate Phosphotransferase System Deletion by Inactivating *arcA* and Overexpressing the Genes Coding for Glucokinase and Galactose Permease J. Mol. Microbiol. Biotechnol. 13: 105-16.
110. Brückner R, Titgemeyer F. (2002). Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. FEMS Microbiol. Lett. 209(2):141-8.
111. Flores S, Flores N, de Anda R, González A, Escalante A, Sigala JC, Gosset G, Bolívar F. (2005). Nutrient-scavenging stress response in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system, as explored by gene expression profile analysis. J. Mol. Microbiol. Biotechnol. 10(1):51-63.
112. Sigala JC, Flores S, Flores N, Aguilar C, de Anda R, Gosset G, Bolívar F. (2008). Acetate Metabolism in *Escherichia coli* Strains Lacking Phosphoenolpyruvate: Carbohydrate Phosphotransferase System; Evidence of Carbon Recycling Strategies and Futile Cycles. J. Mol. Microbiol. Biotechnol. 16(3-4): 224-35.
113. Csonka LN. & Fraenkel DG. (1977). Pathways of NADPH formation in *Escherichia coli*. J. Biol. Chem. 252(10):3382-91.
114. Szyperski, T. (1995). Biosynthetically directed fractional ^{13}C -labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. Eur. J. Biochem. 232:433-448.
115. Singaravelan B, Roshini BR, Munavar MH. (2010). Evidence that the *supE44* mutation of *Escherichia coli* is an amber suppressor allele of *glnX* and that it also suppresses ochre and opal nonsense mutations. J. Bacteriol. 192(22):6039-44.

116. Olvera L, Mendoza-Vargas A, Flores N, Olvera M, Sigala JC, Gosset G, Morett E, Bolívar F. (2009). Transcription analysis of central metabolism genes in *Escherichia coli*. Possible roles of sigma38 in their expression, as a response to carbon limitation. PLoS One. 4(10).
117. Flores N, Escalante A, de Anda R, Báez-Viveros JL, Merino E, Franco B, Georgellis D, Gosset G, Bolívar F. (2008). New insights on the role of the sigma factor RpoS as revealed in *Escherichia coli* strains lacking the Phosphoenolpyruvate: carbohydrate phosphotransferase system. J. Mol. Microbiol. Biotechnol. 14:176–92.
118. Báez-Viveros JL, Flores N, Juárez K, Castillo-España P, Bolívar F, Gosset G. (2007). Metabolic transcription analysis of engineered *Escherichia coli* strains that overproduce L-phenylalanine. Microb. Cell Fact. 6:30.
119. Messing J. (1979). A multipurpose cloning system based on the single stranded DNA bacteriophage M13. Recombinant DNA technical bulletin. 2: 43-48.
120. Ausubel FA, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. (eds.) (2006). Short protocols in molecular biology a compendium of methods from Current protocols in molecular biology. John Wiley & Sons, New York. 740 p.
121. Bachmann BJ. (1996). Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12 In: Neidhart FC, (Ed.). *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, 2nd ed. ASM, Washington DC, USA Vol. II pp. 2460–2488.
122. Bolívar F, Rodríguez R, Greene P, Betlach M, Heyneker H, Boyer H. (1977). Construction and characterization of new cloning vehicles II. A multipurpose cloning system. Gene. 2(2): 95-113.
123. Naas T, Blot M, Fitch WM, Arber W. (1995). Dynamics of IS-related genetic rearrangements in resting *Escherichia coli* K-12. Mol. Biol. Evol. 12(2):198-207.
124. Benito A, Ventoura G, Casadei M, Robinson T, Mackey B. (1999). Variation in resistance of natural isolates of *Escherichia coli* O157 to high hydrostatic pressure, mild heat, and other stresses. Appl. Environ. Microbiol. 65(4):1564-1569.
125. Johnson J, Delavari P, Stell A, Prats G, Carlino U, Russo T. (2001). Integrity of archival strain collections: the ECOR collection. ASM NEWS. 67(6):288-289.
126. Atlung T, Nielsen HV, Hansen FG. (2002). Characterisation of the allelic variation in the *rpoS* gene in thirteen K12 and six other non-pathogenic *Escherichia coli* strains. Mol. Genet. Genomics 2002, 266(5):873-81
127. Faure D, Frederick R, Wloch D, Portier P, Blot M, Adams J. (2004). Genomic changes arising in long-term stab cultures of *Escherichia coli*. J. Bacteriol. 186(19):6437-6442.
128. Ferenci T, Galbiati HF, Betteridge T, Phan K, Spira B. (2011). The constancy of global regulation across a species: the concentrations of ppGpp and RpoS are strain-specific in *Escherichia coli*. BMC Microbiol. 11:62.
129. Mahillon J. & Chandler M. (1998). Insertion sequences. Microbiol. Mol. Biol. Rev. 62(3):725-74.
130. Jordan E, Saedler H, Starlinger P. (1968). O0 and strong-polar mutations in the *gal* operon are insertions. Mol. Gen. Genet. 102(4):353-63.
131. Saedler H, Reif HJ, Hu S, Davidson N. (1974). IS2, a genetic element for turn-off and turn-on of gene activity in *E. coli*. Mol. Gen. Genet.; 132(4):265-89.
132. Schneider D, Duperchy E, Coursange E, Lenski RE, Blot M. (2000). Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. Genetics. 156: 477-488.

133. Hart DL, Medhora M, Green L, Dykhuizen DE. (1986). The evolution of DNA sequences in *Escherichia coli*. Philos Trans. R. Soc. Lond. B. Biol. Sci. 312(1154):191-204.
134. Blot M. (1994). Transposable elements and adaptation of host bacteria. Genetica. 93(1-3): 5-12.
135. Luciano DJ, Hui MP, Deana A, Foley PL, Belasco KJ, Belasco JG. (2012). Differential control of the rate of 5'-end-dependent mRNA degradation in *Escherichia coli*. J Bacteriol. 194(22):6233-9.
136. Harvat EM, Zhang YM, Tran CV, Zhang Z, Frank MW, Rock CO, Saier MH Jr. (2005). Lysophospholipid flipping across the *Escherichia coli* inner membrane catalyzed by a transporter (LpIT) belonging to the major facilitator superfamily. J Biol Chem. 280(12):12028-34.
137. Morein S, Andersson A, Rilfors L, Lindblom G. (1996). Wild-type *Escherichia coli* cells regulate the membrane lipid composition in a "window" between gel and non-lamellar structures. J. Biol. Chem. 271 (12):6801-9.
138. Bentley WE, Mirjalili N, Andersen DC, Davis RH, Kompala DS. (1990). Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. Biotechnol. Bioeng. 35(7):668-81.
139. Flores S, de Anda-Herrera R, Gosset G, Bolívar FG. (2004). Growth-rate recovery of *Escherichia coli* cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway. Biotechnol Bioeng. 87(4):485-94.
140. Ow DS, Nissom PM, Philp R, Oh SK, Yap MG. (2006). Global transcriptional analysis of metabolic burden due to plasmid maintenance in *Escherichia coli* DH5 α during batch fermentation. Enzyme Microb. Technol. 39:391–398.
141. Yu H, Shi Y, Sun X, Luo H, Shen Z. (2003). Effect of poly (β -hydroxybutyrate) accumulation on the stability of a recombinant plasmid in *Escherichia coli*. J. Biosci. Bioeng. 96(2):179-83.
142. Death A, Ferenci T. (1994). Between feast and famine: endogenous inducer synthesis in the adaptation of *Escherichia coli* to growth with limiting carbohydrates. J Bacteriol. 176 (16):5101-7.
143. Hsu L, Jackowski S, Rock CO. (1991). Isolation and characterization of *Escherichia coli* K-12 mutants lacking both 2-acyl-glycerophosphoethanolamine acyltransferase and acyl-acyl carrier protein synthetase activity. J. Biol. Chem. 266 (21):13783-8.
144. Weimar JD, DiRusso CC, Delio R, Black PN. (2002). Functional role of fatty acyl-coenzyme A synthetase in the transmembrane movement and activation of exogenous long-chain fatty acids. Amino acid residues within the ATP/AMP signature motif of *Escherichia coli* FadD are required for enzyme activity and fatty acid transport. J. Biol. Chem. 277(33):29369-76.
145. Lemieux MJ, Huang Y, Wang DN. (2004). Glycerol-3-phosphate transporter of *Escherichia coli*: structure, function and regulation. Res Microbiol. 155(8):623-9.
146. Lemieux MJ, Huang Y, Wang DN. (2004). The structural basis of substrate translocation by the *Escherichia coli* glycerol-3-phosphate transporter: a member of the major facilitator superfamily. Curr. Opin. Struct. Biol. 14(4):405-12.
147. Lemieux MJ, Huang Y, Wang da N. (2005). Crystal structure and mechanism of GlpT, the glycerol-3-phosphate transporter from *E. coli*. J Electron Microsc (Tokyo). 54 Suppl 1:i43-6.
148. Gutknecht R, Beutler R, Garcia-Alles LF, Baumann U, Erni B. (2001). The dihydroxyacetone kinase of *Escherichia coli* utilizes a phosphoprotein instead of ATP as phosphoryl donor. EMBO J. 20(10):2480-6.
149. Cooper VS. & Lenski RE. (2000). The population genetics of ecological specialization in evolving *Escherichia coli* populations. Nature. 407(6805):736-9.
150. Tchieu JH, Norris V, Edwards JS, Saier MH Jr. (2001). The complete phosphotransferase system in *Escherichia coli*. J. Mol. Microbiol. Biotechnol. 3(3):329-46.

151. Schofield MJ. & Hsieh P. (2003). DNA mismatch repair: Molecular Mechanisms and Biological Function. Annu. Rev. Microbiol. 57:579–608.
152. Junop MS, Yang W, Funchain P, Clendenin W, Miller JH. (2003). In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. DNA Repair. 2(4): 387-405.
153. Chen C, Blumentritt CA, Curtis MM, Sperandio V, Torres AG, Dudley EG. (2013). Restrictive streptomycin-resistant mutations decrease the formation of attaching and effacing lesions in *Escherichia coli* O157:H7 strains. Antimicrob. Agents Chemother. [Epub ahead of print]
154. Hall BM, Ma C, Liang P, Singh, KK. (2009). Fluctuation AnaLysis CalculatOR (FALCOR): a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. Bioinformatics. 25(12): 1564-1565.
155. Orren DK, Selby CP, Hearst JE, Sancar A. (1992). Post-incision steps of nucleotide excision repair in *Escherichia coli*. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I. J. Biol. Chem. 267(2):780-8.
156. Dzidic S, Radman M. (1989). Genetic requirements for hyper-recombination by very short patch mismatch repair: involvement of *Escherichia coli* DNA polymerase I. Mol Gen Genet. 217(2-3):254-6.
157. Mao EF, Lane L, Lee J, Miller JH. (1997). Proliferation of mutators in a cell population J Bacteriol. 179(2):417-22.
158. Desai MM, Fisher DS. (2011). The balance between mutators and nonmutators in asexual populations. Genetics. 188(4):997-1014.
159. Denamur E, Lecointre G, Darlu P, Tenaillon O, Acquaviva C, Sayada C, Sunjevaric I, Rothstein R, Elion J, Taddei F, Radman M, Matic I. (2000). Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. Cell. 103(5):711-21.
160. Wielgoss S, Barrick JE, Tenaillon O, Wisner MJ, Dittmar WJ, Cruveiller S, Chane-Woon-Ming B, Médigue C, Lenski RE, Schneider D. (2013). Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. Proc. Natl. Acad. Sci. USA. 110(1):222-7.
161. Denamur E, Lecointre G, Darlu P, Tenaillon O, Acquaviva C, Sayada C, Sunjevaric I, Rothstein R, Elion J, Taddei F, Radman M, Matic I. (2000). Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. Cell. 103(5):711-21.
162. Turrientes MC, Baquero F, Levin BR, Martínez JL, Ripoll A, González-Alba JM, Tobes R, Manrique M, Baquero MR, Rodríguez-Domínguez MJ, Cantón R, Galán JC. (2013). Normal mutation rate variants arise in a Mutator (Mut S) *Escherichia coli* population. PLoS One. 8(9):e72963.
163. Lin E.C. & Iuchi S. (1991). Regulation of gene expression in fermentative and respiratory systems in *Escherichia coli* and related bacteria. Annu. Rev. Genet. 25:361-87.
164. Iuchi S. & Lin EC. (1991). Adaptation of *Escherichia coli* to respiratory conditions: regulation of gene expression. Cell 66(1);5-7.
165. Gunsalus RP, Park SJ. (1994). Aerobic-anaerobic gene regulation in *Escherichia coli*: control by the ArcAB and Fnr regulons. Res. Microbiol. 145(5-6);437-50.
166. Malpica R, Franco B, Rodriguez C, Kwon O, Georgellis D. (2004). Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. Proc. Natl. Acad. Sci. USA. 101(36);13318-23.
167. Pernestig AK, Melefors O, Georgellis D. (2001). Identification of UvrY as the cognate response regulator for the BarA sensor kinase in *Escherichia coli* J. Biol. Chem. 276(1): 225-31.

168. Pernestig AK, Georgellis D, Romeo T, Suzuki K, Tomenius H, Normark S, Melefors O. (2003). The *Escherichia coli* BarA-UvrY two-component system is needed for efficient switching between glycolytic and gluconeogenic carbon sources. J. Bacteriol. 185(3): 843-53.
169. Sahu SN, Acharya S, Tuminaro H, Patel I, Dudley K, LeClerc JE, Cebula TA, Mukhopadhyay S. (2003). The bacterial adaptive response gene, barA, encodes a novel conserved histidine kinase regulatory switch for adaptation and modulation of metabolism in *Escherichia coli*. Mol. Cell. Biochem. 253(1-2):167-77.
170. Pernestig AK, Georgellis D, Romeo T, Suzuki K, Tomenius H, Normark S, Melefors O. (2003). The *Escherichia coli* BarA-UvrY two-component system is needed for efficient switching between glycolytic and gluconeogenic carbon sources. J. Bacteriol. 185(3):843-53.
171. Romeo T. (1998). Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. Mol. Microbiol. 29(6):1321-30.
172. Baker CS, Eöry LA, Yakhnin H, Mercante J, Romeo T, Babitzke P. (2007). CsrA inhibits translation initiation of *Escherichia coli* hfq by binding to a single site overlapping the Shine-Dalgarno sequence. J. Bacteriol. 189(15):5472-81.
173. Aguilar C. (2009). Análisis preliminar de mutaciones puntuales ocurridas en genes reguladores seleccionadas durante evolución adaptativa por crecimiento rápido en glucosa en una cepa de *Escherichia coli* carente del sistema PTS. Instituto de Biotecnología, UNAM. Tesis de Maestría.
174. Campbell E.A., Muzzin O., Chlenov M., Sun J.L., Olson C.A., Weinman O., Trester-Zedlitz M.L., Darst S.A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. Mol. Cell. 9(3):527-39.
175. Rodriguez SM, Panjekar S, Van Belle K, Wyns L, Messens J, Loris R. (2008). Nonspecific base recognition mediated by water bridges and hydrophobic stacking in ribonuclease I from *Escherichia coli*. Protein Sci. 17(4):681-90.
176. Serres, M. H., Gopal, S., Nahum, L. A., Liang, P., Gaasterland, T., Riley, M. (2001). A functional update of the *Escherichia coli* K-12 genome. Genome Biol. 2(9): RESEARCH0035.
177. Mika F. & Hengge R. (2005). A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigmaS (RpoS) in *E. coli*. Genes Dev. 19(22): 2770-81.
178. Gaspar AH, Machner MP. (2014). VipD is a Rab5-activated phospholipase A1 that protects *Legionella pneumophila* from endosomal fusion. Proc. Natl. Acad. Sci. USA. 111(12):4560-5.
179. Yates C.M., Filippis I., Kelley L.A., Sternberg M.J.E. (2014). SuSPect: Enhanced prediction of single amino acid variant (SAV) phenotype using network features. Journal of Molecular Biology. In press.
180. Fried L, Behr S, Jung K. (2013). Identification of a target gene and activating stimulus for the YpdA/YpdB histidine kinase/response regulator system in *Escherichia coli*. J. Bacteriol. 195(4):807-15.
181. Behr S, Fried L, Jung K. (2014). Identification of a novel nutrient-sensing histidine kinase/response regulator network in *Escherichia coli*. J. Bacteriol. 196(11):2023-9.
182. Martínez K, de Anda R, Hernández G, Escalante A, Gosset G, Ramírez OT, Bolívar FG. (2008). Couitilization of glucose and glycerol enhances the production of aromatic compounds in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. Microb. Cell Fact. 7:1.
183. Chávez-Béjar MI, Lara AR, López H, Hernández-Chávez G, Martinez A, Ramírez OT, Bolívar F, Gosset G. (2008). Metabolic engineering of *Escherichia coli* for L-tyrosine production by expression of genes coding for the chorismate

- mutase domain of the native chorismate mutase-prephenate dehydratase and a cyclohexadienyl dehydrogenase from *Zymomonas mobilis*. Appl. Environ. Microbiol. 74(10):3284-90.
184. Escalante A, Calderón R, Valdivia A, de Anda R, Hernández G, Ramírez OT, Gosset G, Bolívar F. (2010). Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. Microb. Cell Fact. 12;9:21.
185. Finkel SE. & Kolter R. (1999). Evolution of microbial diversity during prolonged starvation. Proc. Natl. Acad. Sci. USA. 96(7):4023-4027.
186. Sutton A, Buencamino R, Eisenstark A. (2000). rpoS mutants in archival cultures of *Salmonella enterica* serovar typhimurium. J. Bacteriol. 182(16):4375-4379.
187. Vulic M & Kolter R. (2001). Evolutionary cheating in *Escherichia coli* stationary phase cultures. Genetics. 158(2):519-526.
188. Finkel SE. (2006). Long-term survival during stationary phase: evolution and the GASP phenotype. Nat. Rev. Microbiol. (2):113-20.
189. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. (1997). The complete genome sequence of *Escherichia coli* K-12. Science. 1997, 277(5331):1453–1462.
190. Morgan M, Anders S, Lawrence M, Aboyoun P, Pagès H, Gentleman R. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. Bioinformatics. 25:2607–8.
191. Ariyaratne PN. & Sung WK. (2011). PE-Assembler: de novo assembler using short paired-end reads. Bioinformatics. 27:167–74.
192. Tsai IJ, Otto TD, Berriman M. (2010). Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. Genome Biol. 11:R41.
193. Langmead B, Trapnell C, Pop M, Salzberg SL. (2009). Ultrafast and memoryefficient alignment of short DNA sequences to the human genome. Genome Biol. 10:R25.
194. Darling AC, Mau B, Blattner FR, Perna NT. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14:1394–1403.
195. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. (2009). Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics. 25:2071–2073.
196. Popendorf K, Tsuyoshi H, Osana Y, Sakakibara Y. (2010). Murasaki: a fast, parallelizable algorithm to find anchors from multiple genomes. PLoS One. 5:e12651.
197. Kent WJ. (2002). BLAT -the BLAST-like alignment tool. Genome Res. 12:656–664.
198. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics. 25:2283–2285.
199. Li H. & Durbin R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25(14):1754-60.
200. Datsenko, K. A. & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA. 97(12): 6640-5.

201. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett HJ, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J. & Wittwer CT. (2009). The MIQE Guidelines: Minimum information for publication of quantitative realtime PCR experiments. *Clinical Chemistry*. 55:611–622.
202. Livak KJ. & Schmittgen TD. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*. 25:402–408.
203. Galán JC, Turrientes MC, Baquero MR, Rodríguez-Alcayna M, Martínez-Amado J, Martínez JL, Baquero F. (2007). Mutation rate is reduced by increased dosage of mutL gene in *Escherichia coli* K-12. *FEMS Microbiol. Lett.* 275(2):263-9.
204. Bligh EG. & Dyer WJ. (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37(8):911-7.
205. de Rudder KE, Thomas-Oates JE, Geiger O. (1997). *Rhizobium meliloti* mutants deficient in phospholipid *N*-methyltransferase still contain phosphatidylcholine. *J Bacteriol.* 179(22):6921-8.

XII. Anexos:

Participación en congresos:

2013

5th Congress of European Microbiologists, (FEMS 2013). Leipzig, Germany.

Organizer: Federation of European Microbiological Societies.

Project: –Genetic changes during a laboratory adaptive evolution process allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system”

Mode: Poster presentation

2013

XV National Congress of Biotechnology and Bioengineering and the 12th International Symposium on the Genetics of Industrial Microorganisms (GIM-2013). Cancún, Q.R., México.

Organizer: Mexican Society of Biotechnology and Bioengineering

Project: –Genetic changes during a laboratory adaptive evolution process allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system”

Mode: Oral presentation

2011

II Congress of Biochemistry and Molecular Biology on Bacteria. Huatusco, Veracruz.

Organizer: Mexican Society of Biochemistry

Project:–Characterization of mutations effects in regulatory genes originated through an adaptive evolution process by rapid growth on glucose in an *Escherichia coli* strain lacking the PTS system”

Mode: Oral presentation

2011

III International Congress of Biology, Chemistry and Agronomy. Universidad Autónoma de Guadalajara. Zapopan, Jalisco

Organizer: Universidad Autónoma de Guadalajara

Project:–Characterization of mutations effects in regulatory genes originated through an adaptive evolution process by rapid growth on glucose in an *Escherichia coli* strain lacking the PTS system”

Mode: Symposium

2011

I Congress of Master and PhD students of UNAM. Ciudad Universitaria; México D.F.

Organizer: Universidad Nacional Autónoma de México.

Project:–Characterization of mutations effects in regulatory genes originated through an adaptive evolution process by rapid growth on glucose in an *Escherichia coli* strain lacking the PTS system”

Mode: Oral presentation

2010

I Congress of Biochemistry and Molecular Biology in Bacteria. San Miguel Regla, Hidalgo.

Organizer: Mexican Society of Biochemistry

Project:—Preliminary analysis of mutations originated in regulatory genes during an adaptive evolution process for fast growing on glucose in a *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.

Mode: Oral presentation

2009

XIII National Congress of Biotechnology and Bioengineering and VII International Symposium of alcohol production and yeast (SIPAL). Acapulco, Guerrero.

Organizer: Mexican Society of Biotechnology and Bioengineering

Project:—Preliminary analysis of mutations in regulatory genes generated during an adaptive evolution process in an *Escherichia coli* strain lacking the PTS system”

Mode: Poster

Artículos publicados:

Aguilar C, Escalante A, Flores N, De Anda R, Riveros-McKay F, Gosset G, Morett E, Bolívar F. 2012. Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system. BMC Genomics, 13:385.

Sigala JC, Flores S, Flores N, Aguilar C, de Anda R., Gosset G, Bolívar F. 2008. Acetate Metabolism in *Escherichia coli* Strains Lacking Phosphoenolpyruvate: Carbohydrate Phosphotransferase System; Evidence of Carbon Recycling Strategies and Futile Cycles. J. Mol. Microbiol. Biotechnol. 16(3-4): 224-35.

RESEARCH ARTICLE

Open Access

Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system

César Aguilar¹, Adelfo Escalante^{1*}, Noemí Flores¹, Ramón de Anda¹, Fernando Riveros-McKay^{1,2}, Guillermo Gosset¹, Enrique Morett¹ and Francisco Bolívar¹

Abstract

Background: *Escherichia coli* strains lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system (PTS), which is the major bacterial component involved in glucose transport and its phosphorylation, accumulate high amounts of phosphoenolpyruvate that can be diverted to the synthesis of commercially relevant products. However, these strains grow slowly in glucose as sole carbon source due to its inefficient transport and metabolism. Strain PB12, with 400% increased growth rate, was isolated after a 120 hours adaptive laboratory evolution process for the selection of faster growing derivatives in glucose. Analysis of the genetic changes that occurred in the PB12 strain that lacks PTS will allow a better understanding of the basis of its growth adaptation and, therefore, in the design of improved metabolic engineering strategies for enhancing carbon diversion into the aromatic pathways.

Results: Whole genome analyses using two different sequencing methodologies: the Roche NimbleGen Inc. comparative genome sequencing technique, and high throughput sequencing with Illumina Inc. GAllx, allowed the identification of the genetic changes that occurred in the PB12 strain. Both methods detected 23 non-synonymous and 22 synonymous point mutations. Several non-synonymous mutations mapped in regulatory genes (*arcB*, *barA*, *rpoD*, *rna*) and in other putative regulatory loci (*yjjU*, *rssA* and *ypdA*). In addition, a chromosomal deletion of 10,328 bp was detected that removed 12 genes, among them, the *rppH*, *mutH* and *galR* genes. Characterization of some of these mutated and deleted genes with their functions and possible functions, are presented.

Conclusions: The deletion of the contiguous *rppH*, *mutH* and *galR* genes that occurred simultaneously, is apparently the main reason for the faster growth of the evolved PB12 strain. In support of this interpretation is the fact that inactivation of the *rppH* gene in the parental PB11 strain substantially increased its growth rate, very likely by increasing glycolytic mRNA genes stability. Furthermore, *galR* inactivation allowed glucose transport by GalP into the cell. The deletion of *mutH* in an already stressed strain that lacks PTS is apparently responsible for the very high mutation rate observed.

Keywords: PTS system, Resequencing, Laboratory evolution, Mutation, Deletion, Adaptation, *rppH*, Glycolysis

* Correspondence: adelfo@ibt.unam.mx

¹Departamento de Ingeniería Celular y Biotecnología, Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Cuernavaca, Morelos 62210, México

Full list of author information is available at the end of the article

Background

Genome changes including point mutations, duplications, and recombination with homologous and heterologous DNA are the driving force of evolution. Bacteria, the most diverse and adapted types of cells in the biosphere, have permanently been evolving during millions of years to survive different environmental changes. Comparative genomics is a very powerful tool to analyze bacterial evolution occurring over short periods of time [1-3]. Moreover, whole genome resequencing of evolved *Escherichia coli* strains using simultaneously two different methodologies has recently been reported. This strategy is certainly very useful for understanding bacterial evolution, such as pathogen emergence, adaptation to environmental perturbations or during fermentation events used to generate derivative strains with enhanced industrial capacities [3-5].

We have constructed and characterized *Escherichia coli* strains that lack the phosphoenolpyruvate: carbohydrate phosphotransferase system (PTS), by deletion of the *ptsH*, *ptsI*, and *crr* genes, which is the major bacterial component involved in glucose transport and its phosphorylation. One of these strains, PB11, in spite of growing very slow in glucose (with a specific growth rate (μ) = 0.1 vs. 0.7 h⁻¹ as compared to the parental strain JM101), accumulates high amounts of phosphoenolpyruvate, which can be diverted to the synthesis of aromatic compounds. PTS deletion results in a carbon stress response when the PB11 strain is grown in glucose as the sole carbon source that induces carbon scavenging. Strains lacking PTS can co-utilize several carbon sources due to the lack of catabolite repression exerted by PTS, and their glycolytic flux is reduced as part of a carbon limitation response [6-13]. As a metabolic engineering strategy, an adaptive laboratory evolution process for the selection of faster growing derivatives of the PB11 strain was carried out in a fermentor in minimal medium with glucose as the sole carbon source. In this process, after entering the stationary phase this carbohydrate was fed by progressively increasing the dilution rate. The resulting strain, PB12, which achieved a very reasonable growth rate (μ = 0.44 h⁻¹), was selected in a process that lasted 120 hours (hr) (Figure 1) [9,10,12,13]. The evolved PB12 strain that in the absence of PTS uses the galactose permease (GalP), as the parental PB11 strain for glucose transport, has been utilized for overproduction of aromatic compounds [7,9,12,14-17].

It is well known that *E. coli* cells can adapt their metabolism to achieve higher growth rates as a result of specific mutations [2,5,18]. To get insights of the faster growth of the PB12 strain, we have compared its transcript levels with those of the parental PB11 strain, by reverse transcriptase quantitative real time PCR (RT-qPCR), of critical metabolic pathways. Interestingly,

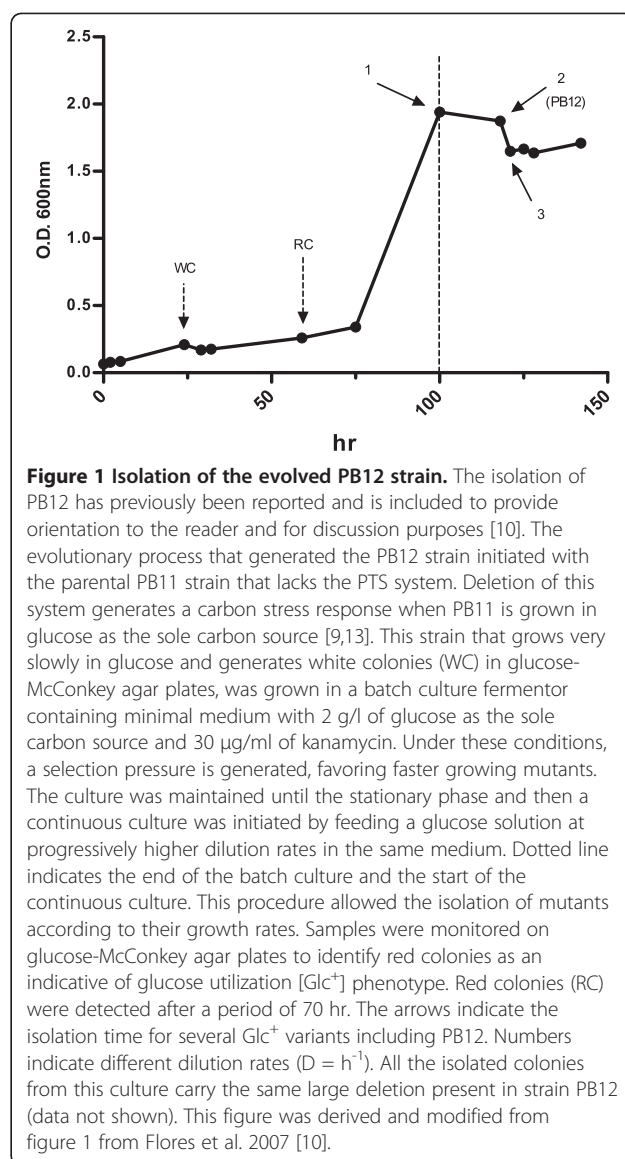


Figure 1 Isolation of the evolved PB12 strain. The isolation of PB12 has previously been reported and is included to provide orientation to the reader and for discussion purposes [10]. The evolutionary process that generated the PB12 strain initiated with the parental PB11 strain that lacks the PTS system. Deletion of this system generates a carbon stress response when PB11 is grown in glucose as the sole carbon source [9,13]. This strain that grows very slowly in glucose and generates white colonies (WC) in glucose-McConkey agar plates, was grown in a batch culture fermentor containing minimal medium with 2 g/l of glucose as the sole carbon source and 30 μ g/ml of kanamycin. Under these conditions, a selection pressure is generated, favoring faster growing mutants. The culture was maintained until the stationary phase and then a continuous culture was initiated by feeding a glucose solution at progressively higher dilution rates in the same medium. Dotted line indicates the end of the batch culture and the start of the continuous culture. This procedure allowed the isolation of mutants according to their growth rates. Samples were monitored on glucose-McConkey agar plates to identify red colonies as an indicative of glucose utilization [Glc⁺] phenotype. Red colonies (RC) were detected after a period of 70 hr. The arrows indicate the isolation time for several Glc⁺ variants including PB12. Numbers indicate different dilution rates ($D = h^{-1}$). All the isolated colonies from this culture carry the same large deletion present in strain PB12 (data not shown). This figure was derived and modified from figure 1 from Flores et al. 2007 [10].

we found that all glycolytic and several other central carbon metabolism genes, including those that code for the tricarboxylic acid (TCA) cycle enzymes, are overexpressed, suggesting a very efficient carbon utilization by the evolved strain [7-13,19]. We have previously shown that a mutation in the *arcB* gene could be responsible for the overexpression of the TCA genes [9,20-22]. In addition a second mutation responsible of amber stop codon at position 98 in the *rpoS* gene which codes for the sigma factor RpoS, was detected in PB12 when compared against strain MG1655 [9,11]. Nevertheless, to get a detailed knowledge at the molecular level, of all the different genetic changes that occurred in the PB12 strain, a complete genomic analysis is required. This information will allow a better understanding of the basis of growth adaptation, plasticity, and the physiology of

this evolved *E. coli* strain, and also will be useful in the design of improved laboratory adaptive evolution and metabolic engineering strategies for enhancing carbon diversion into the aromatic pathway utilizing strains lacking PTS.

In this work, using the Roche NimbleGen Inc. comparative genome sequencing technique (CGS) and high throughput sequencing with Illumina Inc. GAIIX, we identified all the genetic changes that occurred in the evolved PB12 strain during the selection process and analyzed and characterized the most relevant ones. Results of the whole genome sequencing, supported by transcript quantification by RT-qPCR and by knockout inactivation of selected genes in the parental PB11 strain, indicate that a simultaneous deletion of several contiguous genes including *rppH*, *mutH* and *galR*, is the main reason for the fast growth in glucose. *galR* codes for the repressor of the *gal* operon that includes *galP* that codes for the GalP permease [23], *rppH* codes for the RNA pyrophosphohydrolase (RppH), which initiates mRNA degradation [24], while *mutH* codes for the endonuclease of the MutHLS system involved in the mismatch DNA repair system [25]. In addition, several non-synonymous point mutations were detected as one located in the RNase I coding gene *rna*, involved in the degradation of RNA [26,27], while others were located in known and in putative regulatory genes, such as *arcB* [21,22], *barA* [28,29] *rpoD* [30], *rssA* [31] and *yjjL* [32,33]. Finally, other mutation was mapped on *ypdA*, which code for a putative histidine kinase [34].

Results and discussion

Detection and characterization of non-synonymous point mutations in the evolved PB12 strain

Two comparative whole genome nucleotide sequence analyses of the evolved PB12, its parental JM101, and the wild type K-12 MG1655 strains were performed. The first was carried out by Roche NimbleGen Inc., Madison, WI (RN), using their CGS method; the second analysis was developed by Winter Genomics Inc., Mexico City (WG), using Illumina's massively parallel sequencing technology (see Materials and Methods). In the RN analysis, 26 non-synonymous point mutations were detected in structural genes; 21 of them were also mapped at the same positions by WG. In addition, 6 non-synonymous mutations were detected only by WG (Table 1A and Tables S1 and S2 presented in Additional file 1 and Additional file 2). Since there was some discrepancies between the two technologies that utilized DNA obtained from the original frozen stock of the PB12 strain (see Materials and Methods), we decided to sequence each of the mutant genes solely reported by one company, after PCR amplification using the Sanger methodology. Only two of the mutants detected by WG

(in the *csfF* and *ytfR* genes) could be confirmed by Sanger resequencing. The mutations in the other four genes (*ftsK*, *stfE*, *C0362* and *rsxC*) reported by WG and the five other genes (*mdlB*, *yagN*, *ydfN*, *ykfA* and *yagG*) reported by RN (Tables S1 and S2 presented in Additional file 1 and Additional file 2), could not be confirmed by the Sanger methodology (data not shown). Therefore, the total number of non-synonymous point mutations detected, comprised 21 reported by both companies, plus 2 additional mutations reported only by WG. 14 of the 21 common mutations including the ones located in regulatory genes (see below), were confirmed by Sanger resequencing (Table 1A). Recently, it has been shown that when both types of methodologies were utilized simultaneously for whole genome resequencing of *E. coli* strains in which growth adaptations by evolution occurred, both techniques reported false positive mutations [4,5]. Therefore, it is likely that the mutations in the PB12 strain not confirmed by the Sanger methodology are false positives. However, since the *mutH* gene deletion in this strain is responsible of increasing the mutation rate in *E. coli* [35] (see below), it could be possible that the two non-synonymous mutations detected only by WG, and confirmed by Sanger, are due to *de novo* changes that occurred in the overnight culture utilized to obtain DNA for genome analysis by WG and not in the fermentation process started with the PB11 strain (Figure 1). Alternatively, these two mutations could be real, but they were not detected by the RN analysis. Importantly, the nucleotide sequence of the parental PB11 strain was also determined by WG (data not shown). None of the point mutations that occurred in PB12 were detected in PB11, indicating that they appeared in the laboratory adaptive evolution process.

Among the non-synonymous substitutions detected some were located at genes with regulatory functions: *arcB*, *barA*, *rpoD*, *rna*, and three in putative regulatory genes: *yjjL*, *rssA* and *ypdA* (Table 1A). The mutation in the *arcB* gene, a tyrosine to cysteine residue substitution at position 71 that apparently modifies the ArcA/B ability to function as a repressor, has previously been reported by our group. We have proposed that this modification could be responsible for the overexpression of the TCA genes in the PB12 derivative as compared to the parental PB11 strain [9,20,21] (Table 1A). A new role for this mutation is proposed (see next sections).

The *barA* mutation resulted in a phenylalanine to leucine residue substitution at position 366 (Table 1A). This residue is located between two functional subregions of the HK domain of this protein; the "H box", where the conserved histidine residue involved in the autophosphorylation of BarA is located, and the "N box", involved in ATP binding [28,29].

Table 1 Mutations that occurred in the evolved PB12 strain during the adaptive process

A) Gene	Basic description	Mutations			
Hypothetical genes		JM101 and PB11	PB12	Nucleotide	Pos/Change
<i>gfcD*</i>	Hypothetical lipoprotein.	GTA	cTA	916	306 V-L
<i>yafV*</i>	Predicted C-N hydrolase family amidase, NAD(P)-binding.	CGC	CaC	452	151 R-H
<i>+ytfR^(WG)</i>	Putative ATP-binding component of a galactose ABC transporter.	GTC	GcC	602	201 V-A
Metabolism and Transport		JM101 and PB11	PB12	Nucleotide	Pos/Change
<i>+actP*</i>	Acetate/glycolate permease in the solute:sodium symporter (SSS) family.	GTA	GcA	731	244 V-A
<i>arnT*</i>	L-Ara4N transferase catalyzes addition of L-Ara4N to lipid A under some conditions.	TAC	TgC	1193	398 Y-C
<i>+chbC*</i>	Integral membrane transport protein, member of the chitobiose PTS transporter.	GCG	GtG	347	116 A-V
<i>+csgF^(WG)</i>	Curli secretion and assembly complex.	AAU	AgU	89	30 N-S
<i>+dgoT*</i>	Probable galactonate transporter, member of the major facilitator superfamily (MFS) of transporters.	GAT	GgT	653	218 D-G
<i>dhaM*</i>	Dihydroxyacetone kinase subunit M, homologous to certain PTS components.	TGG	TGa	1038	346 W-stop
<i>+dppF*</i>	ATP-binding component of the dipeptide ABC transporter.	CAC	CgC	680	227 H-R
<i>fdhD*</i>	Protein with unknown function, required for wild-type formate dehydrogenase-N activity.	AGT	gGT	652	218 S-G
<i>fimH*</i>	Minor fimbrial subunit, D-mannose specific adhesin, subunit of fimbrial complex.	ACT	gCT	535	179 T-A
<i>glpT*</i>	Major uptake transporter for glycerol-3-phosphate, belongs to the major facilitator superfamily (MFS).	CCG	CtG	416	139 P-L
<i>rfe*</i>	Undecaprenyl-phosphate α-N-acetylglucosaminyl transferase .	GCC	aCC	238	80 A-T
<i>+sucA*</i>	E1(0) component of the oxoglutarate dehydrogenase complex.	GAT	GAa	372	124 D-E
<i>ydiQ*</i>	Putative subunit of YdiQ-YdiR flavoprotein.	GTG	GcG	218	73 V-A
Regulatory and putative regulatory genes		JM101 and PB11	PB12	Nucleotide	Pos/Change
<i>+arcB*</i>	Hybrid protein sensory kinase, member of two-component system.	TAC	TgC	212	71 Y-C
<i>+barA*</i>	Hybrid protein sensory kinase, member of two-component system.	TTC	cTC	1096	366 F-L
<i>+rna*</i>	RNase I, cleaves phosphodiester bond between any two nucleotides.	GCC	aCC	268	90 A-T
<i>+rpoD*</i>	Sigma 70 factor, subunit of RNA polymerase.	GTT	aTT	1744	582 V-I
<i>+rssA*</i>	Hypothetical protein, coded by a gene that is part of the <i>rssAB</i> operon.	CGC	CaC	773	258 R-H
<i>+yjiU*</i>	Conserved hypothetical protein, could work as a transcriptional protein.	ACT	gCT	338	179 T-A
<i>+ypdA*</i>	Predicted sensory kinase member of two-component system.	GCG	tCC	598	200 A-S
B) Gene	Basic description	Mutations			
<i>ptsP</i>	Member of a second PTS chain involved in nitrogen metabolism	Present	Absent	—	—
<i>rppH</i>	RNA pyrophosphohydrolase that initiates mRNA degradation by hydrolysis of the 5'-triphosphate end.	Present	Absent	—	—
<i>ygdT</i>	Hypothetical protein.	Present	Absent	—	—
<i>mutH</i>	dGATC endonuclease in the MuthLS complex, the methyl-directed mismatch repair pathway.	Present	Absent	—	—
<i>ygdQ</i>	Putative transport protein.	Present	Absent	—	—
<i>ygdR</i>	Predicted protein.	Present	Absent	—	—
<i>tas</i>	Putative NAD(P)-linked reductase that acts in starvation-associated mutations.	Present	Absent	—	—
<i>lplT</i>	Lysophospholipid transporter (LplT).	Present	Absent	—	—
<i>aas</i>	2-acylglycerophosphoethanolamine acyltransferase/acyl-ACP synthetase.	Present	Absent	—	—
<i>omrA</i>	Small RNA that is involved in regulating the protein composition of the outer membrane.	Present	Absent	—	—

Table 1 Mutations that occurred in the evolved PB12 strain during the adaptive process (Continued)

<i>omrB</i>	Small RNA that is involved in regulating the protein composition of the outer membrane.	Present	Absent	—	—
<i>galR</i>	DNA-binding transcription factor; represses transcription of the operons involved in transport and catabolism of D-galactose.	Present	Absent	—	—

Table 1A presents the 23 non-synonymous point mutations in structural genes that changed the code for a different amino acid when compared to the parental PB11 and JM101 strains genomes. The asterisks in the table indicate the 21 non-synonymous point mutations detected by both Roche NimbleGen Inc. (RN) and Winter Genomics Inc. (WG) methods. In addition 22 synonymous point mutations in different genes were also detected (Tables S1 and S2 in Additional file 1 and Additional file 2). The regulatory and possible regulatory genes analyzed in this study are in bold letters. The mutations in these seven regulatory and possible regulatory genes and in seven additional genes were confirmed by the Sanger methodology and are labeled with a +. 26 (21+5) point mutations in 26 genes were detected by RN, 5 of them were false positive. 27 (21+6) point mutations in 27 genes were detected by WG. In addition to the 21 mutations shared with RN, only two of these point mutations (noted as WG) were confirmed by the Sanger methodology. Additional file 1 and Additional file 2 (Tables S1 and S2) include the data from RN and WG. Table 1B presents the absent genes in the largest deletion reported by both companies. (Figures. 2, 3, 4 and S1, presented in Additional file 3). Basic descriptions of these genes were taken from www.ecocyc.org.

The change in the *rpoD* gene resulted in a valine to isoleucine residue substitution at position 582, which is located in a helix-turn-helix (HTH) motif of the RpoD coded protein (Table 1A). This motif is involved in the binding to the -35 promoter region, but it is unlikely that the conservative nature of the substitution, and the fact that this residue does not make any direct contact with the DNA [36], had any significant effect on promoter recognition.

The mutation in the *rna* gene, an alanine to threonine residue substitution at position 90 of the coded RNase I, is also unlikely to have any consequence, since it is located in a nonconserved structural part outside the catalytic region [37].

The *yjjU* mutation resulted in a threonine to alanine residue substitution at position 179 in the coded YjjU protein (Table 1A). It has been proposed that this protein could have a regulatory role [32,33].

The mutation in the *rssA* gene resulted in an arginine to histidine residue substitution at position 258 of its product. The importance of this mutation, as well as the function of the RssA protein, are unknown; however, it might be functionally related to the RssB protein, which is involved in the degradation of the RpoS, since both conserved proteins are coded by genes located in the same operon [31]. An esterase function has also been predicted for RssA [Uniprot.org].

The point mutation in the *ypdA* gene located at position 200, caused an alanine to serine residue substitution. This gene codes for a predicted sensory histidine kinase of the two-component system YpdA-YpdB [34].

The products of most of the other mutated genes (*sucA*, *dhaM*, *ydiQ*, *chbC*, *glpT*, *arnT*, *dppF*, *rfe*, *fdhD*, *fimH*, *csgF*, *dgoT*, and *actP*) detected by both companies and the two mutations detected only by WG (*csgF* and *ytfR*), are involved in bacterial metabolism and transport (Table 1A).

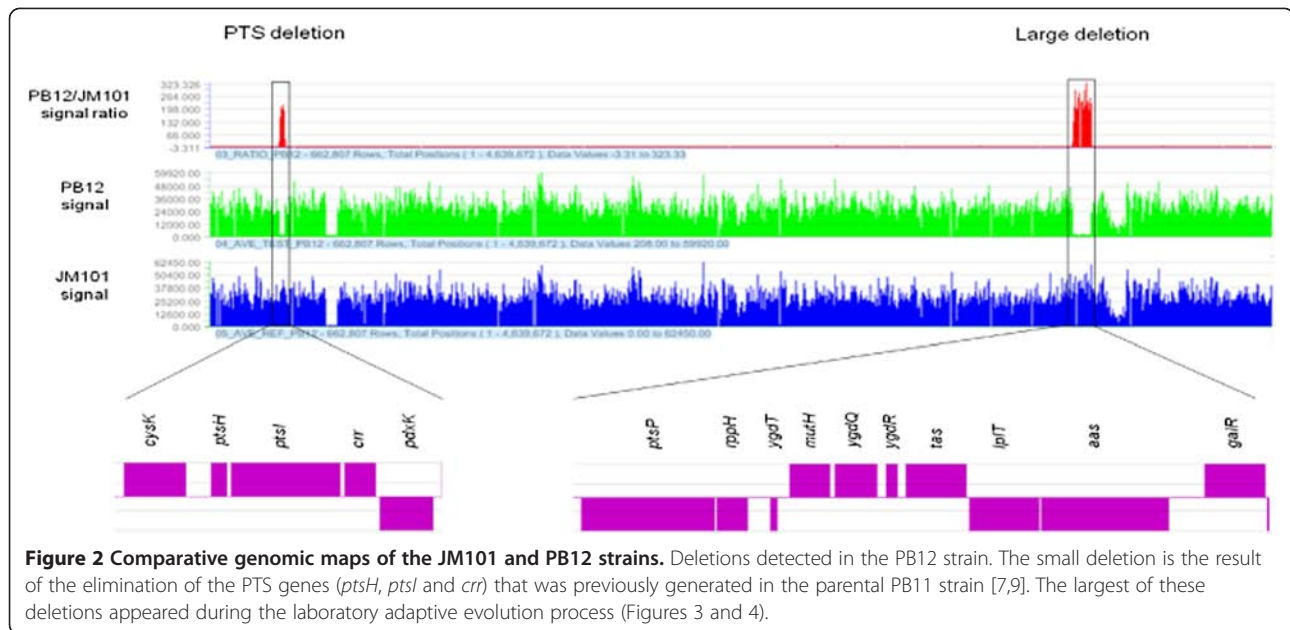
In addition, 22 synonymous mutations were reported; 16 of them detected by both companies. These mutations were not analyzed since it is unlikely that they could have any significant effect on the phenotype (Tables S1 and S2

in Additional file 1 and Additional file 2). Also, several point mutations were detected in non-coding regions by WG. It should be noted that these unconfirmed mutations are unlikely to be located in regulatory regions (data not shown).

The presence of a mutation in the *rpoS* gene in the PB12 strain has been reported. This change generates a stop codon instead of a glutamine coding residue at position 98. It is known, that this strain being a derivative of JM101, carries a *supE* mutation, which suppresses amber stop codons [11,38,39]. Originally this mutation was considered to have occurred during the adaptive evolution process since the change was detected when compared against the sequence of strain MG1655 [9,11,40]. However, both comparative genome sequencing strategies showed that this mutation was already present in the parental PB11 and the JM101 strains. Sanger resequencing confirmed the presence of this mutation in both of the parental strains (data not shown). So, this mutation was incorporated at sometime during the development of the JM101 strain from its parental strains JC3130, CSH51 and 71.18 [38,39].

Detection, characterization of a chromosomal deletion in the evolved PB12 strain, and analysis of the effects of the deleted genes

A deletion of 10,328 bp located at minute 64 on the chromosome of the PB12 strain that removed simultaneously the *rppH*, *ygdT*, *mutH*, *ygdQ*, *ygdR*, *tas*, *lplT*, *aas*, *omrA*, *omrB*, and the part of *ptsP* and *galR* genes, was detected by both RN and WG analyses (Table 1B and Figure 2). This deletion was confirmed by PCR (Figure 3) and its limits mapped within the *ptsP* and *galR* genes, resulting in a fusion of the remaining segments of these two genes (Figure 4). The nucleotide sequence of the fused fragments was confirmed by Sanger and it is presented in Additional file 3 (Figure S1). Neither repeated sequences nor insertion sequences were detected in the chromosomal DNA regions flanking the deleted genes; therefore, the molecular bases of this deletion are unknown.



The analyses of three of these genes, *galR*, *mutH* and *rppH* whose deletion could be the main cause (see below) of the faster growth in glucose of the evolved PB12 strain is presented and discussed in the following sections.

The *galR* gene codes for the repressor of the *gal* regulon that includes the *galP* gene [23]. Therefore, the inactivation of this gene in the PB12 strain is apparently responsible for the high transcription levels of most of the *gal* genes. In agreement with this proposition, we have previously demonstrated that the PB12 strain that lacks PTS is dependent on GalP for glucose transport [9,12].

The *rppH* gene codes for an RNA pyrophosphohydrolase that initiates mRNA degradation by hydrolysis of the 5' triphosphate-end. After this modification, RNase E can initiate further mRNA degradation. It has been reported that the level of 382 transcripts increased significantly in *E. coli* cells lacking RppH [24]. Accordingly, in the PB12 strain higher transcript levels of many genes, among them glycolytic and TCA genes were detected by RT-qPCR as compared to the parental JM101 and PB11 strains. We could not explain the basis of the simultaneous “overexpression” of the genes involved in these metabolic pathways (with the exception of the TCA cycle genes due to the *arcB* mutation, as previously mentioned) [9,11]. In the light of the phenotype of the strains lacking *rppH*, it is likely that the higher levels of the transcripts observed in the PB12 strain is a result of impaired mRNA degradation instead of overexpression (see below). In fact, *rppH* inactivation in the PB11Δ*rppH* strain increased its μ 261% (Table 2). In agreement with this μ increment, as it will be presented in detail in the next sections, the RT-qPCR values of most central metabolic genes were increased 1.5- to 18-fold when the

rppH gene was inactivated in the PB11Δ*rppH* derivative as compared to the parental PB11 strain, allowing an improved glucose metabolism, probably due to higher levels of central metabolic genes transcripts (see next sections). Interestingly, in contrast to the PB11 strain, inactivation of the *rppH* gene in the parental JM101Δ*rppH* strain decreased 27% its μ (data not shown), indicating that this mutation disrupted the glucose metabolism in this strain that does not have impaired the glucose transport. From these results it is tempting to speculate that inactivation of *rppH* could be considered as a tool for increasing the half-life of certain mRNAs families, and therefore growth rates, in strains with certain limited growth conditions.

The absence of the *mutH* gene, which codes for the MutH endonuclease that is part of the MutHLS complex involved in the DNA mismatch repair pathway, is probably responsible for the appearance of the large number of mutations detected in the PB12 strain during the short term adaptive laboratory evolution process that lasted for only 120 hr [5,7,9,10,12,18,25]. It is known that the absence of the *mutH* gene increases the mutation frequency in *E. coli* at least 200 fold [35]. In addition, starvation-induced mutagenesis among hundreds of *E. coli* natural isolates is increased on average 7-fold, but in certain strains up to 1000-fold [41-43]. The presented and analyzed information indicates that the 21 shared non-synonymous point mutations as well as the deletion are real mutations that appeared after a 120 hr period in this adaptive laboratory evolution process. A possible explanation for this high number of mutations is that the deletion of the *mutH*, *rppH* and *galR* genes occurred early in the very short laboratory

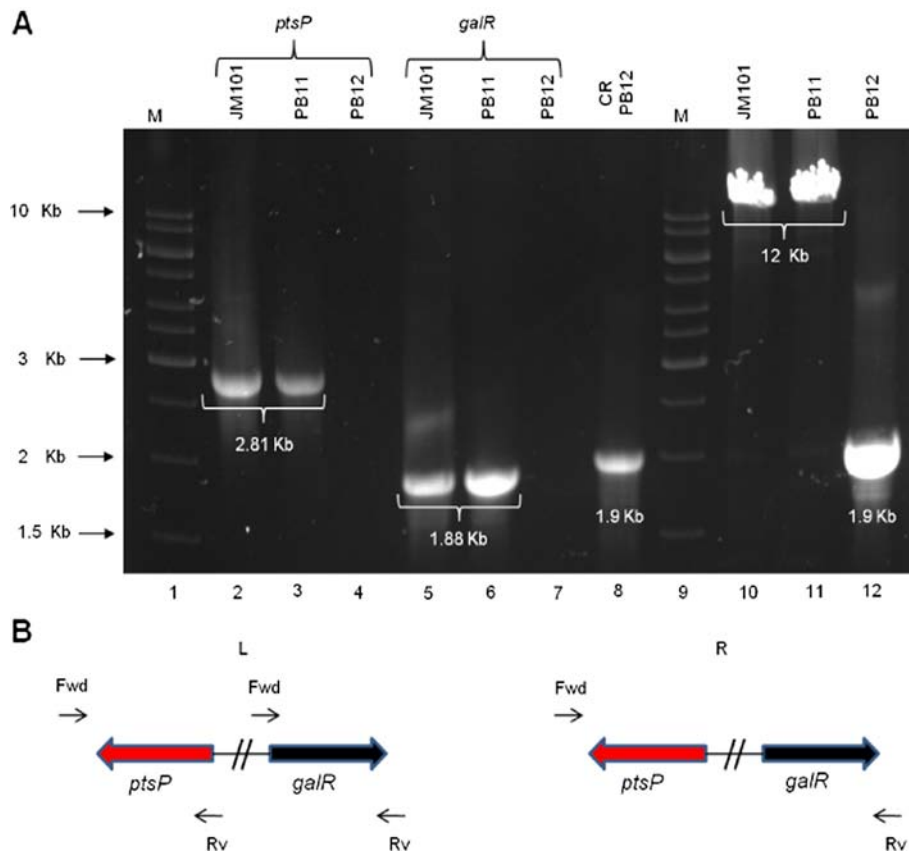


Figure 3 Chromosomal deletion markers in the PB12 strain. The absence of a chromosomal fragment in the PB12 strain was confirmed by PCR and by Sanger resequencing. Ten genes were deleted and the *galR* and *ptsP* genes were fused. Section **A** shows the *ptsP* and *galR* genes that were amplified in the JM101, PB11 and PB12 strains: line 1, (M) molecular weight markers; lines 2, 3 and 4, *ptsP* amplification in the JM101, PB11 and PB12 strains, respectively; lines 5, 6 and 7, *galR* amplification in the JM101, PB11 and PB12 strains, respectively; line 8, amplification of the chromosomal region in the PB12 strain; line 9, (M) molecular weight markers; lines 10, 11 and 12, amplification of the chromosomal region using DNA from strains JM101, PB11 and PB12, respectively. Section **B** presents the oligonucleotides utilized for DNA amplifications. The left section (L) includes the oligonucleotides employed for the amplification of the *ptsP* and *galR* genes of the three strains (lines 2–7), and the right section (R) presents the entire chromosomal regions of the same three strains amplified using *ptsP*-fwd and *galR*-rv oligonucleotides (lines 8, 10–12). The nucleotide sequences of the oligos utilized are included in table S3 presented in additional file 4.

evolutionary process allowing faster growth by itself. In the absence of *mutH*, some or several of the point mutations could then have occurred during a few replication cycles, favouring the selection of even faster growing variants in glucose. In addition, in this genetic background in which the PTS deletion generates a carbon stress response, further mutagenesis induction is expected [9,10,12,41,42,44,45].

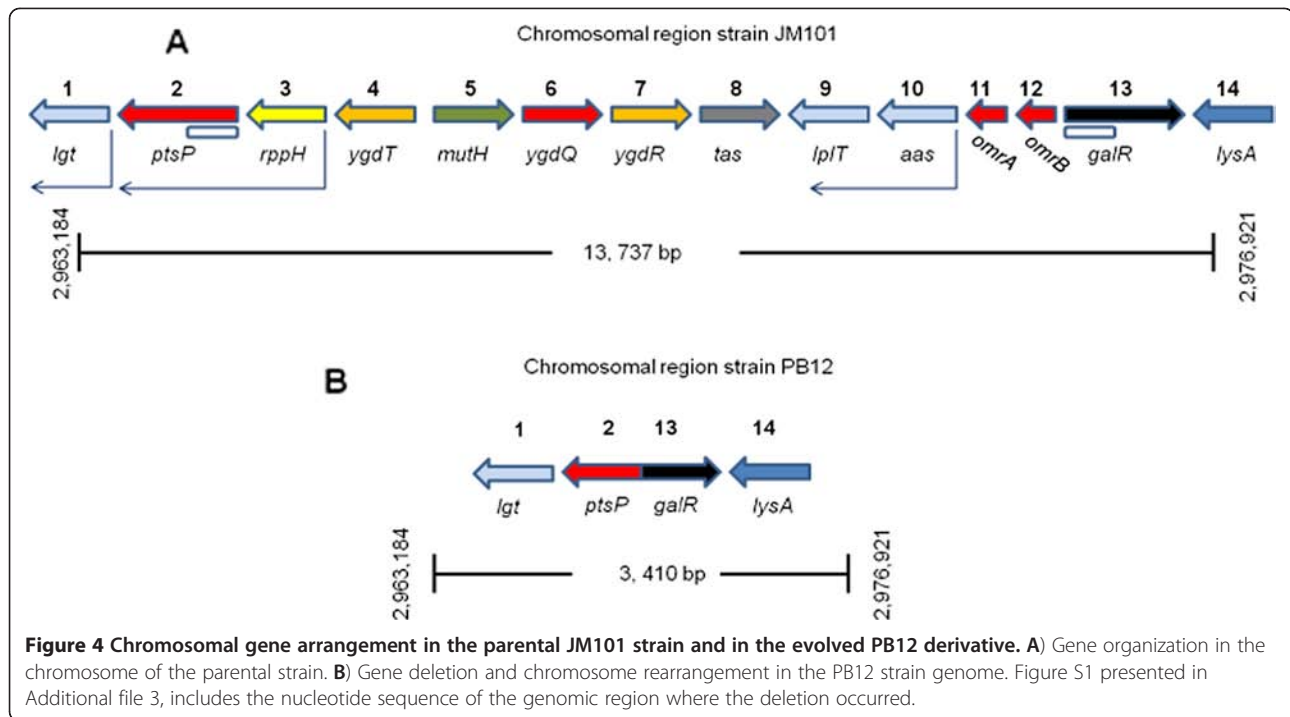
RT-qPCR mRNA expression values of relevant mutated genes and of central metabolic genes in the evolved PB12 strain

Table 3 shows mRNA expression values of relevant (mainly regulators or possible regulators) mutated genes and of some other genes as controls determined by RT-qPCR (see Materials and Methods). RT-qPCR values of more than 100 genes from the evolved PB12 strain, the parental PB11 and

JM101 strains have previously been reported by our group, including the *arcB* gene in which a non-synonymous point mutation appeared, and are presented in Table 3 for comparison and analyses purposes [9–11]. Results indicate that the RT-qPCR values of some of the regulatory and possible regulatory genes were not substantially modified (except for *rpoS* and *barA*) with respect to the JM101 parental strain. As anticipated, except for *galR*, no transcripts of the 12 deleted contiguous genes were detected.

Inactivation of mutated regulatory and putative regulatory genes in the PB11 and PB12 strains. RT-qPCR values of carbon central metabolism genes of the PB11ΔrppH strain

With the aim of understanding the possible roles of some of the regulatory and possible regulatory genes



mutated in PB12, as well as in its parental PB11 and JM101 strains (Table 1A), they were individually inactivated by a cassette insertion using the Datsenko-Wanner method [46] (see Materials and Methods), with the exceptions of the *rpoD* gene, since its inactivation is lethal [47], and the *galR* gene, since its regulated target, the *galP* gene, is already overexpressed in the PB11 and PB12 strains [9,23]. Interestingly, as shown in Table 2, individual cassette inactivation of the *arcA*, *rppH*, *yjiU*, and *rssA* genes in the PB11 strain increased their specific growth rates in glucose as the only carbon source,

Table 2 Specific growth rates (μ) of the PB11 strain and its derivatives

Regulatory genes		
Strain	μ (h^{-1})	%
PB11	0.13 ± 0.001	100
PB11 Δ <i>barA</i>	0.12 ± 0.001	92
PB11 Δ <i>rppH</i>	0.34 ± 0.019	261
PB11 Δ <i>rpoS</i> *	0.06	46
PB11 Δ <i>arcA</i> *	0.32	246
Possible regulatory genes		
Strain	μ (h^{-1})	%
PB11 Δ <i>yjiU</i>	0.16 ± 0.005	123
PB11 Δ <i>rssA</i>	0.17 ± 0.003	130
PB11 Δ <i>rna</i>	0.13 ± 0.007	100
PB11 Δ <i>ypdA</i>	0.13 ± 0.002	100

*These values have been previously reported [9-11].

supporting our hypothesis that the inactivation of some of these genes, especially *rppH*, was the result of direct selection during the evolution process for faster growth.

Some of these regulatory genes were also inactivated by the Datsenko-Wanner method in the PB12 strain; the effects on their respective specific growth rate values are shown in Table 4. Knockout inactivation of the *barA*, *arcA*, and *yjiU* genes decreased the μ by about 5%, 10%, and 23%, respectively, while no substantial growth rate difference were observed when the *rna*, *rssA*, or *ypdA* genes were inactivated.

Since the inactivation of the *rppH* gene in the PB11 Δ *rppH* strain is responsible for a markedly (261%) μ increment, the expression of several genes, including carbon central metabolism, transport and regulators in this strain, as well as in the JM101 Δ *rppH* strain, were determined (Table 3). In PB11 Δ *rppH*, most of the RT-qPCR values of the central metabolism genes, including TCA (except *pgk*, *fbaA*, *talB*, *pckA*), increased from 2- to 18-fold, as compared to PB11. RT-qPCR values of all genes involved in growth under stress-limited carbon conditions, which most of them are overexpressed in the PB11 strain due to the lack of PTS (such as the *gal* operon, *poxB*, *acs*, and the glyoxylate shunt genes), were also high in the PB11 Δ *rppH* strain and, for some genes, the increase was up to 10-fold, as for the *maeB* gene (Table 3) [8-12]. The mRNA levels of all these previously mentioned genes, except for *aceEF*, *fbaA*, *eno*, *pckA*, *pfkA*, *pykA* and *pgk*, in the PB11 Δ *rppH* strain were higher than in the parental JM101 strain. Importantly,

Table 3 RT-qPCR values of central metabolism and regulatory genes

Gene	PB11*	PB12*	JM101ΔrppH	PB11ΔrppH	PB11ΔrpoS*	PB12ΔrpoS*	PB11ΔarcA*
Glycolysis							
<i>aceE</i> +	0.40 ± 0.02	1.40 ± 0.02	1.57 ± 0.23	1.32 ± 0.21	0.20 ± 0.00	0.85 ± 0.00	0.40 ± 0.00
<i>aceF</i> +	0.60 ± 0.00	1.20 ± 0.00	2.66 ± 0.51	1.00 ± 0.12	0.40 ± 0.10	0.70 ± 0.00	0.60 ± 0.00
<i>eno</i>	0.50 ± 0.10	0.50 ± 0.14	0.89 ± 0.01	1.28 ± 0.41	0.20 ± 0.00	0.37 ± 0.00	0.30 ± 0.00
<i>fbaA</i>	0.90 ± 0.00	1.10 ± 0.05	0.82 ± 0.08	1.46 ± 0.18	0.90 ± 0.00	0.37 ± 0.10	0.40 ± 0.00
<i>gapA</i>	0.40 ± 0.10	1.70 ± 0.20	6.98 ± 1.17	3.76 ± 0.67	0.10 ± 0.00	1.40 ± 0.20	0.30 ± 0.00
<i>glk</i>	1.00 ± 0.00	2.20 ± 0.10	3.43 ± 0.72	6.04 ± 0.03	0.70 ± 0.10	1.70 ± 0.10	0.80 ± 0.10
<i>gpmA</i>	0.90 ± 0.00	1.80 ± 0.00	3.07 ± 0.21	2.46 ± 0.13	0.60 ± 0.10	1.70 ± 0.30	0.50 ± 0.00
<i>lpdA</i> +	1.10 ± 0.10	2.00 ± 0.10	2.33 ± 0.00	3.16 ± 0.36	1.30 ± 0.30	1.70 ± 0.10	2.00 ± 0.10
<i>pfkA</i>	0.30 ± 0.00	0.50 ± 0.00	1.49 ± 0.11	1.48 ± 0.11	0.00 ± 0.00	0.20 ± 0.00	0.10 ± 0.00
<i>pgi</i>	1.00 ± 0.10	6.60 ± 0.30	3.47 ± 0.10	7.71 ± 0.91	0.40 ± 0.10	3.70 ± 0.10	0.40 ± 0.00
<i>pgk</i>	0.70 ± 0.00	1.20 ± 0.00	1.24 ± 0.20	1.08 ± 0.11	N.D.	N.D.	0.80 ± 0.00
<i>ppc</i>	0.60 ± 0.05	0.80 ± 0.00	3.81 ± 0.27	1.97 ± 0.40	0.80 ± 0.10	0.50 ± 0.10	N.D.
<i>pykA</i>	0.40 ± 0.00	1.00 ± 0.00	1.16 ± 0.12	0.83 ± 0.03	0.10 ± 0.00	0.51 ± 0.10	0.50 ± 0.10
<i>pykF</i>	0.80 ± 0.00	0.90 ± 0.00	0.89 ± 0.31	2.00 ± 0.26	0.20 ± 0.00	0.50 ± 0.10	0.40 ± 0.00
<i>tpiA</i>	0.50 ± 0.00	1.80 ± 0.00	4.84 ± 0.12	8.83 ± 2.61	0.20 ± 0.00	0.30 ± 0.00	0.40 ± 0.00
TCA cycle							
<i>acnA</i>	5.20 ± 0.60	5.70 ± 0.20	2.88 ± 0.31	10.80 ± 1.63	0.70 ± 0.10	3.50 ± 0.30	4.30 ± 0.20
<i>acnB</i> +	1.70 ± 0.36	1.70 ± 0.42	1.76 ± 0.06	24.80 ± 2.18	0.20 ± 0.00	0.70 ± 0.00	2.60 ± 0.30
<i>fumA</i> +	3.30 ± 0.67	3.60 ± 0.67	2.62 ± 0.72	24.07 ± 3.41	2.30 ± 0.50	3.80 ± 0.10	2.10 ± 0.10
<i>gltA</i> +	0.80 ± 0.10	1.30 ± 0.20	1.73 ± 0.31	6.68 ± 0.13	1.30 ± 0.20	1.20 ± 0.20	1.70 ± 0.10
<i>icdA</i> +	1.00 ± 0.10	1.90 ± 0.00	1.03 ± 0.11	4.66 ± 0.78	1.20 ± 0.00	2.40 ± 0.00	1.30 ± 0.00
<i>mdh</i>	1.20 ± 0.10	1.80 ± 0.50	1.93 ± 0.16	3.64 ± 0.32	1.40 ± 0.30	1.40 ± 0.10	1.10 ± 0.10
<i>sdhA</i> +	0.90 ± 0.00	2.00 ± 0.30	4.75 ± 0.98	5.96 ± 0.67	3.70 ± 0.20	4.00 ± 0.00	2.60 ± 0.10
<i>sdhB</i> +	0.80 ± 0.00	1.70 ± 0.00	4.40 ± 0.86	8.05 ± 0.86	2.20 ± 0.00	2.40 ± 0.30	1.30 ± 0.00
<i>sucA</i> +	1.40 ± 0.10	1.70 ± 0.10	3.28 ± 0.20	5.42 ± 1.39	1.20 ± 0.20	1.00 ± 0.00	1.20 ± 0.00
<i>sucB</i> +	0.80 ± 0.00	1.40 ± 0.10	3.31 ± 0.62	8.28 ± 1.91	N.D.	N.D.	1.20 ± 0.00
<i>sucC</i> +	0.90 ± 0.10	1.60 ± 0.00	1.99 ± 0.43	7.41 ± 1.21	1.30 ± 0.10	1.50 ± 0.00	1.20 ± 0.20
<i>sucD</i> +	1.00 ± 0.20	1.40 ± 0.30	1.89 ± 0.32	7.16 ± 0.11	1.70 ± 0.20	1.00 ± 0.10	1.10 ± 0.10
Gluconeogenic, glyoxylate and <i>poxB</i> shunt							
<i>aceA</i> +	12.00 ± 1.20	1.90 ± 0.20	1.66 ± 0.38	24.20 ± 3.16	9.10 ± 2.80	0.80 ± 0.10	8.81 ± 1.00
<i>aceB</i> +	15.60 ± 2.50	3.70 ± 0.30	2.54 ± 0.09	28.65 ± 5.91	11.00 ± 2.10	0.74 ± 0.00	24.20 ± 2.60
<i>acs</i>	5.60 ± 0.74	8.00 ± 0.13	1.83 ± 0.19	21.82 ± 3.20	4.10 ± 0.10	3.40 ± 0.10	11.90 ± 0.10
<i>glcB</i> +	11.70 ± 1.4	3.70 ± 0.50	1.04 ± 0.28	12.62 ± 0.93	2.50 ± 0.20	1.80 ± 0.20	7.60 ± 0.30
<i>maeB</i>	1.20 ± 0.20	2.50 ± 0.03	1.18 ± 0.26	10.06 ± 1.19	1.60 ± 0.00	1.70 ± 0.00	N.D.
<i>pckA</i>	0.80 ± 0.06	2.30 ± 0.02	0.83 ± 0.21	2.81 ± 0.89	0.90 ± 0.20	1.30 ± 0.10	N.D.
<i>poxB</i>	4.20 ± 0.35	5.70 ± 0.48	3.30 ± 0.48	5.30 ± 0.56	1.50 ± 0.10	0.50 ± 0.00	2.80 ± 0.90
<i>ppsA</i>	3.70 ± 0.21	2.40 ± 0.09	7.34 ± 2.08	5.57 ± 1.34	3.30 ± 0.20	1.10 ± 0.20	N.D.
<i>sfcA</i>	1.90 ± 0.33	1.70 ± 0.10	2.56 ± 0.85	3.72 ± 0.28	1.10 ± 0.10	1.00 ± 0.00	N.D.
<i>gal</i> regulon							
<i>galE</i>	38.80 ± 6.80	46.60 ± 6.50	0.53 ± 0.12	40.25 ± 1.97	1.40 ± 0.10	25.00 ± 2.10	N.D.
<i>galK</i>	39.00 ± 2.30	48.21 ± 4.00	2.12 ± 0.09	43.78 ± 3.43	N.D.	N.D.	N.D.
<i>galM</i>	8.10 ± 3.94	3.30 ± 0.70	0.96 ± 0.18	2.67 ± 0.85	N.D.	N.D.	N.D.
<i>galP</i>	12.40 ± 1.08	13.10 ± 1.86	1.98 ± 0.36	25.11 ± 0.98	N.D.	N.D.	11.70 ± 1.90

Table 3 RT-qPCR values of central metabolism and regulatory genes (Continued)

<i>galR</i> ***	3.20 ± 0.10	1.20 ± 0.40	2.04 ± 0.25	5.20 ± 0.10	N.D.	N.D.	N.D.
<i>galS</i>	4.90 ± 0.47	3.20 ± 0.52	4.14 ± 0.35	23.48 ± 0.07	N.D.	N.D.	N.D.
<i>galT</i>	35.60 ± 2.37	42.50 ± 4.00	1.10 ± 0.12	64.34 ± 5.20	N.D.	N.D.	N.D.
<i>galU</i>	1.70 ± 0.30	2.10 ± 0.36	1.27 ± 0.01	2.36 ± 0.22	N.D.	N.D.	N.D.
Gene	Respiration						
<i>cyoA</i> +	0.60 ± 0.01	1.40 ± 0.07	2.16 ± 0.07	4.33 ± 0.38	N.D.	N.D.	1.10 ± 0.10
<i>cyoB</i> +	0.30 ± 0.00	1.40 ± 0.10	2.42 ± 0.58	3.11 ± 0.04	N.D.	N.D.	N.D.
<i>cyoC</i> +	0.30 ± 0.00	1.20 ± 0.20	2.93 ± 0.19	2.77 ± 0.55	N.D.	N.D.	N.D.
<i>cyoD</i> +	0.30 ± 0.02	1.10 ± 0.10	1.90 ± 0.43	3.05 ± 0.03	N.D.	N.D.	N.D.
<i>nuoA</i> +	0.40 ± 0.00	2.40 ± 0.10	3.53 ± 0.54	9.67 ± 1.98	N.D.	N.D.	1.00 ± 0.00
<i>nuoF</i> +	0.40 ± 0.10	1.40 ± 0.10	4.68 ± 0.11	6.84 ± 0.28	N.D.	N.D.	N.D.
<i>nuoN</i> +	0.40 ± 0.02	1.00 ± 0.08	0.44 ± 0.02	1.55 ± 0.18	N.D.	N.D.	N.D.
<i>ubiE</i> +	0.40 ± 0.00	1.60 ± 0.00	4.39 ± 1.31	12.01 ± 1.26	N.D.	N.D.	N.D.
Gene	ppGpp metabolism						
<i>gpp</i>	1.40 ± 1.20	5.20 ± 0.30	9.17 ± 1.65	4.77 ± 0.93	0.00 ± 0.00	0.10 ± 0.00	N.D.
<i>ndk</i>	0.40 ± 0.00	3.10 ± 0.10	3.49 ± 0.52	11.08 ± 0.43	0.50 ± 0.00	3.30 ± 0.10	N.D.
<i>ppa</i>	1.00 ± 0.00	6.10 ± 1.40	7.42 ± 0.33	9.29 ± 0.32	0.50 ± 0.10	3.80 ± 0.10	N.D.
<i>relA</i>	1.30 ± 0.00	1.60 ± 0.20	3.13 ± 0.20	1.51 ± 0.00	0.90 ± 0.10	1.10 ± 0.00	N.D.
<i>spoT</i>	1.10 ± 0.20	4.80 ± 0.20	6.00 ± 0.77	7.36 ± 0.82	0.50 ± 0.10	0.70 ± 0.10	N.D.
Gene	pts operon						
<i>crr</i>	0.00 ± 0.00	0.00 ± 0.00	2.46 ± 0.37	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
<i>ptsH</i>	0.00 ± 0.00	0.00 ± 0.00	1.07 ± 0.12	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
<i>ptsI</i>	0.00 ± 0.00	0.00 ± 0.00	1.34 ± 0.14	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Gene	Genes in deleted chromosomal fragment in PB12						
<i>galR</i> ***	3.20 ± 0.10	1.20 ± 0.40	2.04 ± 0.25	5.20 ± 0.10	N.D.	N.D.	N.D.
<i>lpIT</i> ***	3.95 ± 0.53	0.00 ± 0.00	0.95 ± 0.12	2.83 ± 0.32	N.D.	N.D.	N.D.
<i>muth</i> ***	2.76 ± 0.02	0.00 ± 0.00	16.49 ± 0.22	30.48 ± 0.00	N.D.	N.D.	N.D.
<i>rppH</i> ***	1.84 ± 0.20	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	N.D.	N.D.	N.D.
<i>ptsP</i> ***	1.80 ± 0.20	0.00 ± 0.00	0.55 ± 0.07	0.30 ± 0.06	N.D.	N.D.	N.D.
<i>tas</i> ***	1.31 ± 0.07	0.00 ± 0.00	1.07 ± 0.21	12.80 ± 2.44	N.D.	N.D.	N.D.
<i>ygdQ</i> ***	8.86 ± 1.86	0.00 ± 0.00	1.92 ± 0.06	6.10 ± 1.58	N.D.	N.D.	N.D.
Gene	Genes with point mutations in PB12						
<i>arcB</i> **	2.30 ± 0.70	1.30 ± 0.72	0.99 ± 0.06	1.23 ± 0.19	N.D.	N.D.	N.D.
<i>barA</i> **	4.14 ± 0.91	3.46 ± 0.46	3.25 ± 0.09	2.30 ± 0.53	N.D.	N.D.	N.D.
<i>rna</i> **	0.89 ± 0.18	3.40 ± 0.75	2.06 ± 0.42	8.03 ± 0.00	N.D.	N.D.	N.D.
<i>rssA</i> **	4.89 ± 1.27	4.04 ± 0.43	6.45 ± 0.31	6.07 ± 0.95	N.D.	N.D.	N.D.
<i>rpoD</i> **	1.00 ± 0.20	0.90 ± 0.20	1.00 ± 0.20	1.18 ± 0.24	0.70 ± 0.20	1.10 ± 0.10	N.D.
<i>rpoS</i> +	2.90 ± 0.86	2.20 ± 0.20	2.03 ± 0.07	4.62 ± 0.34	0.10 ± 0.00	0.20 ± 0.00	2.60 ± 0.10
<i>yjJ</i> **	0.86 ± 0.26	8.51 ± 0.00	1.44 ± 0.03	4.31 ± 0.35	N.D.	N.D.	N.D.
<i>ypdA</i> **	1.59 ± 0.49	1.38 ± 0.15	1.20 ± 0.31	1.36 ± 0.11	N.D.	N.D.	N.D.
Gene	Regulators						
<i>arcA</i>	1.02 ± 0.02	1.30 ± 0.09	1.68 ± 0.49	3.86 ± 0.60	0.70 ± 0.00	1.10 ± 0.10	0.00 ± 0.00
<i>arcB</i> **	2.30 ± 0.70	1.30 ± 0.72	0.99 ± 0.06	1.23 ± 0.19	N.D.	N.D.	N.D.
<i>barA</i> **	4.14 ± 0.91	3.46 ± 0.46	3.25 ± 0.09	2.30 ± 0.53	N.D.	N.D.	N.D.
<i>crp</i>	0.80 ± 0.02	1.00 ± 0.05	1.38 ± 0.01	2.22 ± 0.64	N.D.	N.D.	N.D.

Table 3 RT-qPCR values of central metabolism and regulatory genes (Continued)

<i>cydA</i>	1.40 ± 0.10	1.80 ± 0.10	2.37 ± 0.78	3.62 ± 0.68	N.D.	N.D.	N.D.
<i>fadR</i>	0.60 ± 0.03	1.60 ± 0.08	3.52 ± 0.03	3.34 ± 0.19	N.D.	N.D.	N.D.
<i>fruR</i>	1.30 ± 0.00	1.50 ± 0.07	1.41 ± 0.02	4.34 ± 0.34	N.D.	N.D.	N.D.
<i>galR</i> ***	3.20 ± 0.10	1.20 ± 0.40	2.04 ± 0.25	5.20 ± 0.10	N.D.	N.D.	N.D.
<i>iclR</i>	0.40 ± 0.03	1.40 ± 0.07	0.81 ± 0.20	2.62 ± 0.39	N.D.	N.D.	N.D.
<i>ihfA</i>	1.80 ± 0.04	1.70 ± 0.06	1.97 ± 0.32	2.07 ± 0.22	N.D.	N.D.	N.D.
<i>ihfB</i>	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	N.D.	N.D.	N.D.
<i>mlc</i>	1.10 ± 0.00	1.70 ± 0.06	2.37 ± 0.49	10.66 ± 2.07	N.D.	N.D.	N.D.
<i>rpoD</i> **	1.00 ± 0.20	0.90 ± 0.20	1.00 ± 0.20	1.18 ± 0.22	0.70 ± 0.20	1.10 ± 0.10	N.D.
<i>rpoS</i> +	2.90 ± 0.86	2.20 ± 0.20	2.03 ± 0.07	4.62 ± 0.34	0.10 ± 0.00	0.20 ± 0.00	2.60 ± 0.10

Relative mRNA concentrations of central metabolism and regulatory genes of JM101, PB11, PB12 and their derivatives strains, grown in glucose as the sole carbon source, were determined by RT-qPCR. Most of these values have previously been reported (*) and are presented here for discussion and comparison purposes [9-11]. ArcA/B regulated genes are marked with (+). Data in this table are reported as relative expression levels of the parental JM101 strain, assigning it the value of one (see Materials and Methods). The RT-qPCR values of mutated (**) or deleted (***) genes in the PB12 strain have not been reported, with the exception of *arcB* and *rpoS*. N.D. Not determinate.

RT-qPCR values of several of these genes were also increased in the JM101Δ*rppH* as compared to JM101, but in general at lower level, with the exception of the *gapA*, *pgi*, *glk*, *sdhA*, *sdhB*, and *sucAB* genes, which increased more than 3-fold (Table 3). The RT-qPCR values of some of the ArcA/B regulated genes were also highly increased in the PB11Δ*rppH* (Table 3). When comparing the RT-qPCR values of central metabolic genes in the PB11Δ*rppH* strain with those of the PB12 strain, the TCA cycle genes were reduced in the PB12 strain (Table 3). The analysis and possible explanations of these differences is presented in the next section. Remarkably, RT-qPCR values of most of the regulatory genes analyzed were also increased and, in some cases, highly increased in the PB11Δ*rppH* derivative, as compared to the parental PB11, whereas RT-qPCR values of most regulatory genes in the JM101Δ*rppH* derivative

were not substantially modified as compared to its parental JM101 strain.

As mentioned, higher RT-qPCR values of most of the genes in the PB11Δ*rppH* derivative are mainly the result of the increase in the mRNA half-life time due to the absence of RppH. However, since higher levels of most of the transcriptional regulators were detected in the PB11Δ*rppH* strain, it is possible that the enhanced expression of some genes could be, in addition of increased mRNA half-life time, the result of higher or lower expression of their regulatory genes.

Analysis and possible effects of some of the mutated regulatory and putative regulatory genes in the evolved PB12 strain

It has previously been proposed that the point mutation in the *arcB* gene, detected in the PB12 strain, is apparently responsible for diminishing the ArcA/B function as a repressor, since it is known that inactivation of ArcB or enhancing of its ArcA-P dephosphorylating activity, could contribute to the overexpression of ArcA/B regulated genes [9,20-22]. However, in order to explain the particularly higher mRNA levels of most of the TCA cycle and respiratory genes mainly controlled by ArcA/B in the PB11Δ*rppH* strain, as compared to the PB12 strain (except for *mdh* and *nuoN*; Table 3), we now propose a different role for this *arcB* mutation. This mutation is apparently responsible for modifying the ArcA/B repressor function in the PB12 strain by reducing -not enhancing-, its ArcA-P dephosphorylating capacity, which in turn could contribute to higher repression of ArcA/B regulated genes, explaining the reduction in the RT-qPCR values of most of the genes regulated by the ArcA-P in PB12 as compared to the PB11Δ*rppH* strain. Therefore, this change in the *arcB* gene apparently

Table 4 Specific growth rates (μ) of the PB12 strain and its derivatives

Regulatory genes		
Strain	μ (h ⁻¹)	%
PB12	0.44 ± 0.016	100
PB12Δ <i>arcA</i>	0.4 ± 0.0038	91
PB12Δ <i>barA</i>	0.46 ± 0.001	105
PB12Δ <i>rpoS</i> *	0.38	86
Possible regulatory genes		
Strain	μ (h ⁻¹)	%
PB12Δ <i>yjiU</i>	0.34 ± 0.003	77
PB12Δ <i>rssA</i>	0.43 ± 0.018	98
PB12Δ <i>rna</i>	0.43 ± 0.035	98
PB12Δ <i>ypdA</i>	0.44 ± 0.003	100

*This value has been previously reported [9-11].

reduced both, transcription of ArcA/B-dependent genes [9,20,21], and metabolic burden, allowing better growth capacities to the PB12 strain as compared to the PB11 Δ *rppH* derivative (Table 3). In agreement with this proposition, the knockout inactivation of the *arcA* gene in the PB12 Δ *arcA* strain reduced 10% the μ (Table 4), because higher transcription levels of the ArcA/B-controlled genes resulted in this derivative (data not shown) and this was probably sensed as metabolic burden. The same growth diminishing effect occurred in the JM101 Δ *rppH* strain, probably due to higher transcription levels of many central metabolism genes, including some of the TCA cycle, which were apparently responsible for reducing 27% the μ (data not shown), as compared to the parental JM101 strain. In agreement with the important role of the ArcA/B regulator, inactivation of *arcA* in the PB11 Δ *arcA* strain increase substantially the μ and the transcription levels of most of the ArcA/B-regulated genes as compared to PB11 strain [10] (Table 3). From these results, it is tempting to propose that inactivation of the *arcA* gene in *E. coli* could be used as a tool for allowing better growth capabilities to cells growing aerobically in certain stress conditions, in which the lack of regulation of the TCA cycle and respiratory genes would be an advantage [9,10].

It has been proposed that YjjU could be involved in regulatory processes [32,33]. The inactivation of *yjjU* in the PB11 Δ *yjjU* strain increased its μ from 0.13 to 0.16 h⁻¹. This 23% increment is not as high as the values obtained with the inactivation of *arcA* (243%) and *rppH* (261%) (Table 2). However, *yjjU* inactivation in the PB12 Δ *yjjU* strain reduced its μ 23% (from 0.44 to 0.34 h⁻¹), as compared to the parental PB12 strain (Table 4). These results suggest that if this protein really functions as a regulatory factor, as has been proposed, the point mutation could allow stronger capabilities to the cell for faster growth in glucose. Cassette inactivation of *yjjU* is the only case in which a gene knockout increased the μ in the PB11 Δ *yjjU* derivative, and reduced the μ in the same percentage in the PB12 Δ *yjjU* derivative. This mutation has to be investigated further, initially analyzing the transcription pattern of critical genes in the strain PB12 Δ *yjjU* as compared to the parental PB12.

The mutation in the *rpoD* gene is responsible of a conserved valine 482 to isoleucine substitution located in the HTH motif of region 4.2 of RpoD that is involved in the recognition of the -35 promoter region. In the co-crystal structure of region 4.2 of *Thermus aquaticus* with promoter DNA, which is almost identical to the *E. coli*, this position is located at the turn of the HTH motif and does not make any direct contact with the DNA [36]. Thus, it is likely that this particular substitution does

not affect the affinity of this sigma subunit for the promoter DNA sequences.

Since the knockout inactivation of the *barA*, *rssA*, *rna* and *ypdA* genes did not modify substantially the μ in the PB11 and PB12 derivatives, it appears that these genes played minor or not role at all in the growth recovery observed in the evolved strain.

Conclusions

We propose that the deletion event that simultaneously removed the *mutH*, *rppH*, and part of the *galR* genes, mainly responsible for the faster growth (4x) in glucose, occurred as one of the initial events in the adaptive laboratory evolution process which resulted in the evolved PB12 strain. This deletion caused simultaneously: a) a very high mutagenesis rate due to the removal of *mutH*, in a strain lacking PTS that is already responsible of a carbon stress response, b) higher glucose transport, by increased levels of GalP in this strain lacking PTS, due to the inactivation of *galR* [9,12], and c) higher mRNA levels resulting in enhanced glycolytic and TCA fluxes and better respiratory capacity to the precursor of the PB12 strain due to the absence of RppH.

In addition, lower mRNA levels of most of the ArcA/B regulated genes were detected in the PB12 strain as compared to the PB11 Δ *rppH* derivative. This can be explained as an enhanced ArcA-P repressor capacity due to the *arcB* mutation that apparently appeared after the deletion of the *rppH* gene in the evolved strain, allowing lower levels of transcription of ArcA/B-regulated genes.

Knockout inactivation of the *barA*, *rssA*, *rna* and *ypdA* genes in the PB11 and in PB12 strains did not modify substantially the μ of the derivatives, suggesting that each of these mutations alone apparently played minor or no roles at all in the growth recovery in the evolved strain. Some of these changes could in fact be neutral mutations [48].

From these considerations, the evidences indicate that the main reasons for fast growth on glucose are apparently the deletion of the *rppH*, *galR*, and *mutH* genes and, perhaps, the point mutation in the *arcB* gene. These two changes could have been fixed in a short period of time during the fermentation process. Nevertheless, it cannot be ruled out that other point mutation, as those in the *yjjU*, or in the *barA* genes that have not been completely characterized in this study, could also play a minor role in the growth recovery in glucose.

In this study, as in others [4,5], we used two different whole genome sequencing strategies which produced slightly different results. True changes had to be discerned from false positives by conventional Sanger sequencing. Therefore, it is important to emphasize the relevance of using more than one genome resequencing

method for this type of studies to have high confidence in the results.

Finally, the results presented here show the physiological plasticity of *E. coli* and could be useful in the design of more robust adaptive laboratory evolution strategies.

Methods

Bacterial strains, growth conditions and recombinant DNA techniques

E. coli strains JM101 [F' *traD36 proAB⁺ lacI^q lacZΔM15/supE thi Δ(lac-proAB) rpoS(33 am)*], PB11 [JM101Δ(*ptsH, ptsI, crr*):: *kan*] and PB12 (PB11, PTS⁻ Glc⁺) and derivatives have previously been described [7,9-11,38,39]. The derivatives of these strains utilized in this report, in which the *barA, yjjU, rssA, rna, ypdA*, and *rppH* genes were knockout inactivated, were obtained by the Datsenko and Wanner method [46], using the oligonucleotides listed in table S3 presented in Additional file 4. All gene disruptions were confirmed by PCR (data not shown). For *inoculum*s preparations, strains stored at -72°C in glycerol were inoculated into Luria broth (LB) for overnight growth.

The culture of the PB12 strain that was also utilized for preparing the DNA for genome sequencing, was obtained from the original culture that has been kept frozen in glycerol (Figure 1) [10]. For μ determinations, cells were grown in LB and then inoculated into M9 minimal medium with 2 g/l of glucose as the only carbon source; when the cultures were growing exponentially, they were inoculated into the same prewarmed (50 ml) medium at 37°C and stirred at 300 rpm with a starting optical density at 600 nm ($O.D._{600nm}$) = 0.1. $O.D._{600nm}$ were measured using a Klett/Summerson photocolormeter, model 800-3. All specific growth rate values presented in Tables 2 and 4 are the averages of at least two independent cultures, each one in duplicate. For RNA isolation and RT-qPCR analyses, duplicate cultures were grown on 1 L fermentors on M9 medium with 2 g/l of glucose as the sole carbon source, at 37°C, stirred at 600 rpm and air flow rate at 1 vvm, with a starting $O.D._{600nm}$ = 0.1. For RT-qPCR determinations cells of the different fermentations were collected in the log phase at $O.D._{600nm}$ = 1 [9].

DNA extraction from parental and evolved PB12 strains for genomic analysis

Two overnight cultures of the *E. coli* strains JM101, PB11 and PB12 were grown from their frozen original stocks in liquid LB medium. One set of these cultures (not including PB11) of these strains was utilized for DNA purification submitted to RN, and the DNA of the other set was submitted to the UNAM Massive Sequencing Unit, for genome resequencing (see below). DNA

was extracted by a maxiprep phenol extraction and ethanol precipitation method [49] and purified with the Pure Link PCR purification kit (Invitrogen, USA). Quality and quantity of extracted DNA was verified as recommended by RN and by UNAM Massive DNA Sequencing Unit.

Roche NimbleGen Inc. sequencing

DNA samples from the JM101 and PB12 strains were submitted to RN for CGS analysis using *E. coli* K-12 MG1655 (ATC #47076) as the reference strain [40]. The results provided by RN are included in Table 1, Figure 2 and in table S1 presented in Additional file 1.

Paired and paired end (PE) library construction and GAIIX sequencing

DNA samples from the JM101, PB11 and PB12 strains were submitted to the Massive DNA Sequencing Unit of UNAM for its paired ended (PE) library construction and genome sequencing. PE library was constructed following Illumina Inc. recommendations. Briefly, 5 μ g of chromosomal DNA of each strain was fragmented by nitrogen nebulization during 6 min at a pressure of 32 psi. Fragmented DNA was purified using the QIAquick PCR purification kit and resuspended in 30 μ l of elution buffer (EB: 10 mM Tris-HCl, pH 8.5). DNA end repairs were performed using a mixture of T4 and Klenow DNA polymerases and T4 polynucleotide kinase for 5' ends. In order to facilitate the ligation of double stranded adapters, an adenine residue was incorporated at each 3' end of fragmented DNA before this step using a Klenow exo minus (exo⁻) enzyme and dATP. Illumina Inc. adapters with overhang thymine residues at 3' ends were ligated at each end of fragmented DNA using 2x rapid ligation buffer (Illumina Inc.) and T4 DNA ligase during 15 min at room temperature. Ligated DNA was purified using a Qiagen MinElute purification kit (Qiagen, USA) and resuspended in 15 μ l of EB. Modified DNA pool was loaded on a 2% gel of Ultra Low Range Agarose (Bio Rad Laboratories USA) and \approx 500 bp DNA fragments were purified using a QIAquick gel extraction kit. To enrich the adapter-modified DNA fragments, purified DNA was used as template for a 12-cycle PCR reaction (98, 65 and 72°C), using PCR primers PE 1.0 and 2.0 and Phusion DNA polymerase (included in the Illumina Inc. PE sample prep kit). PCR products were purified using a QIAquick PCR purification kit (Qiagen, USA) and eluted in 50 μ l of EB. Validation and quantification of the libraries were performed using an Agilent Bioanalyzer 2100 (DNA 1000 chip). Finally, 18 pM of DNA library were used for a PE sequence of 2x36 cycles on a GAIIX instrument that performs sequencing by a synthesis method based on reversible fluorescent terminators accordingly to Illumina, Inc.

Genome “de novo” assembly and variant identification by Winter Genomics Inc

Low quality reads produced by the Illumina GAIIX method were filtered using the ShortRead 1.8.0 package [50]. Assembly for each strain was performed with the PE-Assembler 1.1 [51]. IMAGE 2.1 [52] was used to close gaps as it locally assembles reads aligning to contig ends. Bowtie 0.12.5 short read aligner [53] was used to align reads to the resulting contigs and unsupported bases were removed with the Biostrings 2.18.0 package. Contigs were re-ordered along the *E. coli* K12 MG1655 genome [40] by using the Mauve 2.3.1. software [54,55]. Then contigs were compared against the reference genome using both Mauve and Murasaki 1.68.6. softwares [56]. Using the PTS operon deletion as marker, it was possible to correctly identify each strain. BLAT v34 software [57] was used to perform alignments of strain PB12, against JM101. VarScan 1.2 software [58] was used to identify variants using the BLAT alignments as input. Ambiguous variants were filtered out using a custom Perl script. For most of the analyses local cluster resources of the Instituto de Biotecnología-UNAM were used. The results provided by WG for JM101 and PB12 strains are included in Table 1 and in table S2 presented in Additional file 2. None of the point mutations detected in PB12 appeared in the genomic sequencing of the PB11 strain (data not shown).

DNA sequencing of putative mutations by Sanger methodology

DNA regions containing putative mutations in regulatory genes detected by RN and WG were PCR amplified using oligonucleotide primers listed in table S3 presented in Additional file 4, purified by the Pure Link PCR purification kit and sequenced by the Sanger methodology with the Taq FS Dye Terminator Cycle Sequencing Fluorescence-Based Sequencing, in a Perkin Elmer/Applied Biosystems Model 3730. Sequence differences of 14 of the mutations presented in Table 1A were confirmed by examination of the trace data (data not shown).

RNA Extraction, DNase treatment of RNA and cDNA synthesis for RT-qPCR analysis

Total RNA from the utilized strains was isolated and purified using the hot-phenol method, with some modifications. Samples containing 50 ml of the different strains growing logarithmically in the fermentor were collected at 1 OD_{600nm}. 1 ml of RNA later buffer (Ambion Inc., USA) was added to each sample, mixed and centrifuged 10 min/4°C/5000 rpm. Cells were resuspended with 1 ml of buffer I (0.3 M sucrose, 0.1 M sodium acetate), treated with 20 µl of lysozyme (10 mg/ml in TE buffer) for 10 min at room temperature. 2 ml of buffer II (0.01 M sodium acetate, 2% SDS) were added and the

mixtures incubated for 3 min at 65°C. The lysates were extracted with 2 ml of hot phenol and heated for 3 min at 65°C. A second extraction with hot phenol was performed without heating the mixtures. Samples were then extracted with 2 ml of a phenol:chloroform mixture (1:1), precipitated with 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.5 volume of ethanol and centrifuged for 15 min at 4°C/10000 rpm. Samples were then suspended in 300 µl of DNase and RNase-free water (Ambion Inc, USA) with RNase inhibitor (Fermentas Life Sciences, USA) and extracted twice with 1 volume of chloroform. Finally, samples were precipitated as before and suspended in 300 µl of TE buffer (Ambion Inc, USA). RNA was analyzed on formaldehyde agarose gel for integrity. RNA concentrations were quantified using Nanodrop 2000c (Thermo Scientific); the 260_{nm}/280_{nm} and 260_{nm}/230_{nm} ratios were examined for protein and solvent contamination. For all samples the 260_{nm}/280_{nm} absorbances values were between 1.9-2.0 and in the range of 2.0-2.3 for the 260_{nm}/230_{nm} ratio. RNA samples were stored at -70°C. Three RNA extractions and purifications were carried out from three independent fermentations for each strain.

For DNase treatment, total RNA samples were treated with TURBO DNA-free kit (Ambion Inc, USA) at 37°C for 30 min, following manufacturer's instructions. To determine whether RNA samples were significantly contaminated with genomic DNA, samples were subjected to conventional PCR with primers for the *arcA* gene (Table S3 presented in Additional file 4). Since these primers were designed to recognize genomic DNA, the presence of a detectable PCR product on an ethidium bromide-stained agarose gel confirmed that the specific RNA sample was contaminated with genomic DNA. Contaminated samples were discarded. PCR reactions were performed with Taq polymerase (Fermentas Life Sciences, USA). The cycling parameters were: 95°C for 5 min, 30 cycles at 95°C for 1 min, 55°C for 1 min and 72°C for 1 min plus an extension step at 72°C for 5 min. Additionally, the DNase-treated RNA samples were used for RT-qPCR analyses of the same *arcA* gene, using the appropriate oligos [*arcAa* (forward) and *arcAb* (reverse)] (Table S3 presented in Additional file 4). As in the PCR case, all utilized samples did not produce a 101 bp amplicon, indicating that small fragments of genomic DNA were not present. cDNA was synthesized using RevertAidTM H minus First Strand cDNA Synthesis kit following the manufacturer's instructions (Fermentas LifeSciences, USA.). For each reaction approximately 5 µg of RNA and a mixture of 10 pmol/µl of specific DNA reverse primers (b) for the utilized genes, were used. Nucleotide sequences of these genes have been previously published [9-11] or are listed in table S3 presented in Additional file 4. cDNA were used as

templates for RT-qPCR assays. cDNAs were synthesized using specific oligonucleotides, since this condition ensures the synthesis of only one copy of cDNA per each RNA molecule [9,59].

RT-qPCR

RT-qPCR was performed with the ABI Prism 7000 Sequence Detection System and 7300 Real Time PCR System (Perkin Elmer/Applied Biosystems, USA) using the Maxima^R SYBR Green/ROX qPCR Master Mix (2X) kit (Fermentas LifeSciences, USA). MicroAmp Optical 96-well reaction plates (Applied Biosystems, USA) and Plate Max ultraclear sealing films (Axygen Inc, USA) were used in these experiments. Amplification conditions were 10 min at 95°C, followed by a two-step cycle at 95°C for 15 sec and 60°C for 60 sec for a total of 40 cycles, to finish with a dissociation protocol (95°C for 15 sec, 60°C for 1 min, 95°C for 15 sec and 60°C for 15 sec). DNA sequences of the primers for specific amplifications were designed using the Primer Express software (Perkin Elmer/Applied Biosystems, USA). Some of these sequences have been previously published [9-11] and the rest are included in table S3 presented in Additional file 4. All RT-qPCR experiments complied with the MIQE guidelines (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) [60,61]. The length of all the utilized oligonucleotides (forward and reverse), was between 18 and 21 nucleotides, with % of GC between 45 to 60 and T_m between 58 to 60°C. The size of all amplimers was 101 bp. The final primer concentration was 0.2 μM, in a total volume of 12 μl. Five ng of target cDNA for each gene were added to the reaction mixture, since higher cDNA concentrations (>10 ng) are not in the dynamic range of the reference *ihfB* gene (see below). Hence the obtained values cannot be correctly normalized for this higher cDNA concentration. All experiments were performed at least in triplicate from three different fermentations, for each gene of each strain, obtaining very similar values (differences <0.3 SD). A non-template control reaction mixture was included for each gene and values appeared for all genes after cycle 31. Standard curves were constructed to evaluate PCR efficiency and all the genes had R² values above 0.9976 with slopes between -3.4 to -3.7. The quantification technique used to analyze data was the 2^{-ΔΔC_q} method described by Livak and Shmittgen [62]. Data were normalized using the *ihfB* gene as an internal control (reference gene). The same reproducible expression level of this gene was detected in all the strains in the conditions in which bacteria were grown and analyzed, since this is the most important characteristic that a reference gene should have in accordance with

the MIQE guidelines. Additional file 5 (Figure S2) presents the *ihfB* gene values detected for the utilized strains. These results demonstrate the stability of the expression of this reference gene in all the analyzed derivatives for the used conditions in this report and also on previous reports utilizing these strains and other derivatives [9-11,60].

For each analyzed gene in all strains the transcription level of the corresponding JM101 gene, was considered equal to one, and it was used as control to normalize the data. Therefore, data are reported as relative expression levels, compared to the expression level of the same gene in the JM101 strain. Results presented in Table 3 are the averages of at least three independent measurements of the RT-qPCR expression values for each gene. Values were obtained from different cDNAs generated from two independent bioreactor samples [9].

Additional files

Additional file 1: Table S1. Mutations in coding regions detected by Roche NimbleGen Inc. This table includes the data provided by Roche NimbleGen Inc. (RN) for the whole genome sequence analysis of the evolved PB12 strain. Section A lists 26 (21+5) non-synonymous point mutations in structural genes that accordingly to RN changed the coding regions of these genes. In fact only in 21 of these genes detected also by Winter Genomics Inc (WG), the mutations occurred (Table 1A). Section B presents the list of the 12 genes included in a large deletion in this strain. This list also includes the genes deleted in the parental PB11 strain, for the construction of this derivative lacking PTS (Figure 2). The table also includes (Section C), the list of the 20 genes in which synonymous point mutations occurred accordingly to RN. Those 16 in common with WG are in bold letters.

Additional file 2: Table S2. Mutations in coding regions detected by Winter Genomics Inc. This table lists the data provided by Winter Genomics Inc. (WG) obtained for the whole genome sequence of the PB12 strain in comparison to the parental strains JM101 and PB11. Importantly, PB11 strain was sequenced by WG and the same nucleotide sequences as in the parental strain JM101 were determined (data not shown). Therefore all the point mutations detected in PB12 by RN and WG appeared during the laboratory evolution process. Section A includes a list of 27 genes (21+6) in which, accordingly to this company non-synonymous point mutations occurred changing the coding regions in structural genes. 21 of these genes were also detected by RN (Table 1A). The table also includes (Section B) the list of 18 genes in which synonymous mutations also occurred, accordingly to WG. Those 16 in common with RN are in bold letters.

Additional file 3: Figure S1. Nucleotide sequence of the chromosomal genes fusion that occurred in the evolved PB12 strain. This figure includes the nucleotide sequence of the genomic region where the deletion occurred (Figures 3 and 4) between the *ptsP* and the *galR* genes in the PB12 strain.

Additional file 4: Table S3. Oligonucleotides employed in this study. This table lists the oligonucleotides utilized in this work. Section A shows the oligonucleotides used for DNA sequencing with the Sanger method, including those for the confirmation of the deletion that occurred in the PB12 strain (Figures 2, 3 and 4), the reported mutations provided by RN and WG (Table 1 and Tables S1 and S2 in Additional file 1 and Additional file 2) and the ones employed for gene disruption confirmation. Section B lists the oligos utilized for gene disruption with the Datsenko-Wanner methodology [46]. Section C lists the oligonucleotides utilized for RT-qPCR analysis not previously reported. The sequences of the oligos utilized for the remaining genes listed in Table 3, have been previously published [9-11] (see Materials and Methods).

Additional file 5: Figure S2. Amplification curves for the *ihfB* gene in different strains. This figure shows the positions of the amplification curves for the *ihfB* gene (Section A) and the Ct values of this gene (Section B), in the different strains employed in this study. As can be seen, all the amplification curves of the *ihfB* gene that has been used as the reference gene for the determination of the RT-qPCR levels, have very similar values. The values presented in the table are from three different fermentations (F1, F2 y F3) for each utilized strain. Since all the values included in section B are very similar, only one third of them are presented (labeled with an asterisk *) in section A. These results demonstrate that the same reproducible expression levels were obtained for the *ihfB* gene in all strains. This is the most important characteristic that a reference gene should have, in agreement to the MIQE guidelines [59-61]. These results corroborate the stability of the expression of the reference *ihfB* gene in these strains in the utilized conditions.

Competing interests

The authors have declared that no competing interests exist.

Authors' contributions

CA carried out all the molecular, genetic and bacterial growth experiments. CA and RA participated in the fermentor cultures of the strains. CA and NF performed the RT-qPCR analysis. FR-M and AE participated in the genome assembly and variant identification of the strains. CA, AE, EM and FB carried out the data analysis. CA, AE, GG, EM and FB conceived the study and designed the experiments. CA, EM and FB wrote the paper. All authors read and approved the final manuscript.

Funding

Consejo Nacional de Ciencia y Tecnología (CONACyT/México grants 105782, FONSEC/SSA/ISSSTE/CONACyT 44126, 126793 and INOVAPYMMME 137117,155519; Dirección General de Asuntos del Personal Académico-Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica-Universidad Nacional Autónoma de México (DGAPA-PAPIIT-UNAM) grants IN213508, IN224709, IN205811 and IN221106. The founders had no role in study design, data collection and analyses, decision to publish, or preparation of the manuscript.

Acknowledgements

We thank Paul Gaytán, Jorge Yáñez and Eugenio López for the synthesis of oligonucleotides and Sanger DNA sequencing; Ricardo Grande and Verónica Jiménez-Jacinto for the complete genome sequence performed at Unidad Universitaria de Secuenciación Masiva de ADN, UNAM. Leonardo Collado and Salvador A. Romero-Martínez from Winter Genomics Inc., for their assistance during the assembly and variant calling steps. We also thank Georgina Hernández and Mercedes Enzaldo for technical support and Jerome Verleyen for his help with the computer Cluster.

Author details

¹Departamento de Ingeniería Celular y Biotecnología, Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Cuernavaca, Morelos 62210, México. ²Winter Genomics, México D.F. 07300, Cuernavaca, Morelos 62210, México.

Received: 10 April 2012 Accepted: 2 August 2012

Published: 10 August 2012

References

1. Albert TJ, Dailidene D, Dailide G, Norton JE, Kalia A, Richmond TA, Molla M, Singh J, Green RD, Berg DE: **Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*.** *Nat Methods* 2005, **2**:951-953.
2. Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK: **Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale.** *Nat Genet* 2006, **38**:1406-1412.
3. Herring CD, Palsson BØ: **An evaluation of Comparative Genome Sequencing (CGS) by comparing two previously-sequenced bacterial genomes.** *BMC Genomics* 2007, **8**:274.
4. Conrad TM, Joyce AR, Applebee MK, Barrett CL, Xie B, Gao Y, Palsson BØ: **Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations.** *Genome Biol* 2009, **10**:R118.
5. Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y, Palsson BØ: **Genetic basis of growth adaptation of *Escherichia coli* after deletion of *pgi*, a major metabolic gene.** *PLoS Genet* 2010, **6**:1001186.
6. Postma PW, Lengeler JW, Jacobson GR: **Phosphoenolpyruvate: carbohydrate phosphotransferase systems.** In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. 2nd edition. Edited by Neidhardt FC. Washington DC, USA: ASM; 1996:1149-1174.
7. Flores N, Xiao J, Berry A, Bolívar F, Valle F: **Pathway engineering for the production of aromatic compounds in *Escherichia coli*.** *Nat Biotechnol* 1996, **14**:620-623.
8. Flores N, de Anda R, Flores S, Escalante A, Hernández G, Martínez A, Ramírez OT, Gosset G, Bolívar F: **Role of pyruvate oxidase in *Escherichia coli* strains lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *J Mol Microbiol Biotechnol* 2004, **8**:209-221.
9. Flores N, Flores S, Escalante A, de Anda R, Leal L, Malpica R, Georgellis D, Gosset G, Bolívar F: **Adaptation for fast growth on glucose by differential expression of central carbon metabolism and *gal* regulon genes in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *Metabol Eng* 2005, **7**:70-87.
10. Flores N, Leal L, Sigala JC, de Anda R, Escalante A, Martínez A, Ramírez OT, Gosset G, Bolívar F: **Growth recovery on glucose under aerobic conditions of an *Escherichia coli* strain carrying a phosphoenolpyruvate: carbohydrate phosphotransferase system deletion by inactivating *arcA* and overexpressing the genes coding for glucokinase and galactose permease.** *J Mol Microbiol Biotechnol* 2007, **13**:105-116.
11. Flores N, Escalante A, de Anda R, Báez-Viveros JL, Merino E, Franco B, Georgellis D, Gosset G, Bolívar F: **New insights into the role of the sigma factor RpoS as revealed in *Escherichia coli* strains lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *J Mol Microbiol Biotechnol* 2008, **14**:176-192.
12. Flores S, Gosset G, Flores N, de Graaf A, Bolívar F: **Analysis of carbon metabolism in *Escherichia coli* strains with an inactive phosphotransferase system by ¹³C labelling and NMR spectroscopy.** *Metabol Eng* 2002, **4**:124-137.
13. Flores S, Flores N, de Anda R, González A, Escalante A, Sigala JC, Gosset G, Bolívar F: **Nutrient-scavenging stress response in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system, as explored by gene expression profile analysis.** *J Mol Microbiol Biotechnol* 2005, **10**:51-63.
14. Báez JL, Bolívar F, Gosset G: **Determinations of 3-deoxy-D-arabino-heptulonate 7- phosphate productivity and yield from glucose in *Escherichia coli* devoided of the glucose phosphotransferase system.** *Biotechnol Bioeng* 2001, **73**:530-535.
15. Báez-Viveros JL, Osuna J, Hernández-Chávez G, Soberon X, Bolívar F, Gosset G: **Metabolic engineering and protein directed evolution increase the yield of L-phenylalanine synthesized from glucose in *Escherichia coli*.** *Biotechnol Bioeng* 2004, **87**:516-524.
16. Martínez K, de Anda R, Hernández G, Escalante A, Gosset G, Ramírez OT, Bolívar FG: **Couitilization of glucose and glycerol enhances the production of aromatic compounds in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *Microb Cell Fact* 2008, **7**:1.
17. Escalante A, Calderon R, Valdivia A, de Anda R, Hernandez G, Ramírez OT, Gosset G, Bolívar F: **Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *Microb Cell Fact* 2010, **9**:21.
18. Raghunathan A, Palsson BØ: **Scalable method to determine mutations that occur during adaptive evolution of *Escherichia coli*.** *Biotechnol Lett* 2003, **25**:435-441.
19. Olvera L, Mendoza-Vargas A, Flores N, Olvera M, Sigala JC, Gosset G, Morett E, Bolívar F: **Transcription analysis of central metabolism genes in *Escherichia coli*. Possible roles of sigma38 in their expression, as a response to carbon limitation.** *PLoS One* 2009, **4**:e7466.
20. Malpica R, Franco B, Rodríguez C, Kwon O, Georgellis D: **Identification of a quinone sensitive redox switch in the ArcB sensor kinase.** *Proc Natl Acad Sci USA* 2004, **101**:13318-13323.

21. Iuchi S, Lin ECC: **Mutational analysis of signal transduction by ArcB, a membrane sensor protein responsible for anaerobic repression of operons involved in the central aerobic pathways in *Escherichia coli*.** *J Bacteriol* 1992, **174**:3972–3980.
22. Liu X, DeWulf P: **Probing the ArcP modulon of *Escherichia coli* by whole transcriptional analysis and sequence recognition profiling.** *J Biol Chem* 2004, **279**:12588–12597.
23. Geanakopoulos M, Adhya S: **Functional characterization of the roles of GalR and GalS as regulators of the *gal* regulon.** *J Bacteriol* 1997, **179**:228–234.
24. Deana A, Celesnik H, Belasco JG: **The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal.** *Nature* 2008, **451**:355–358.
25. Schofield MJ, Hsieh P: **DNA mismatch repair: molecular mechanisms and biological function.** *Annu Rev Microbiol* 2003, **57**:579–608.
26. Kaplan R, Apirion D: **Decay of ribosomal ribonucleic acid in *Escherichia coli* cells starved for various nutrients.** *J Biol Chem* 1975, **250**:3174–3178.
27. Cohen L, Kaplan R: **Accumulation of nucleotides by starved *Escherichia coli* cells as a probe for the involvement of ribonucleases in ribonucleic acid degradation.** *J Bacteriol* 1977, **129**:651–657.
28. Pernestig AK, Meleforts O, Georgellis D: **Identification of UvrY as the cognate response regulator for the BarA sensor kinase in *Escherichia coli*.** *J Biol Chem* 2001, **276**:225–231.
29. Pernestig AK, Georgellis D, Romeo T, Suzuki K, Tomenius H, Normark S, Meleforts O: **The *Escherichia coli* BarA-UvrY two-component system is needed for efficient switching between glycolytic and gluconeogenic carbon sources.** *J Bacteriol* 2003, **185**:843–853.
30. Lengeler JW, Drews G, Schlegel HG: *Biology of the Prokaryotes*. Blackwell Science: New York; 1999:366–367.
31. Ruiz N, Peterson CN, Silhavy TJ: **RpoS-dependent transcriptional control of *sprE*: regulatory feedback loop.** *J Bacteriol* 2001, **183**:5974–5981.
32. Serres MH, Gopal S, Nahum LA, Liang P, Gaasterland T, Riley M: **A functional update of the *Escherichia coli* K-12 genome.** *Genome Biol* 2001, **2**(9):0035.
33. Banerji S, Flieger A: **Patatin-like proteins: a new family of lipolytic enzymes present in bacteria?** *Microbiology* 2004, **150**:522–525.
34. Arifuzzaman M, Maeda M, Itoh A, Nishikata K, Takita C, Saito R, Ara T, Nakahigashi K, Huang HC, Hirai A, Tsuzuki K, Nakamura S, Altaf-Ul-Amin M, Oshima T, Baba T, Yamamoto N, Kawamura T, Ioka-Nakamichi T, Kitagawa M, Tomita M, Kanaya S, Wada C, Mori H: **Large-scale identification of protein-protein interaction of *Escherichia coli* K-12.** *Genome Res* 2006, **16**:686–691.
35. Junop MS, Yang W, Funchain P, Clendenin W, Miller JH: ***In vitro* and *in vivo* studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination.** *DNA Repair* 2003, **2**(4):387–405.
36. Campbell EA, Muzzin O, Chlenov M, Sun JL, Olson CA, Weinman O, Trester-Zedlitz ML, Darst SA: **Structure of the bacterial RNA polymerase promoter specificity σ subunit.** *Mol Cell* 2002, **9**:527–539.
37. Rodriguez SM, Panjikar S, Van Belle K, Wyns L, Messens J, Loris R: **Nonspecific base recognition mediated by water bridges and hydrophobic stacking in ribonuclease I from *Escherichia coli*.** *Protein Sci* 2008, **17**:681–690.
38. Messing J: **A multipurpose cloning system based on the single stranded DNA bacteriophage M13.** *Recombinant DNA technical bulletin* 1979, **2**:43–48.
39. Bachmann BJ: **Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12.** In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. 2nd edition. Edited by Neidhart FC. Washington DC, USA: ASM; 1996:2460–2488.
40. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y: **The complete genome sequence of *Escherichia coli* K-12.** *Science*. 1997, **277**(5331):1453–1462.
41. Tenaillon O, Denamur E, Matic I: **Evolutionary significance of stress-induced mutagenesis in bacteria.** *Trends Microbiol*. 2004, **12**(6):264–70.
42. Bjedov I, Tenaillon O, Gerard B, Souza V, Denamur E, Radman M, Taddei F, Matic I: **Stress-induced mutagenesis in bacteria.** *Science* 2003, **300**:1404–1409.
43. Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, Calteau A, Chiappello H, Clermont O, Cruveiller S, Danchin A, Diard M, Dossat C, Karoui ME, Frapy E, Garry L, Ghigo JM, Gilles AM, Johnson J, Le Bouguéne C, Lescat M, Mangenot S, Martinez-Jéhanne V, Matic I, Nassif X, Oztas S, Petit MA, Pichon C, Rouy Z, Ruf CS, Schneider D, Tourret J, Vacherie B, Vallenet D, Médigue C, Rocha EP, Denamur E: **Organized genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths.** *PLoS Genet* 2009, **5**:e1000344.
44. Hua Q, Yand C, Oshima T, Mori H, Shimizu K: **Analysis of gene expression in *Escherichia coli* in response to changes of growth-limiting nutrient in chemostat cultures.** *Appl Environ Microbiol* 2004, **70**:2354–2366.
45. Silander OK, Tenaillon O, Chao L: **Understanding the evolutionary fate of finite populations: the dynamics of mutational effects.** *PLoS Biology* 2007, **5**:e94.
46. Datsenko KA, Wanner BL: **One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.** *Proc Natl Acad Sci USA* 2000, **97**:6640–6645.
47. Gerdes SY, Scholle MD, Campbell JW, Balázs G, Ravasz E, Daugherty MD, Somera AL, Kyrpides NC, Anderson I, Gelfand MS, Bhattacharya A, Kapatral V, D'Souza M, Baev MV, Grechkin Y, Mseeh F, Fonstein MY, Overbeek R, Barabási AL, Oltvai ZN, Osterman AL: **Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655.** *J Bacteriol* 2003, **185**:5673–84.
48. Ochman H: **Neutral mutations and neutral substitutions in bacterial genomes.** *Mol Biol Evol* 2003, **12**:2091–2096.
49. Ausubel FA, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: *Current Protocols in Molecular Biology*. New York: John Wiley & Sons; 1999.
50. Morgan M, Anders S, Lawrence M, Abyouan P, Pagès H, Gentleman R: **ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data.** *Bioinformatics* 2009, **25**:2607–8.
51. Ariyaratne PN, Sung WK: **PE-Assembler: *de novo* assembler using short paired-end reads.** *Bioinformatics* 2011, **27**:167–74.
52. Tsai IJ, Otto TD, Berriman M: **Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps.** *Genome Biol* 2010, **11**:R41.
53. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10**:R25.
54. Darling AC, Mau B, Blattner FR, Perna NT: **Mauve: multiple alignment of conserved genomic sequence with rearrangements.** *Genome Res* 2004, **14**:1394–1403.
55. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT: **Reordering contigs of draft genomes using the Mauve aligner.** *Bioinformatics* 2009, **25**:2071–2073.
56. Popendorf K, Tsuyoshi H, Osana Y, Sakakibara Y: **Murasaki: a fast, parallelizable algorithm to find anchors from multiple genomes.** *PLoS One*. 2010, **5**:e12651.
57. Kent WJ: **BLAT -the BLAST-like alignment tool.** *Genome Res* 2002, **12**:656–664.
58. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD: **VarScan: variant detection in massively parallel sequencing of individual and pooled samples.** *Bioinformatics* 2009, **25**:2283–2285.
59. Bustin SA: **Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems.** *J Mol Endocrinol* 2002, **29**:23–39.
60. Bustin SA, Benes V, Garson JA, Hellems J, Huggett HJ, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **The MIQE Guidelines: Minimum information for publication of quantitative real-time PCR experiments.** *Clinical Chemistry* 2009, **55**:611–622.
61. Taylor S, Waken M, Dijkman G, Alsarraj M, Nguyen M: **A practical approach to RT-qPCR- Publishing data that conform to the MIQE guidelines.** *Methods* 2010, **50**:51–55.
62. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-Delta Delta} C(T) method.** *Methods* 2001, **25**:402–408.

doi:10.1186/1471-2164-13-385

Cite this article as: Aguilar et al.: Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system. *BMC Genomics* 2012 **13**:385.