

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE CIENCIAS MARINAS
INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS

MAESTRÍA EN CIENCIAS EN OCEANOGRAFÍA COSTERA



**IDENTIFICACIÓN DE SISTEMAS DE DOS COMPONENTES QUE
PARTICIPAN EN LA REGULACIÓN DE LOS MECANISMOS DE
DEGRADACIÓN DE HIDROCARBUROS EN *Rhodococcus ruber* MSA14**

T E S I S

QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS PARA
OBTENER EL GRADO DE MAESTRA

PRESENTA

ANAHÍ SAMPAYO MARTÍNEZ

Ensenada, Baja California, México
Junio, 2025

**FACULTAD DE CIENCIAS MARINAS
INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS
POSGRADO EN OCEANOGRAFÍA COSTERA**

**IDENTIFICACIÓN DE SISTEMAS DE DOS COMPONENTES QUE
PARTICIPAN EN LA REGULACIÓN DE LOS MECANISMOS DE
DEGRADACIÓN DE HIDROCARBUROS EN *Rhodococcus ruber*
MSA14**

T E S I S

**QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS PARA
OBTENER EL GRADO DE**

MAESTRA EN CIENCIAS EN OCEANOGRAFÍA COSTERA

PRESENTA:

ANAHÍ SAMPAYO MARTÍNEZ

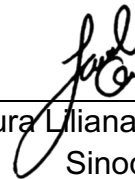
Aprobada por:



Dra. Hortencia Silva Jiménez
Director de tesis



Dr. Dimitris Georgellis
Sinodal



Dra. Laura Liana López Galindo
Sinodal

RESUMEN

La secuenciación masiva de ARN (RNA-seq) se ha consolidado como una herramienta fundamental para el análisis de la regulación transcripcional mediada por sistemas de transducción de señales como los sistemas de dos componentes (SDC), en la adaptación bacteriana a condiciones ambientales y metabólicas variables. *Rhodococcus ruber* MSA14 es una cepa hidrocarbonoclasta aislada de sedimentos marinos del Puerto de Rosarito, en la Costa de Baja California, México, que utiliza el pireno como única fuente de carbono y energía. Con la finalidad de comprender el proceso molecular de degradación del pireno, en este trabajo se estudió el transcriptoma de MSA14 en respuesta a este hidrocarburo mediante la secuenciación de ARN de alto rendimiento en la plataforma Illumina NovaSeq 6000.

Se identificaron 61 genes sobreexpresados y 71 subexpresados, incluyendo genes implicados en el catabolismo de hidrocarburos aromáticos como *benB*. La presencia de pireno indujo la expresión de genes relacionados con la detoxificación, la reconfiguración metabólica y la protección celular, lo que sugiere una respuesta adaptativa integral. Aunque no se identificaron explícitamente SDC directamente involucrados en el catabolismo de pireno, la modulación de genes sobreexpresados sugiere un ajuste dinámico de las vías reguladoras. Por lo tanto, esta tesis proporciona una visión global de la regulación transcripcional de *R. ruber* MSA14 en respuesta al pireno.

Palabras clave: *Rhodococcus ruber* MSA14, pireno, RNA-Seq, SDC.

ABSTRACT

RNA sequencing (RNA-seq) has become established as a fundamental tool for analyzing transcriptional regulation mediated by signal transduction systems, such as two-component systems (TCS), in bacterial adaptation to variable environmental and metabolic conditions. *Rhodococcus ruber* MSA14 is a hydrocarbonoclastic strain isolated from marine sediments of the Port of Rosarito, Baja California Coast, Mexico, which utilizes pyrene as its sole carbon and energy source. To understand the molecular process of pyrene degradation, this study investigated the transcriptome of MSA14 in response to this hydrocarbon using high-throughput RNA sequencing on an Illumina NovaSeq 6000 platform.

Sixty-one overexpressed and 71 downregulated genes were identified, including genes implicated in aromatic hydrocarbon catabolism such as *benB*. The presence of pyrene induced the expression of genes related to detoxification, metabolic reconfiguration, and cellular protection, suggesting a comprehensive adaptive response. Although no two-component systems (TCS) were explicitly identified as directly involved in pyrene catabolism, the modulation of overexpressed genes suggests a dynamic adjustment of regulatory pathways. Therefore, this thesis provides a global overview of the transcriptional regulation of *R. ruber* MSA14 in response to pyrene.

Keywords: *Rhodococcus ruber* MSA14, pyrene, RNA-Seq, TCS

ÍNDICE

<i>RESUMEN</i>	III
<i>ABSTRACT</i>	IV
<i>AGRADECIMIENTOS</i>	VII
<i>ÍNDICE DE TABLAS</i>	IX
<i>ÍNDICE DE FIGURAS</i>	X
<i>INTRODUCCIÓN</i>	12
SISTEMA DE DOS COMPONENTES (SDC)	13
HISTIDINAS QUINASAS (HQ)	15
REGULADORES DE RESPUESTA (RR)	17
<i>Rhodococcus ruber</i> MSA14	20
<i>JUSTIFICACIÓN</i>	25
<i>OBJETIVO GENERAL</i>	26
<i>OBJETIVOS ESPECÍFICOS</i>	26
<i>HIPÓTESIS</i>	27
<i>MATERIALES Y METODOLOGÍA</i>	28
CEPA BACTERIANA.....	28
CONDICIONES DE CRECIMIENTO	28
EXTRACCIÓN DE ARN TOTAL CON EL MÉTODO TRIhoz25.	29
SECUENCIACIÓN MASIVA DE ARN (RNA-Seq).....	32
ANÁLISIS DE DATOS DE SECUENCIACIÓN DE ARN PROCARIOTA.....	32
PROCESAMIENTO DE DATOS, ENSAMBLE DEL TRANSCRIPTOMA Y ANÁLISIS DE EXPRESIÓN DIFERENCIAL.	33
IDENTIFICACIÓN DE GENES SOBREENPRESADOS Y SUBEXPRESADOS EN PRESENCIA DE PIRENO.....	35
ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL.....	38
<i>RESULTADOS</i>	40
CRECIMIENTO CELULAR SIN Y CON PIRENO.	40
VERIFICACIÓN DE LA CALIDAD E INTEGRIDAD DEL ARN.....	42
CONTROL DE CALIDAD Y SECUENCIACIÓN MASIVA (RNASeq)	44
ANÁLISIS DE CALIDAD DE DATOS DE SECUENCIACIÓN Y ALINEAMIENTO DE SECUENCIAS. ..	46
ALINEAMIENTO DE LECTURAS Y ENSAMBLE DEL TRANSCRIPTOMA CON GENOMA GUIADO.	48

ANÁLISIS DE EXPRESIÓN GÉNICA MEDIANTE BALLGOWN Y STRINGTIE.....	49
IDENTIFICACIÓN DE GENES EXPRESADOS DIFERENCIALMENTE (DEGs) EN PRESENCIA DE PIRENO.....	53
ENRIQUECIMIENTO DE GENES EXPRESADOS DIFERENCIALMENTE.....	75
DAVID (GO).....	76
ShinyGO 0.81	80
<i>DISCUSIÓN</i>	89
ESTRATEGIAS DE ADAPTACIÓN DE <i>R. ruber</i> AL PIRENO A NIVEL TRANSCRIPCIONAL.	89
<i>CONCLUSIÓN</i>	100
<i>PERPECTIVAS</i>	101
<i>ANEXOS</i>	102
[1] PROTOCOLO ESTANDARIZADO DE EXTRACCIÓN DE ARN “TRIhoz25” en <i>R. ruber</i> MSA14.	102
[2] LIMPIEZA DE LECTURAS (PAIR-END) CON FASTQ “PREFILTRO y POSTFILTRO”	104
[3] COMPARACIÓN DE CONJUNTOS METODOLÓGICOS DE GSA-P MEDIANTE DIAGRAMAS DE VENN GENÉTICO.	109
[4] GSA-P y GsA-P no significativos.....	110
[5] GSA-P y GsA-P.....	113
<i>BIBLIOGRAFÍA</i>	115

AGRADECIMIENTOS

Esta maestría no habría sido posible sin muchísimas personas. Ante todo, gracias a mi familia por su amor y apoyo incondicional que me ha acompañado desde lejos durante todo este proceso. Gracias por impulsarme a lo largo de esta etapa y por enseñarme que, no importa cuántas veces se falle, lo único que importa es seguir trabajando duro con toda la confianza del mundo, los amo.

Gracias a mis mentores científicos:

- Dra. Hortencia Silva Jiménez, por la oportunidad de trabajar en su grupo de investigación, donde he aprendido incontables lecciones y numerosas habilidades.
- Dr. Dimitris Georgellis, por su entusiasmo por la ciencia, así como su capacidad para compartirlos con todos los que conoce. Agradezco la libertad que me brindó en su laboratorio durante mi estancia de posgrado.
- Dra. Laura Liliana López Galindo; por la orientación y el interés brindado en el transcurso del proyecto.

Gracias a la M en C. Claudia Rodríguez Rangel y al Dr. Adrián Fernando Alvarez, ambos del IFC de la UNAM, por el apoyo técnico experimental brindado durante mi estancia.

Gracias Ram por la motivación y apoyo emocional que me brindaste durante mis días más catastróficos en modo berserk. A mi amiga Mely, gracias por siempre hacer ameno mi día, te quiero.

Por último, pero ciertamente no menos importante gracias a SECIHTI por la beca otorgada (824191/1228786), al proyecto "ESTUDIO DE LA DIVERSIDAD BACTERIANA Y DE GENES MARCADORES ASOCIADOS A LA DEGRADACIÓN DE HIDROCARBUROS Y LA RESISTENCIA A ANTIBIÓTICOS EN LA ZONA COSTERA DE BAJA CALIFORNIA" y al proyecto "DISEÑO Y CARACTERIZACIÓN DE CONSORCIOS BACTERIANOS SINTÉTICOS PARA LA DEGRADACIÓN DE HIDROCARBUROS AROMÁTICOS POLICÍCLICOS BASADOS EN INTERACCIONES MICROBIANAS" que permitieron desarrollar esta tesis.

ÍNDICE DE TABLAS

Tabla I. Secuenciación del genoma de <i>Rhodococcus ruber</i> MSA14 (Embarcadero et al., 2024).	21
Tabla II. Sistemas de Dos Componentes, RR huérfanos e HQ huérfanas identificados en <i>Rhodococcus ruber</i> MSA14. RR: Reguladores de respuesta; HQ: Histidinas Quinasas. ...	23
Tabla III. Especificaciones de parámetros de muestras de ARN para RNA-seq procariota.	32
Tabla IV. Concentración y pureza del ARN extraído de <i>R. ruber</i> MSA14 en ausencia y presencia de pireno. La concentración se expresa en nanogramos por microlitro (ng/μL). La pureza se basa en la absorbancia A260/280. Tratamiento CONTROL= C; Tratamiento PIRENO = P. Las columnas representan las muestras y las filas los parámetros evaluados.	43
Tabla V. Secuencias de los adaptadores de secuenciación utilizados por Novogene Co. Ltd. ...	45
Tabla VI. Número de Integridad de ARN (RIN) de muestras experimentales. Tratamiento CONTROL: C1, C2, C3, C4; tratamiento PIRENO: P1, P2, P3, P4.	45
Tabla VII. Recorte de lecturas de secuencias Pair-End en formato FastQC proporcionadas por la plataforma Illumina, empleando TRIMMOMATIC v0.39. Secuencias rescatadas de buena calidad (% SA); lecturas no pareadas (F / R) (% SHF y % SHR, respectivamente) y lecturas descartadas (% SB).	48
Tabla VIII. Cantidad de genes sobreexpresados y subexpresados anotados en <i>R. ruber</i> MSA14 en presencia de pireno, identificados mediante cuatro diferentes estrategias metodológicas. A: Método A (Li et al., 2015; Kuijjer et al., 2020); B: Método B (Li et al., 2015; Kuijjer et al., 2020); C: Método C (Li et al., 2015; Kuijjer et al., 2020); D: Método D (Perlea et al. 2016); GSA-P: Genes Sobreexpresados y Anotados en presencia de Pireno; GsA-P: Genes subexpresados y Anotados en presencia de Pireno.	55
Tabla IX. Genes sobreexpresados en respuesta a pireno. fc: tasa de cambio; pval: valor de probabilidad; qval: p- valor ajustado (qval > 0.1 pero < 0.2 se consideran aceptables si fc > 1 y si pval ≤ 0.05).	60
Tabla X. Genes subexpresados en respuesta a pireno. fc: tasa de cambio; pval: valor de probabilidad; qval: p-valor ajustado (qval > 0.1 pero < 0.2 se consideran aceptables si fc < 1 y si pval ≤ 0.05).	68
Tabla XI. Enriquecimiento BP GSA-P (a) y GsA-P (b). FDR bajo indica mayor confianza; Núm. de Genes: cantidad de genes asociados con el término GO o vía KEGG correspondiente ...	81
Tabla XII. Genes agrupados por categorías funcionales definidas por términos GO de alto nivel	87
Tabla XIII. Genes no significativos, inicialmente identificados como GSA-P, sin embargo, no alcanzaron con los parámetros establecidos (valores fuera del umbral): cambio de expresión superior a uno (fc > 1), un valor pval ≤ 0.05 o un qval ≤ 0.1.	110
Tabla XIV. Genes no significativos, inicialmente identificados como GsA-P, sin embargo, no alcanzaron con los umbrales establecidos: un valor pval ≤ 0.05 o un qval ≤ 0.1.	111

ÍNDICE DE FIGURAS

- Figura 1. Mecanismo de expresión génica bacteriana mediante sistema de dos componentes (SDC). (HQ: Histidina Quinasa; RR: Regulador de Respuesta; CT: Control Transcripcional) (Elaboración propia). Autofosforilación: $\text{HK-His} + \text{ATP} \rightleftharpoons \text{HK-His}\sim\text{P} + \text{ADP}$ / Fosfotransferencia: $\text{HK-His}\sim\text{P} + \text{RR-Asp} \rightleftharpoons \text{HK-His} + \text{RR-Asp}\sim\text{P}$ / Desfosforilación: $\text{RR-Asp}\sim\text{P} + \text{H}_2\text{O} \rightleftharpoons \text{RR-Asp} + \text{P}_i$ (Stock et al., 2000). 15
- Figura 2. Diagrama esquemático que representa la organización modular de SDC. PAS: dominio sensor PAS; D: dominio catalítico; ADN: dominio de unión a ADN. Los dominios que contienen residuos de Asp se indican en verde, los que contienen His en rosa y los dominios auxiliares variables en azul. (A) Vía prototípica de dos componentes ejemplificada por el sistema osmorregulador de *E. coli*, en el que un único evento de transferencia de fosfato ocurre entre la proteína histidina quinasa (HQ) ortodoxa EnvZ y su proteína reguladora de respuesta (RR) OmpR. (B) El sistema Arc de *E. coli* ilustra un relevo fosfórico que involucra a la HQ híbrida ArcB y su regulador de respuesta afín ArcA. (C) Organización del dominio y modo de acción del SDC TodS/TodT; TodS contiene dos dominios de autoquinasa (Fuente: elaboración propia). 17
- Figura 3. Curva de crecimiento de *R. ruber* MSA14. Identificación de la fase media exponencial en medio BH suplementado con 0.3% de extracto de levadura (BH.3eL). De 0-4 h = Fase de latencia; de 4-36 h = Fase exponencial; de 36-48 h = Fase estacionaria; >48 h = Fase de muerte celular. Fase media exponencial = 20 h. El gráfico muestra el crecimiento de MSA14 (eje y) en función del tiempo transcurrido (eje x). Los cuadrados reflejan el crecimiento de MSA14 en ausencia de pireno, mientras que los triángulos marcan su crecimiento en presencia de pireno. 40
- Figura 4. Adaptación metabólica de *R. ruber* MSA14 en medio BH sin extracto de levadura (BhsE). La biomasa de *R. ruber* MSA14 de los tratamientos CONTROL y PIRENO provenientes de la fase media exponencial fueron incubadas en BhsE sin y con pireno, respectivamente, durante 6 horas para la posterior extracción de ARN. Tratamiento CONTROL (negro), representa el crecimiento de la cepa sin adición de compuestos externos; tratamiento PIRENO (gris), representa el crecimiento celular de MSA14 en presencia de pireno. 42
- Figura 5. Electroforesis en gel de agarosa [1%]. Condiciones experimentales = Tratamiento CONTROL: C1, C2, C3, C4; tratamiento PIRENO: P1, P2, P3, P4. 44
- Figura 6. Ejemplo de la evaluación de la librería A1_1 con FastQC; A) Reporte FastQC de la calidad de la secuencia por base (Prefiltrado); B) Reporte FastQC de la calidad de la secuencia por base (Postfiltrado con Trimomatic). Eje y= Qphred, eje x= posición de par de bases "bp" en secuencia de 150 pares de bases de longitud (Resto de limpieza de lecturas consultar en ANEXOS 2). 47
- Figura 7. Gráfico de la distribución de los valores de expresión génica transformados a $\log_2(\text{FPKM}+1)$. Muestras experimentales del tratamiento control en rosa (C1, C3, C4) y tratamiento con pireno en amarillo (P2, P3, P4). Eje y = distribución numérica de expresión génica (expresada como logaritmo base 2 de los FPKM, medida normalizada de la abundancia de transcritos), en base a cada muestra experimental (eje x). Ilustrado mediante cajas (que representan el rango intercuartílico-IQR) con bigotes (muestran el rango de los datos) y una línea central en el medio (mediana). Los puntos representan valores atípicos candidatos (de naturaleza indeterminada). 51
- Figura 8. Gráfico de calor de la expresión diferencial de genes con alta y baja expresión en tratamientos control (sin pireno) y con pireno involucrados en el crecimiento de *Rhodococcus ruber* MSA14, empleando valores de expresión transformados a logaritmo base 2 (FPKM+1). La normalización de datos logarítmica con +1 evita valores cero. Escala adaptativa, en rojo se indican valores altos de expresión y en azul, valores bajos de expresión. 52
- Figura 9. Gráfico de volcán de los cambios en la expresión génica entre el tratamiento con pireno y la condición control (sin pireno). Los genes sobreexpresados significativamente se

representan en rojo, los subexpresados significativamente en azul y aquéllos que no representan cambios significativos en gris. Los expresados marginalmente se representan en color rojo claro y azul claro, respectivamente. Cada punto representa un gen, cuya posición refleja dos aspectos clave: en el eje x, la magnitud del cambio en la expresión génica entre ambas condiciones expresada como el logaritmo base 2 de la tasa de cambio (\log_2 fold change) y en el eje y, la significancia estadística de dicho cambio representada como el logaritmo negativo en base 10 del valor q ($-\log_{10}$ q-value). Un valor positivo en el eje x indica que el gen está más expresado en la condición experimental (con pireno) en comparación con el control. La línea discontinua horizontal en $-\log_{10} q_{val} (0.05) = 1.3$ marca el umbral de significancia estadística, considerándose significativos aquellos valores superiores a esta línea. A mayor valor en el eje y, mayor es la significancia del cambio en la expresión génica. 54

- Figura 10. Relación de expresión de GsA-P (azul) y GSA-P (rojo). Eje y: Símbolo de genes, eje x: \log_2 fc. 75
- Figura 11. Análisis de enriquecimiento funcional de GSA-P en *R. ruber* MSA14 en respuesta al pireno empleando la base de datos DAVID. La anotación de los procesos biológicos (BP) se muestran en azul usando UP_KW y en naranja usando GOTERM. MF: Función Molecular; BP: Proceso biológico; CC: Composición celular. Gráfico circular: grupos de términos con mayor cobertura funcional relacionados a los GSA-P. 77
- Figura 12. Enriquecimiento funcional de GsA-P empleando la base de datos DAVID. Gráfico de barras: anotación de los procesos biológicos (BP) se muestran en color azul usando UP_KW y en color turquesa usando GOTERM_BP. MF: Función Molecular; BP: Proceso biológico; CC: Composición celular; Gráfico circular: grupos de términos con mayor cobertura funcional relacionados a los GsA-P. 79
- Figura 13. Obtención de ARN de *Rhodococcus ruber* MSA14 empleando el método TRIhoz25 (elaboración propia). A) Cinética de crecimiento en medio BH.3eL sin pireno y con 0.2 g L de pireno; B) Cinética de crecimiento en medio BHsE con 0.2 de pireno y sin pireno; C) Homogeneización de células de MSA14, inicio de extracción de ARN de MSA14 con TRIhoz25; D) Extracción y precipitación de ARN de MSA14, continuación de protocolo de extracción de ARN con TRIhoz25; E) Limpieza de ARN de MSA14, finalización de protocolo de extracción de ARN con TRIhoz25. 103
- Figura 14. Análisis cualitativo de limpieza de lecturas empleando FASTQ; Muestras de A1_2 – B1_2. 104
- Figura 15. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: C1_1 – D1_1. 105
- Figura 16. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: D1_2 – A2_2. 106
- Figura 17. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: B2_1 – C2_1. 107
- Figura 18. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: C2_2 - D2_2. 108
- Figura 19. Proporción de GSA-P (a) y GsA-P (b) únicos y compartidos, obtenidos de las cuatro estrategias metodológicas implementadas. 109
- Figura 20. Validación del conjunto de GSA-P, empleando la abundancia de transcritos (FPKM) de cada gen respecto al tratamiento control. Ejemplos: *nuoI*, *acpM*, *arcA*, *steA*, *pgfF* y *soxR*; sobreexpresión aproximada de 1.39, 1.16, 1.19, 1.71, 1.73 y 1.42 veces más, respecto al tratamiento control. Expresión de GSA-P (negro) respecto a su control sin pireno (gris). Eje y: Símbolo de genes, eje x: \log_2 FPKM. 113
- Figura 21. Validación del conjunto de GsA-P, empleando la abundancia de transcritos (FPKM) de cada gen respecto al tratamiento control. Ejemplos: *ahpF*, *nphA*, *metE*, *aceA*, *benB* y *manA*, con subexpresión aproximada de 1.96, 2, 1.95, 1.46, 1.14 y 1.67 veces más respecto al control. Expresión de GsA-P (negro) respecto a su control sin pireno (gris). Eje y: Símbolo de genes, eje x: \log_2 FPKM. 114

INTRODUCCIÓN

En los ecosistemas marinos, los sedimentos marinos son el sumidero principal de productos derivados del petróleo. Algunos de estos compuestos se biodegradan fácilmente, mientras que otros como los Hidrocarburos Aromáticos Policíclicos (HAP) son más persistentes y tóxicos. Estos compuestos orgánicos tóxicos pertenecen a los contaminantes de mayor impacto ambiental debido a sus efectos mutagénicos y/o carcinogénicos. Una forma de eliminar los hidrocarburos de petróleo que ingresan al ambiente marino es mediante la degradación mediada por comunidades microbianas hidrocarbonoclastas autóctonas (Head *et al.*, 2006; Yakimov *et al.*, 2007; Prince, 2010; Atlas & Hazen, 2011; Silva-Jiménez *et al.*, 2018; Patel *et al.*, 2020; Ahmad *et al.*, 2020).

Las bacterias hidrocarbonoclastas tienen la capacidad de regular, a nivel génico, su función celular con rasgos fisiológicos más complejos en respuesta al estrés por nutrición y a la presencia de contaminantes orgánicos tóxicos como los HAP. Debido a su hidrofobicidad, la disponibilidad de estos compuestos para estos organismos suele ser limitada; sin embargo, se han identificado cepas bacterianas capaces de degradar HAP y utilizarlos como fuente de carbono y energía (Seo *et al.*, 2009; Yu & Zhang *et al.*, 2012; Fuentes *et al.*, 2015; Guermouche *et al.*, 2015; Silva-Jiménez *et al.*, 2018; Ahmad *et al.*, 2020; Patel *et al.*, 2020; Embarcadero-Jiménez *et al.*, 2024).

Las bacterias, al ser microorganismos dinámicos y complejos, presentan variaciones en su densidad, diversidad y distribución dentro del ambiente. Esta

variabilidad refleja su capacidad para desarrollar sistemas internos sofisticados que les permiten detectar cambios en su entorno y responder adecuadamente. Entre estos mecanismos se encuentran los sistemas de transducción de señales, cuya función principal es activar respuestas adaptativas que favorecen la supervivencia de los microorganismos ante condiciones ambientales cambiantes (Parales *et al.*, 2008; Wuichet & Zhulin, 2010; Lacal *et al.*, 2013).

SISTEMA DE DOS COMPONENTES (SDC)

Entre los mecanismos moleculares de reconocimiento de señales celulares, los Sistemas de Dos Componentes (SDC) representan una de las principales vías de señalización intracelular en bacterias (Casino *et al.*, 2010; Krell *et al.*, 2010; Álvarez & Georgellis, 2023; Matilla *et al.*, 2025). Estos sistemas permiten detectar cambios ambientales repentinos y modulan la expresión génica como respuesta adaptativa al estímulo. Las respuestas reguladas por SDC abarcan desde funciones metabólicas básicas, como la utilización de carbono-nitrógeno, la asimilación de fosfato o el crecimiento aeróbico-anaeróbico, hasta procesos más especializados y complejos como la producción de biosurfactantes, la formación de esporas, biopelículas o cuerpos fructíferos, así como la quimiotaxis (Stock *et al.*, 2000; Galperin, 2005; Cock & Whitworth, 2007; Maneerat & Phetrong, 2007; Krell *et al.*, 2010; Groisman, 2016; Li *et al.*, 2020; Ahmad *et al.*, 2020).

La vía prototípica de transducción de señales en un SDC se basa en un mecanismo de fosforilación transitoria que involucra dos proteínas altamente conservadas. La primera es una proteína Histidina Quinasa (HQ)

transmembranal bifuncional, que actúa como sensor del estímulo externo. En ausencia de autofosforilación, esta proteína actúa como fosfatasa (Kenney, 2010). La segunda es una proteína Reguladora de Respuesta (RR) citosólica, que en la mayoría de los casos interactúa con regiones promotoras específicas del ADN para modular la transcripción génica (Stock *et al.*, 2000; West & Stock, 2001; Oshima *et al.*, 2002; Inouye & Dutta, 2002; Gao *et al.*, 2007; Capra & Laub, 2012; Zschiedrich *et al.*, 2016).

La mecánica de un SDC (Fig. 1) se basa en la detección de los estímulos extracelulares por parte del dominio sensor de HQ (Fig. 1.1). Esta detección activa la función enzimática de la HQ mediante la autofosforilación, proceso mediado por su dominio catalítico y de unión a ATP, que fosforila un residuo de Histidina (His) conservado en la propia HQ (Fig. 1.2). Posteriormente, el RR cataliza la transferencia del grupo fosfato desde la histidina fosforilada hacia un residuo de aspartato conservado (Fig. 1.3). Una vez fosforilado, el RR se activa y actúa como factor de transcripción (Fig. 1.4), uniéndose al ADN en la región promotora del gen u operón blanco, lo que garantiza una señalización precisa y específica (Fig. 1.5) (Stock *et al.*, 2000; Galperin, 2005; Groisman, 2016; Capra y Laub, 2012; Zschiedrich *et al.*, 2016; Terán-Melo *et al.*, 2019).

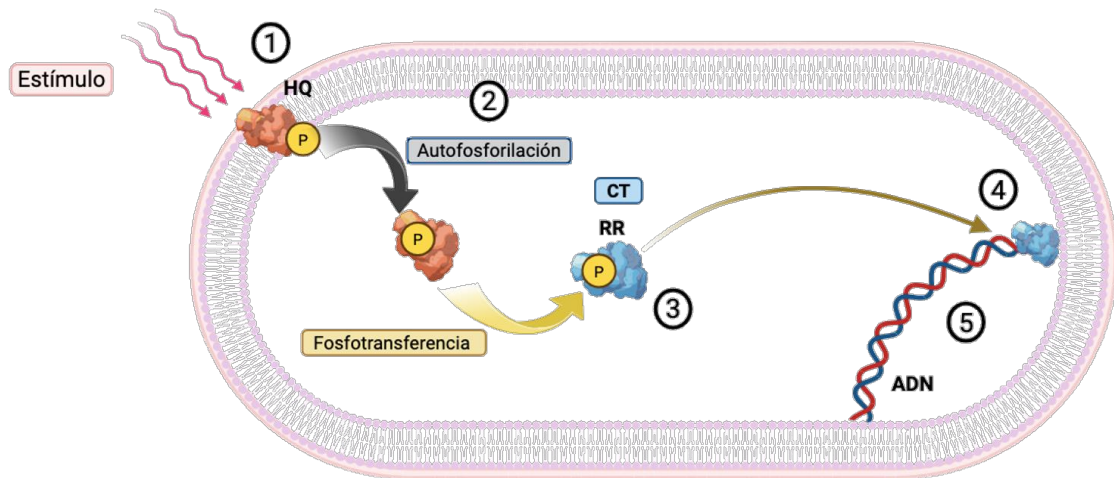


Figura 1. Mecanismo de expresión génica bacteriana mediante sistema de dos componentes (SDC). (HQ: Histidina Quinasa; RR: Regulador de Respuesta; CT: Control Transcripcional) (Elaboración propia). Autofosforilación: $HK-His + ATP \rightleftharpoons HK-His \sim P + ADP$ / Fosfotransferencia: $HK-His \sim P + RR-Asp \rightleftharpoons HK-His + RR-Asp \sim P$ / Desfosforilación: $RR-Asp \sim P + H_2O \rightleftharpoons RR-Asp + Pi$ (Stock et al., 2000).

La fosforilación o eliminación del dominio receptor induce cambios conformacionales que permiten su interacción con el dominio ATPasa de una segunda subunidad, lo cual posibilita la respuesta a una amplia variedad de estímulos dependiendo de su naturaleza (De Carlo, 2006). De este modo, las bacterias pueden adaptarse y sobrevivir en entornos muy diversos, ya que destinan una proporción significativa de su genoma a genes implicados en la regulación del metabolismo, como los circuitos de transducción de señales, en particular los SDC (Ulrich & Zhulin, 2007, 2009).

HISTIDINAS QUINASAS (HQ)

La familia de las histidina quinasa (HQ) exhibe una notable diversidad en cuanto a tamaño molecular, abarcando desde pequeñas proteínas hasta grandes complejos proteicos compuestos por cinco o seis dominios. Estas HQ se

clasifican en dos grupos principales: ortodoxas e híbridas, siendo éstas últimas aproximadamente el 25% del total (Stock *et al.*, 2000; Cock & Whitworth, 2007; Krell *et al.*, 2010). El dominio transmembranal de la mayoría de las HQ ortodoxas, como las proteínas *EnvZ* (Fig. 2A), *PhoQ* (Cheung *et al.*, 2008), *CitA* (Sevvana *et al.*, 2008), *LuxQ* (Min *et al.*, 2021), por mencionar algunas, representan la topología de membrana más común y funcionan como receptor de membrana periplásmica. Por otro lado, las HQ ortodoxas citoplásmicas, como la HQ de quimiotaxis *CheA* (Vázquez *et al.*, 2020; Peña & Georgellis, 2010) y la HQ reguladora de nitrógeno *NtrB* (North *et al.*, 2023) responden a estímulos intracelulares y/o interacciones con dominios citoplasmáticos de otras proteínas.

Por el contrario, la arquitectura de las HQ híbridas se caracteriza por la presencia de múltiples sitios fosfodonantes y fosfoaceptores, lo que les permite utilizar esquemas de fosforrelevo. Esta organización facilita la integración de múltiples señales y niveles de control dentro de una misma vía de señalización, como es el caso de *ArcB*, (Fig. 2B) una HQ representativa de este tipo. *ArcB* forma parte del sistema de control redox anóxico (*Arc*) en *E. coli*, (Vázquez *et al.*, 2020; Peña & Georgellis, 2010). Además, existen HQ híbridas con una estructura aún más compleja y que son parte de una familia de HQ denominadas “doble HQ”. Un ejemplo destacado de este grupo es *TodS* (Fig. 2C), una HQ que regula la ruta catabólica del tolueno a través de la vía de la tolueno dioxigenasa (TOD) en *Pseudomonas putida* DOT-T1E (Krell *et al.*, 2010; Silva-Jiménez *et al.*, 2012).

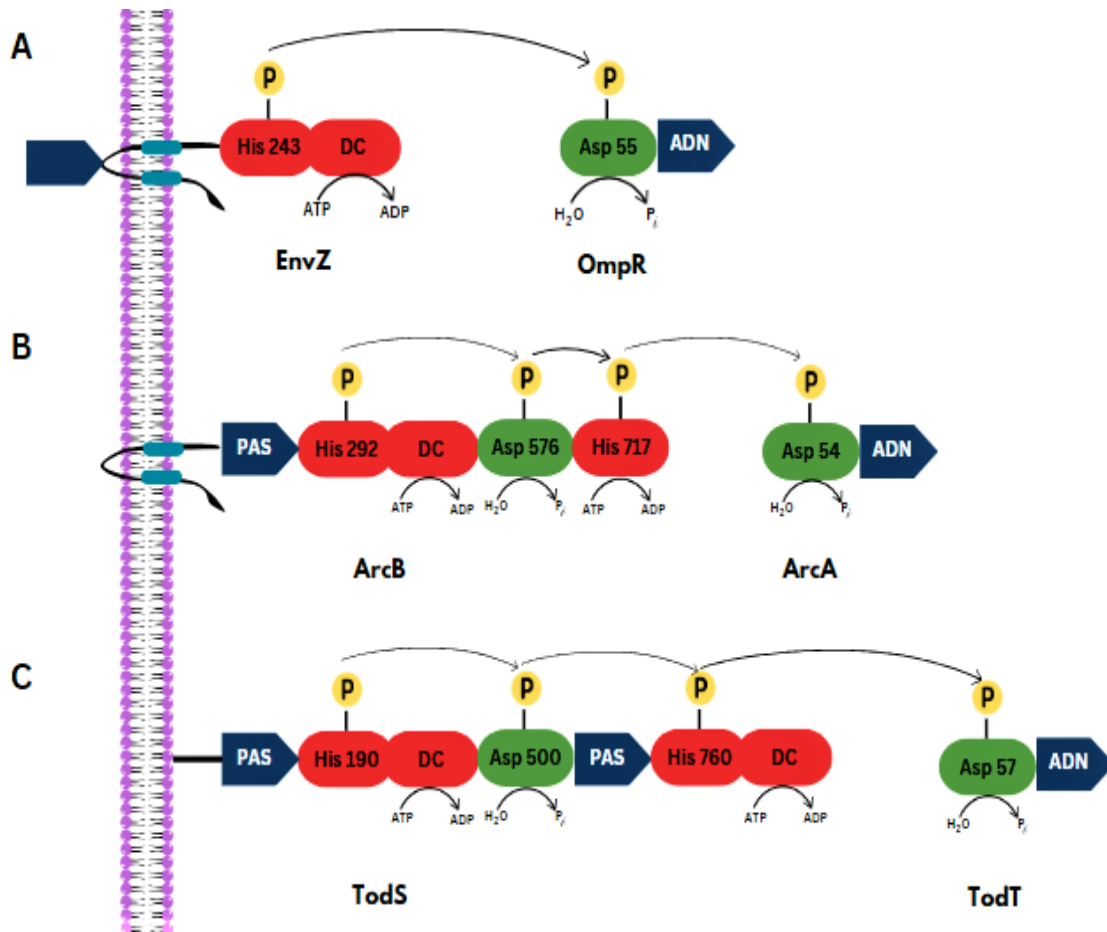


Figura 2. Diagrama esquemático que representa la organización modular de SDC. PAS: dominio sensor PAS; D: dominio catalítico; ADN: dominio de unión a ADN. Los dominios que contienen residuos de Asp se indican en verde, los que contienen His en rosa y los dominios auxiliares variables en azul. (A) Vía prototípica de dos componentes ejemplificada por el sistema osmorregulador de *E. coli*, en el que un único evento de transferencia de fosfato ocurre entre la proteína histidina quinasa (HQ) ortodoxa EnvZ y su proteína reguladora de respuesta (RR) OmpR. (B) El sistema Arc de *E. coli* ilustra un relevo fosfórico que involucra a la HQ híbrida ArcB y su regulador de respuesta afín ArcA. (C) Organización del dominio y modo de acción del SDC TodS/TodT; TodS contiene dos dominios de autoquinasa (Fuente: elaboración propia).

REGULADORES DE RESPUESTA (RR)

Los reguladores de respuesta suelen ser más abundantes que las histidinas quinasa y constituyen una familia clave de proteínas de señalización. Estas proteínas presentan actividad de autodesfosforilación, lo que limita la duración de

su estado fosforilado, con vidas medias que van desde segundos hasta horas. Típicamente, los RR son proteínas multidominio con una arquitectura modular compuesta por un dominio receptor conservado y conformacionalmente dinámico que puede activarse incluso sin estar fosforilado y un dominio efector variable, que actúa como interruptor en la respuesta de salida. Esta respuesta suele ser la regulación transcripcional, dependiente del nivel de fosforilación del RR (Volkman *et al.*, 2001; Mascher *et al.*, 2006; Maneerat & Phetrong, 2007). Así los RR pueden actuar como activadores, represores o ambos, regulando la expresión desde un único operón o de cientos de genes (Capra & Laub, 2012).

Distintos linajes bacterianos muestran variaciones significativas en las familias de RR que codifican, y estas diferencias tienden a correlacionarse con el tamaño del genoma del organismo. Sin embargo, especies estrechamente relacionadas suelen compartir perfiles similares de familias de RR; esta tendencia se mantiene a nivel de género y, con frecuencia, puede rastrearse hasta el nivel de filo (Toro *et al.*, 2005; Gao & Stock, 2009; Galperin, 2010). El análisis de aproximadamente 400 genomas bacterianos ha identificado cerca de 9000 RR, de los cuales la subfamilia *OmpR/PhoB* representa alrededor de un tercio (30%) evidenciando su alto grado de conservación evolutiva (Toro *et al.*, 2005). Alrededor del 17% de los RR consisten únicamente en un dominio receptor aislado, cuya actividad es regulada mediante interacciones intermoleculares. El porcentaje restante se agrupa en un conjunto relativamente pequeño de dominios funcionales con capacidad de unión a ácidos nucleicos, actividad enzimática o interacción con

proteínas/ligandos, lo que les permite ejercer funciones regulatorias a nivel transcripcional, postranscripcional o postraducciona (Gao & Stock, 2009; Galperin, 2010).

Aproximadamente entre el 60-70% de los RR contienen dominios efectores de unión al ADN, siendo las subfamilias *OmpR* y *NarL* las más representativas. En contraste, alrededor del 1% de los RR actúan como dominios de unión al ARN, pertenecientes en su mayoría a la subfamilia *AmiR*. Por otro lado, los RR con actividad enzimática constituyen aproximadamente el 13% y están representados principalmente por la subfamilia *CheB*. Finalmente, un pequeño pero diverso grupo de dominios efectores (alrededor del 3%) media interacciones con otras proteínas o ligandos como es el caso de la subfamilia *CheV* (Galperin, 2006; Gao *et al.*, 2007; Ulrich & Zhulin, 2007; Gao & Stock, 2009; Zschiedrich *et al.*, 2016).

Si bien no todas las bacterias tienen SDC como es el caso de *Mycoplasma genitalium* y *Candidatus Blochmannia floridanus*, muchas otras presentan más de 70 SDC como *Synechocystis sp.*, *Bacillus anthracis* y *B. subtilis* (Ashby & Houmard, 2006; Laub, 2011; Madan & Teichmann, 2003). No obstante, diversos análisis genómicos han revelado una correlación entre la cantidad de SDC y el tipo de ambiente en que se desarrollan los microorganismos (Galperin, 2005; Alm *et al.*, 2006; Han *et al.*, 2024).

En el ambiente marino, la documentación sistemática de la diversidad bacteriana se remonta a 1944, cuando Zobell y Upham caracterizaron 60 especies con un metabolismo versátil, capaces de albergar vías catabólicas que les permiten

aprovechar diversos compuestos orgánicos tóxicos (tal es el caso de los hidrocarburos) como única fuente de carbono y energía. Entre las bacterias que habitan en ambientes cambiantes y poseen capacidad catabólica frente a compuestos tóxicos, destaca el género *Rhodococcus* (Gonçalves *et al.*, 2006; Martinkova *et al.*, 2009; Ghosh *et al.*, 2010; Gürtler & Seviour, 2010; Jones & Goodfellow, 2010; Álvarez, 2019), el cual exhibe una notable capacidad para degradar una amplia variedad de compuestos xenobióticos como nitrilos, nitroaromáticos, compuestos heterocíclicos, bifenilos policlorados e hidrocarburos alifáticos y aromáticos, entre otros (Martínková *et al.*, 2009; Takeda *et al.*, 2010; Ghosh *et al.*, 2010; Di Gennaro *et al.*, 2010; Shimodaira *et al.*, 2012; Zhang *et al.*, 2012; Zampolli *et al.*, 2019).

Rhodococcus ruber MSA14

Entre las especies del género *Rhodococcus* que poseen una amplia reserva de rasgos genéticos y diversidad funcional, se encuentra *Rhodococcus ruber* MSA14 (Embarcadero-Jiménez *et al.*, 2024). Esta cepa fue aislada de sedimentos marinos someros contaminados con hidrocarburos en la costa del puerto de Rosarito, Baja California, México por el equipo de trabajo del Laboratorio de Investigación en Microbiología Ambiental del Instituto de Investigaciones Oceanológicas (IIO) de la Universidad Autónoma de Baja California (UABC) (Silva-Jiménez *et al.*, 2018).

Entre sus peculiaridades, *R. ruber* MSA14 se distingue por ser una bacteria Gram-positiva, aerobia, quimioorganotrófica, no móvil, no esporulante, con un

metabolismo oxidativo, productora de ácidos micólicos, mesófila, haloalcalitolerante y con un alto contenido de Guanina-Citosina en su ADN (70.4%). Esta cepa presenta la capacidad de producir biosurfactantes y muestra una alta tolerancia a hidrocarburos aromáticos policíclicos (HAP) como el pireno ($> 5 \text{ g L}^{-1}$). En condiciones experimentales, logró degradar hasta un 30% de pireno en 27 días a una concentración de 0.2 g L^{-1} , rendimiento que supera al observado con otros HAP como el naftaleno y el fenantreno (Embarcadero-Jiménez *et al.*, 2024).

El análisis genómico de *Rhodococcus ruber* MSA14 (GCF_038853135.1, Bioproject: PRJNA944780) (Tabla I) reveló la presencia de tres megaplásmidos lineales, una característica descrita en este género. Además, se identificaron múltiples secuencias codificantes para citocromos, transportadores del tipo ABC y MFS, y una notable abundancia de enzimas oxidorreductasas, que representan aproximadamente el 30% del total de enzimas codificadas (Embarcadero-Jiménez *et al.*, 2024).

Tabla I. Secuenciación del genoma de *Rhodococcus ruber* MSA14 (Embarcadero *et al.*, 2024).

Característica	Valor
Tamaño del genoma	6.3 Mpb
Tamaño cromosoma	5.45 Mpb
Número de tARN en cromosoma principal	59
Número de rARN en cromosoma principal	12
Tamaños de los plásmidos	72 Kbp, 110 Kbp, 185 Kbp y 479 Kbp
CDS en plásmidos	733
Contenido de GC	70.4%
Vías catabólicas centrales de compuestos aromáticos.	β -cetoadipato, 2-hidroxipentadienoato, gentisato, homogentisato, homoprotocatecuato y salicilato.
Vías catabólicas periféricas de compuestos aromáticos.	Degradación de benzoato, <i>p</i> -hidroxibenzoato, esteroides e isopropilbenceno.

El estudio genómico de *R. ruber* MSA14 realizado por Embarcadero-Jiménez *et al.* (2024) evidenció que esta cepa posee un repertorio diverso de genes catabólicos que le confieren la capacidad de degradar una amplia gama de hidrocarburos alifáticos, aromáticos y poliméricos. Mediante la base de datos HADEG (*Hydrocarbon Aerobic Degradation Enzymes and Genes*), se identificaron 16 genes ortólogos que codifican enzimas involucradas en la degradación de alcanos y alquenos, seis genes asociados a la degradación de polímeros como poliamidas (nylon) y polihidroxicanoatos, así como 30 genes ortólogos implicados en rutas catabólicas de compuestos aromáticos. Además, el análisis reveló singularidades genómicas en comparación con otras cepas de *R. ruber*, incluyendo la presencia exclusiva de genes ortólogos como *CHMO* (oxidación bi-terminal de alcanos), *xecE1* (degradación de alquenos) y *hpaH* (degradación de antranilatol), así como la duplicación del gen *andAa* involucrado en la asimilación de antranilatol.

Entre los genes ortólogos identificados mediante anotación funcional como participantes en la degradación de compuestos aromáticos, se detectaron oxidorreductasas, tales como *npcB_1*, asociada a la degradación del 4-nitrofenol; *apc3/apc4* implicadas en el metabolismo del etilbenceno; y *linE*, involucrada en la transformación de clorohidroquinona. Así mismo, se identificó un gen que codifica para un citocromo P450 (*thcB*), con posible actividad como dioxigenasa (Embarcadero-Jiménez *et al.*, 2024).

Mediante un análisis de anotación funcional utilizando el servidor RAST (Rapid

Annotation using Subsystems Technology) (Aziz *et al.*, 2008) se determinó que MSA14 cuenta con 22 SDC completos, 6 RR huérfanos y 2 HQ huérfanas (Kenedy-Fregoso, tesis en proceso, comunicación personal) (Tabla II). No obstante, para la mayoría de los SDC identificados en el genoma de *R. ruber* MSA14 aún se desconoce su función específica, así como su posible implicación en la regulación de genes involucrados en la degradación de HAP, como el pireno.

Tabla II. Sistemas de Dos Componentes, RR huérfanos e HQ huérfanas identificados en *Rhodococcus ruber* MSA14. RR: Reguladores de respuesta; HQ: Histidinas Quinasas.

Sensor (HQ)	Respuesta Reguladora (RR)	Función
HQ1	RR1 (NarL/FixJ)	-
HQ2	RR2 (LuxR)	-
HQ3 (PhoR)	RR3 (PhoP)	Inanición de fosfato
HQ4	RR4 (LuxR)	-
HQ5	RR5 (OmpR)	-
HQ6 (prxB)	RR6 (OmpR)	Detección de redox, fijación de nitrógeno
HQ7	RR7 (AlgR, IyT)	Biosíntesis de alginato
HQ8	RR8 (LuxR)	Sensor de hipoxia
HQ9 (MprB)	RR9 (MprA)	Persistencia en micobacterias
HQ10	RR10 (NarL/FixJ)	-
HQ11	RR11 (OmpR)	-
HQ12	RR12 (NarL/FixJ)	-
HQ13 (MtrB)	RR13 (MtrA)	Regulación de la división celular
HQ14	RR14 (LuxR)	-
HQ15	RR15 (LuxR)	-
HQ16	RR16 (LuxR)	-
HQ17 (PhoR, SphS)	RR17	Respuesta a la limitación de fosfato/biosíntesis de antibióticos
HQ18	RR18 (LuxR)	-
HQ19	RR19 (LuxR)	Fitocromo 1 de cianobacterias
HQ20	RR20 (LuxR)	-
HQ21	RR21 (OmpR)	-
HQ22 (KdpD)	RR22 (KdpE)	Canal de K ⁺ osmosensible
HQ23		
HQ24 (KdpD)		Canal de K ⁺ osmosensible (KdpD)
	RR25 (LuxR)	-
	RR26 (LuxR)	-
	RR27 (LuxR)	Diferenciación morfológica y biosíntesis de antibióticos
	RR28 (LuxR)	-
	RR29 (LuxR)	-
	RR30 (LuxR)	-

Diversos estudios han demostrado que la secuenciación de ARN de alto rendimiento (RNA-Seq) constituye una herramienta poderosa y rentable para el análisis del transcriptoma, ya que permite obtener una visión global del perfil de transcripción de ARN de una especie bajo condiciones específicas (Costa *et al.*, 2010; Brousseau *et al.*, 2014).

En este contexto, para comprender el metabolismo de contaminantes como el pireno en la cepa marina *Rhodococcus ruber* MSA14, el presente estudio empleó técnicas de secuenciación masiva RNA-seq para determinar los perfiles de expresión y los mecanismos moleculares involucrados en el catabolismo de pireno. Este trabajo representa el primer acercamiento transcriptómico para caracterizar la respuesta molecular y los mecanismos genéticos activados por *Rhodococcus ruber* MSA14 en respuesta a la presencia de pireno.

JUSTIFICACIÓN

Rhodococcus ruber MSA14 es una cepa con capacidad hidrocarbonoclasta frente al pireno y posee SDC en su genoma. Sin embargo, la mayoría de estos SDC no presentan homología con ortólogos de función conocida, en particular aquellos asociados con la regulación de genes u operones catabólicos y/o con mecanismos implicados en la degradación de hidrocarburos aromáticos policíclicos (HAP). Además, se conoce poco sobre la respuesta molecular y los mecanismos genéticos de *R. ruber* MSA14 frente a contaminantes orgánicos tóxicos como el pireno. Por ello, este estudio propone identificar los SDC potencialmente involucrados en la regulación del catabolismo de pireno en *R. ruber* MSA14 mediante un análisis transcriptómico. Este enfoque permitirá ampliar el conocimiento sobre los mecanismos moleculares que sustentan la degradación de pireno en esta cepa marina.

OBJETIVO GENERAL

Identificar los sistemas de dos componentes de *Rhodococcus ruber* MSA14 asociados a la regulación del catabolismo de hidrocarburos aromáticos policíclicos a través de un análisis transcriptómico.

OBJETIVOS ESPECÍFICOS

- Estandarizar el proceso de extracción de ARN de alta calidad en *R. ruber* MSA14.
- Identificar los sistemas de dos componentes que se expresan diferencialmente en MSA14 crecida en pireno.
- Seleccionar, de acuerdo con su contexto genómico, los sistemas de dos componentes potenciales que podrían regular el catabolismo de pireno.

HIPÓTESIS

Rhodococcus ruber MSA14 en presencia de pireno, expresa sistemas de dos componentes los cuales se encargan de regular genes del catabolismo de pireno.

MATERIALES Y METODOLOGÍA

CEPA BACTERIANA

Rhodococcus ruber MSA14, previamente aislada de sedimentos de la costa de Rosarito, B.C. (Silva-Jiménez *et al.*, 2018).

CONDICIONES DE CRECIMIENTO

La cepa MSA14 se cultivó rutinariamente a 30°C en placas de agar Zobell (pH 7.6 ± 2). El medio de cultivo contenía 5 g de peptona, 1 g de extracto de levadura, 0.01% de $FeCl_3$ y 15 g L^{-1} de agar. Se preparó con una mezcla de agua de mar (75%) y agua destilada (25%) esterilizada en autoclave a 121°C por 15 minutos. Previo a la solidificación del medio se añadió y homogeneizó una solución de pireno (Sigma-Aldrich) disuelto en acetona a una concentración de 20 g L^{-1} , para una concentración final de 0.1 g L^{-1} .

Para determinar la dinámica de crecimiento de la cepa MSA14 se partió de un preinóculo crecido en medio Bushnell Haas (BH) modificado estéril (BH.3eL) ($0.2 \text{ g L}^{-1} \text{ MgSO}_4$, $0.02 \text{ g L}^{-1} \text{ CaCl}_2$, $1 \text{ g L}^{-1} \text{ K}_2\text{HPO}_4$, $1 \text{ g L}^{-1} \text{ NH}_4\text{NO}_3$, $0.05 \text{ g L}^{-1} \text{ FeCl}_3$, $15 \text{ g L}^{-1} \text{ NaCl}$ y 0.3% (p/v) de extracto de levadura (YE) pH 7.0, incubado por 24 horas, a 30°C y en agitación constante (200 rpm).

Posterior a las 24 h, el pellet celular se recuperó por centrifugación a 6,000 rpm por 10 minutos a temperatura ambiente. El cultivo bacteriano se sometió a cuatro lavados celulares que consistieron en centrifugaciones a 13,000 rpm por un minuto, resuspendiendo el pellet celular obtenido en 1 mL de medio de cultivo BH.3eL. Posterior al cuarto lavado, el pellet recuperado se resuspendió en

BH.3eL y se utilizó para inocular 50 mL de medio BH.3eL sin pireno (tratamiento control) y con 0.2 g L^{-1} pireno a una densidad óptica inicial $DO_i = 0.1$ (600 nm). Los cultivos se incubaron a 30°C , en agitación constante (200 rpm) y en oscuridad. El crecimiento bacteriano se monitoreó cada dos horas por 48 horas en un espectrofotómetro rango UV y Visible VE-5100UV (VELAB) para determinar las fases del crecimiento celular e identificar la fase media exponencial de *R. ruber* MSA14. En cada medición se utilizó como blanco medio BH.3eL (Fig. 3).

A fin de asegurar la expresión de genes en presencia de pireno, el cultivo celular de la fase exponencial media de MSA14 sin y con pireno se centrifugaron a 6,000 rpm por 10 min. Seguidamente se realizaron cuatro lavados celulares con BH sin extracto de levadura (BHsE) a 13,000 rpm por 1 min. El pellet celular resultante de cada condición de crecimiento se resuspendió en 50 mL de medio BHsE sin y con pireno, respectivamente. Los cultivos se incubaron por 6 horas a 30°C y a 200 rpm. Al término del tiempo de incubación, se llevó a cabo la extracción de ARN (Fig. 4).

EXTRACCIÓN DE ARN TOTAL CON EL MÉTODO TRIhoz25.

El ARN total se extrajo a partir de cultivos bacterianos de 50 mL cultivados a 30°C en medio BHsE sin y con 0.2 g L^{-1} pireno. La extracción de ARN se obtuvo de cuatro réplicas de cada condición de crecimiento. Para ello, se estandarizó el método de extracción de ARN utilizando TRI Reagent (SIGMA) (proporcionado por la Dra. Laura L. López Galindo, Laboratorio de Genómica Funcional de

Organismos Marinos del IIO-UABC) dando como resultado el método TRIhoz25. En este método estandarizado, las células provenientes de cada réplica y de cada condición de crecimiento, se resuspendieron en 540 μ L de agua libre de ARNasas y se adicionaron 60 μ L de lisozima (0.1 g/mL) (cantidades empleadas por uso de lisozima caduca; con enzima vigente se emplean 10 μ L de lisozima a una concentración de 10 mg/mL). Las mezclas se incubaron a 37°C por 60 min, agitando por inversión cada 10 min.

Pasado el tiempo, la mezcla se centrifugó a 4°C y 14 000 rpm durante 2 min, se descartó el sobrenadante y se añadió 500 μ L de TRI Reagent al pellet recuperado. La mezcla, se homogeneizó suavemente con micropipeta. La suspensión celular, se transfirió a un tubo de 2 mL nuevo, estéril, con rosca conteniendo 100 mg de perlas de zirconia de 1.0 mm de diámetro (BioSpec). Las células, se lisaron empleando un homogeneizador de tejidos FastPrep24(™) (MP Biomedicals), sometiendo las células de *R. ruber* MSA14 a tres ciclos de 30 s a 6 m/s, incubando en frío por 5 min entre cada ciclo. Las células lisadas, se mantuvieron a temperatura ambiente durante 5 min, seguidamente se centrifugaron a 12,000 rpm por 10 min a 4°C (Fig. 17C). El sobrenadante se recuperó evitando las perlas de zirconia y se transfirió a un tubo de 1.7 mL nuevo estéril.

Se continuó con el proceso de extracción de ARN adicionando 100 μ L de cloroformo ACS (Fermont) a cada muestra, se agitaron en vórtex por 15 s y se dejaron reposar durante 5 min a temperatura ambiente. Transcurrido el tiempo, las mezclas se centrifugaron por 15 min a 13,000 rpm a 4°C, transfiriendo la fase

acuosa a un tubo nuevo, estéril (1.7 mL) evitando llevar trazas de la interfase y fase orgánica. Seguidamente se añadieron 250 µl de etanol al 95%, las muestras se agitaron por 10 s por inversión y se incubaron a -40°C por una noche (Fig. 17D).

Transcurridas 12 horas de incubación a -40°C, las muestras se centrifugaron a 13,000 rpm por 10 min a 4°C, proceso del cual se descartó el sobrenadante por pipeteo para adicionar 1 mL de etanol al 75% y resuspender el pellet de ARN total (dichos pasos se repitieron dos veces). Por último, las muestras se centrifugaron a 14,000 rpm por 5 min a 4°C, desechando el sobrenadante por pipeteo y dejando secar a temperatura ambiente el pellet de ARN obtenido. Posteriormente, los pellets fueron eluidos en 40 µL de agua libre de ARNasas y se almacenaron a -40°C (Fig. 17E).

La concentración y pureza del ARN extraído se determinó cuantificando espectrofotométricamente en un NanoDrop™ Lite (Thermo Scientific) con una relación A260/280 de 2.0 ± 0.2 . Para determinar la integridad del ARN de *R. ruber* MSA14, se preparó un gel de agarosa al 1% con solución amortiguadora TAE 1X (Tris 40 mM; ácido acético glacial 20 mM; EDTA 5 mM, pH 8.0) libre de ARNasas. El ARN se mezcló con SYBR Gold (PROMEGA) (para soluciones incoloras) y buffer de carga, dando paso a la corrida electroforética que consistió en sumergir el gel en solución TAE 1X por 45 min a 80 V, utilizando una cámara horizontal de Pharmacia Biotech GNA-100, así como una fuente de alimentación Fisher Scientific FB300. Finalmente, el gel se visualizó empleando un UView™ Mini-transilluminator (BIO-RAD).

SECUENCIACIÓN MASIVA DE ARN (RNA-Seq)

Las bibliotecas de secuenciación fueron generadas a partir de ARN total de cada muestra con y sin pireno (8 bibliotecas) de acuerdo con las especificaciones solicitadas por la empresa Novogene Co., Ltd (Beijing, China) (Tabla III), adicionando un análisis de calidad de las muestras de ARN total mediante el equipo Bioanalyzer. Una vez comprobada la calidad se generaron las bibliotecas para cada uno de los tratamientos con sus réplicas biológicas. La secuenciación se llevó a cabo utilizando la plataforma Illumina NovaSeq 6000 en formato pair-end de 150 ciclos con una cobertura genómica profunda y de alta calidad.

Tabla III. Especificaciones de parámetros de muestras de ARN para RNA-seq procariota.

Tipo de biblioteca	Tipo de muestra	Cantidad mínima	Volumen mínimo	Concentración mínima	Pureza (OD260/A280)
ARN procariótico	ARN total	≥ 500 ng	≥ 10 µL	≥ 50 ng/ µL	≥ 2.0

ANÁLISIS DE DATOS DE SECUENCIACIÓN DE ARN PROCARIOTA.

A partir de los datos crudos generados mediante la secuenciación masiva de RNA-Seq, se verificó el valor MD5 (sello de autenticidad) de cada archivo de lecturas crudas. Este procedimiento aseguró la integridad y trazabilidad de los datos crudos.

La calidad de las lecturas crudas se evaluó utilizando FastQC (versión 0.11.3). Posteriormente, las lecturas crudas se procesaron con Trimmomatic (versión 0.39) para eliminar adaptadores y secuencias de baja calidad, utilizando los

parámetros: TruSeq2-PE.fa:2:30:10:2:TRUE LEADING:5 TRAILING:30 SLIDINGWINDOW:4:15 MINLEN:36. Este proceso implicó el corte de los adaptadores y otras secuencias de adaptadores de Illumina (TruSeq2-PE.fa), un corte en bases al inicio de la lectura con un umbral de 5 (LEADING:5), un corte de bases al final de la lectura con un umbral de 30 (TRAILING:30), un escaneo con ventana móvil de 4 bases eliminando regiones con calidad inferior a 15 (SLIDINGWINDOW:4:15), y la eliminación de lecturas con una longitud inferior a 36 bases (MINLEN:36). Se realizó un segundo análisis de calidad con FastQC (versión 0.11.3) para confirmar la calidad de las lecturas limpias.

PROCESAMIENTO DE DATOS, ENSAMBLE DEL TRANSCRIPTOMA Y ANÁLISIS DE EXPRESIÓN DIFERENCIAL.

Para llevar a cabo el análisis bioinformático, se utilizó el clúster modelo HPE (OMICA) a través de un servidor VPN del Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Baja California, constituido de 30 nodos de procesamiento con 24 núcleos cada uno. Para el mapeo de lecturas de las muestras con el genoma de referencia de *Rhodococcus ruber* MSA14 (GCF_038853135.1, Bioproject: PRJNA944780), se empleó el software HISAT2 (Hierarchical Indexing for Spliced Alignment of Transcripts; Pertea *et al.*, 2016). Esta herramienta utiliza un sistema de indexación jerárquica que cubre integralmente el genoma, optimizando tanto la velocidad como la precisión del alineamiento. Este proceso incluyó:

1. La unión de los transcritos identificados en cada muestra para generar un conjunto de transcritos consenso.
2. La comparación de dichos transcritos con la anotación génica de referencia del genoma.
3. La estimación de la abundancia relativa de cada transcrito en las muestras, expresadas en valores de FPKM.

Mediante el software StringTie (version 2.2.1), se realizó el ensamblaje continuo de transcritos a partir de lecturas de buena calidad, utilizando como referencia el archivo de anotación genómica del NCBI, en formato GTF (Gene Transfer Format). Este proceso permitió reconstruir con precisión el transcriptoma, así como estimar la abundancia relativa de los transcritos generados para cada uno de los conjuntos de datos analizados.

Tras el ensamblaje inicial, los transcritos obtenidos se procesan mediante la función “merge”, la cual fusiona todas las estructuras génicas identificadas utilizando la herramienta gffcompare (versión 0.12.8) de StringTie. Este paso permite determinar cuántos de los transcritos ensamblados coinciden con los genes previamente anotados. A continuación, se ejecuta StringTie para reestimar las abundancias de los transcritos fusionados, lo que facilita la detección automática de nuevos genes y la generación de conteos de lecturas para cada transcrito, datos necesarios para su posterior análisis con Ballgown.

IDENTIFICACIÓN DE GENES SOBREEXPRESADOS Y SUBEXPRESADOS EN PRESENCIA DE PIRENO.

Finalmente, se procedió con el análisis exploratorio, la visualización y el modelado estadístico. Este proceso tuvo como objetivo medir cuantitativamente las diferencias en los niveles de expresión de los transcritos entre los dos tratamientos experimentales. Para ello se utilizó el lenguaje de programación R (versión 4.4.0) en conjunto con paqueterías de Bioconductor, una plataforma especializada para el análisis de datos genómicos.

El paquete Ballgown integró los transcritos y las abundancias generadas por StringTie, junto con la información fenotípica de las muestras en formato CSV (Comma-Separated Values). Posteriormente, agrupó los datos por condición experimental y determinó los genes que presentaban una expresión diferencial entre las condiciones analizadas.

A continuación, se aplicó un filtro para descartar genes con baja abundancia, considerando únicamente aquellos con valores de expresión > 1 . Además, se identificaron los transcritos con un valor $q < 0.05$ para establecer diferencias estadísticamente significativas entre las muestras. Este enfoque permitió clasificar los genes como regulados positiva o negativamente, según su perfil de expresión diferencial.

Se prosiguió con el análisis de la distribución de la abundancia de genes, expresada en valores de FPKM (*Fragments Per Kilobase of transcript per Million mapped reads*), entre las muestras. Dos de las tres bibliotecas por tratamiento

presentaron patrones de expresión muy distintos a los de sus réplicas biológicas por lo cual fueron eliminadas del estudio. El análisis se centró en las muestras CONTROL (1, 3, 4) y PIRENO (2, 3, 4), lo que permitió visualizar los genes sobreexpresados, subexpresados y de expresión basal mediante un diagrama de volcán en R. Este análisis empleó un *q-value* ajustado para controlar la tasa de descubrimientos falsos, además del fold change (fc - tasa de cambio) y su transformación a \log_2 , utilizados para cuantificar y normalizar los cambios en la expresión génica, facilitando la interpretación de los datos.

El análisis de expresión génica identificó genes marginales, los cuales presentan cambios en los niveles de expresión entre tratamientos, pero no alcanzan un umbral de significancia estricto; sin embargo, podrían tener relevancia biológica. Empleando los genes identificados por StringTie como expresados diferencialmente y los valores FPKM, se implementaron estrategias metodológicas adicionales enfocadas en la identificación de genes con posible importancia biológica, las cuales se describen a continuación:

- *Método A* (Li et al., 2015; Kuijjer et al., 2020): Análisis exploratorio y selección de genes con niveles de expresión por encima del tercer cuartil (Q3, percentil 75) considerados como de alta expresión y por debajo del primer cuartil (Q1, percentil 25), considerados de baja expresión.

```
> log2(FPKM+1)
```

```
> q ← c(0.25, 0.75)
```

- *Método B (Li et al., 2015; Kuijjer et al., 2020)*: Análisis básico de expresión diferencial entre condiciones experimentales, utilizando un enfoque basado en el umbral de cambio de expresión basado en valores de FPKM.

```
> log2PC ← log2(FPKM experimental/ FPKM control)
```

```
> umbral_log2 ← 1
```

```
> up experimental ← ifelse(log2PC > umbral_log2 & FPKM
experimental > FPKM control, TRUE, FALSE)
```

```
> low experimental ← ifelse(log2PC < -umbral_log2 & FPKM
experimental < FPKM control, TRUE, FALSE)
```

- *Método C (Li et al., 2015; Kuijjer et al., 2020)*: Evaluación de la expresión génica aplicando un umbral basado en la mediana y el cálculo de cuartiles para identificar genes diferencialmente expresados.

```
> data ← log2(FPKM+1)
```

```
> median ← median(data, na.rm = TRUE)
```

```
> high_expr_control <- names(log_data)[high_expr_genes]
```

```
> q3 ← quantile(data[high_expr_genes], probs = 0.75)
```

```
> high_q3 ← data > q3
```

```
> high_q3 ← names(data)[high_q3]
```

- *Método D (Pertea et al. 2016)*: Análisis de expresión diferencial estándar (DEA) utilizando umbrales de fold change (fc) y valores de significancia estadística (q-value).

```
> fc > 1 & qval < 0.05 S (S= Sobreexpresado)
```

```
> fc < 1 & qval < 0.05 s (s= subexpresado)
```

$>fc > 1 \ \& \ qval < 0.1 \ SM$ ($SM=$ Sobreexpresado [*Marginal*]).

$>fc < 1 \ \& \ qval < 0.1 \ sM$ ($sM=$ Subexpresado [*Marginal*]).

Posteriormente, a partir de las estrategias metodológicas e implementando la herramienta bioinformática Venny 2.1 para generación de diagramas de Venn (Anexo 3), se compararon las listas de los grupos de genes sobreexpresados y subexpresados en presencia de pireno para identificar aquéllos que fueran únicos de cada condición.

Como segundo filtro, se evaluaron los valores de *fold change*, considerando $fc > 1$ para los genes sobreexpresados anotados en presencia de pireno GSA-P y $fc < 1$ para genes subexpresados. Además, se aplicaron umbrales de significancia estadística con valores de $p \leq 0.05$ y $q \leq 0.1$.

ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL

Finalmente, se realizó un análisis de enriquecimiento funcional utilizando las plataformas DAVID (Database for Annotation, Visualization, and Integrated Discovery) y ShinyGO (v0.82). Para ello, se seleccionó la lista de genes expresados diferencialmente en respuesta a pireno. Este análisis permitió caracterizar la ontología génica e identificar términos biológicos sobrerrepresentados, agrupados en tres categorías principales: función molecular (FM), procesos biológicos (PB) y componentes celulares (CC). Adicionalmente, se compararon los resultados obtenidos con DAVID con aquellos generados por ShinyGO. En ambos casos, se utilizó el símbolo oficial de cada

gen junto con una lista de referencia con los genes anotados del genoma de *R. ruber* MSA14. No obstante, a diferencia de DAVID, en ShinyGO también se incluyó como referencia el genoma de *Rhodococcus ruber* Chol-4 (Guevara *et al.*, 2019), lo que permitió un análisis más amplio y comparativo.

RESULTADOS

CRECIMIENTO CELULAR SIN Y CON PIRENO.

El análisis de crecimiento de *R. ruber* MSA14 en medio BH.3eL sin (CONTROL) y con pireno (PIRENO) mostró que, en ambas condiciones, se estimula el crecimiento celular. Se observó una fase de latencia entre 0 y 4 horas, una fase exponencial de 4-36 horas, una fase estacionaria de 36-48 horas y una fase de muerte celular a partir de las 48 horas. En ambas condiciones, la cepa alcanzó el punto medio de la fase exponencial a las 20 horas (Fig. 3), con una $D_{0.600nm}$ de 1.6 en la curva CONTROL y de 1.2 en la curva PIRENO.

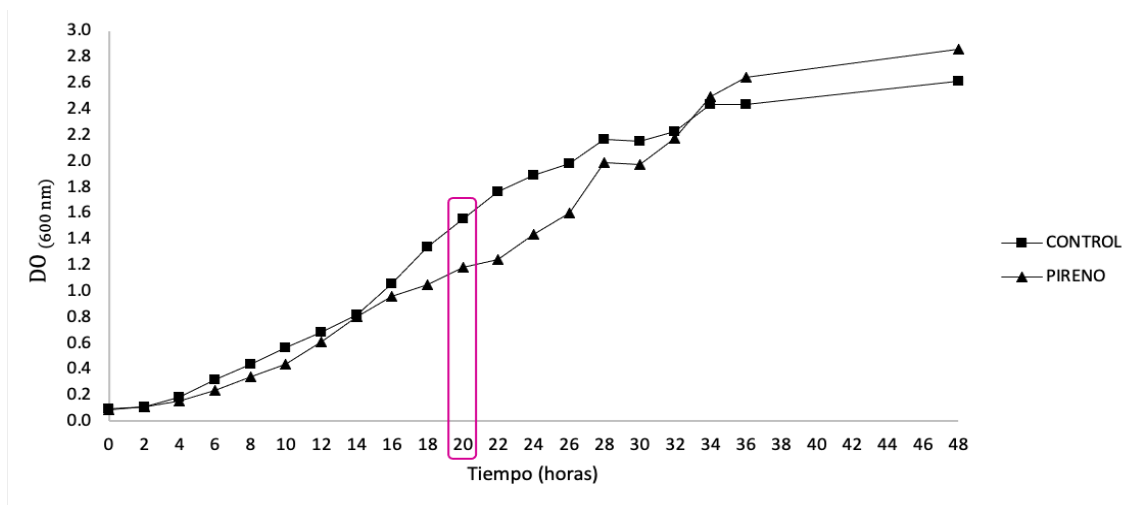


Figura 3. Curva de crecimiento de *R. ruber* MSA14. Identificación de la fase media exponencial en medio BH suplementado con 0.3% de extracto de levadura (BH.3eL). De 0-4 h = Fase de latencia; de 4-36 h = Fase exponencial; de 36-48 h = Fase estacionaria; >48 h = Fase de muerte celular. Fase media exponencial = 20 h. El gráfico muestra el crecimiento de MSA14 (eje y) en función del tiempo transcurrido (eje x). Los cuadrados reflejan el crecimiento de MSA14 en ausencia de pireno, mientras que los triángulos marcan su crecimiento en presencia de pireno.

La biomasa de *R. ruber* MSA14 recolectada durante la fase media exponencial en ambas condiciones (CONTROL y PIRENO) se inoculó completamente en medio BhsE sin y con pireno, respectivamente, seguido de una incubación de 6

horas. Tras este periodo se observó en la cepa MSA14 un comportamiento adaptativo en ausencia de una fuente de carbono adicional (extracto de levadura). En la figura 4 se puede observar que en el tratamiento CONTROL, la biomasa disminuyó un 50% respecto a su concentración inicial después de 6 horas, indicando que, al no contar con los nutrientes esenciales en el medio, el crecimiento celular se vio limitado.

En el tratamiento PIRENO, se observó una reducción de biomasa menos pronunciada respecto al tratamiento CONTROL. Esto sugiere que el pireno posiblemente actuó como fuente de carbono para *R. ruber* MSA14, aprovechando la adaptación metabólica previamente adquirida durante el cultivo precedente (BH.3eL). No obstante, al parecer, este beneficio coexistió con efectos tóxicos inherentes al compuesto, que indujeron un ajuste metabólico compensatorio. A pesar de esta posible dualidad funcional (uso como sustrato energético y estrés químico), la concentración celular final no mostró diferencias significativas respecto al CONTROL, alcanzando valores similares al término de la incubación.

Esta compleja interacción puede explicar las fluctuaciones en la biomasa registradas durante el monitoreo (Fig. 4), donde la capacidad adaptativa de la cepa le permite equilibrar la posible utilización del HAP como recurso con los mecanismos de resistencia al estrés.

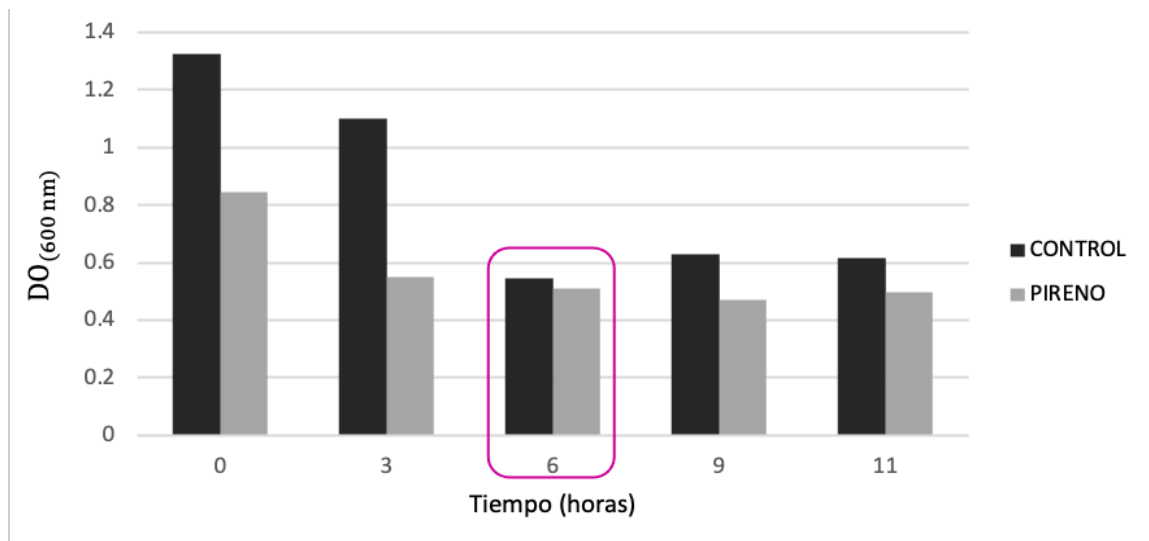


Figura 4. Adaptación metabólica de *R. ruber* MSA14 en medio BH sin extracto de levadura (BhsE). La biomasa de *R. ruber* MSA14 de los tratamientos CONTROL y PIRENO provenientes de la fase media exponencial fueron incubadas en BhsE sin y con pireno, respectivamente, durante 6 horas para la posterior extracción de ARN. Tratamiento CONTROL (negro), representa el crecimiento de la cepa sin adición de compuestos externos; tratamiento PIRENO (gris), representa el crecimiento celular de MSA14 en presencia de pireno.

VERIFICACIÓN DE LA CALIDAD E INTEGRIDAD DEL ARN

A partir de los pellets obtenidos de los cultivos de *R. ruber* MSA14 tras 6 horas de incubación en medio BHsE (sin y con pireno), se llevó a cabo la extracción de ARN utilizando el protocolo "TRIhoz25". El ARN obtenido de las ocho muestras divididas en cuatro réplicas por condición: CONTROL (C1-C4; sin pireno) y PIRENO (P1-P4; con pireno), fue cuantificado mediante Nanodrop Lite (Thermo Scientific).

La muestra C4 presentó la mayor concentración de ARN total (2,053.8 ng/μL), mientras que la muestra P1 mostró la menor concentración (542.8 ng/μL) (Tabla IV). En cuanto a la pureza del ARN, todas las muestras presentaron valores dentro del rango óptimo para ARN ($A_{260}/A_{280} = 2.0 \pm 0.2$), siendo la muestra C2,

la que registró la relación A260/280 ligeramente inferior (1.88) respecto al resto de muestras.

Tabla IV. Concentración y pureza del ARN extraído de *R. ruber* MSA14 en ausencia y presencia de pireno. La concentración se expresa en nanogramos por microlitro (ng/μL). La pureza se basa en la absorbancia A260/280. Tratamiento CONTROL= C; Tratamiento PIRENO = P. Las columnas representan las muestras y las filas los parámetros evaluados.

Muestra	C1	P1	C2	P2	C3	P3	C4	P4
[ng/μL]	1090.7	542.8	625.6	1938.8	1057.1	756.7	2053.8	1705.1
A260/280	2.02	1.92	1.88	2.05	1.96	2.01	2.05	1.96

Para evaluar la integridad del ARN extraído de *R. ruber* MSA14 en las condiciones CONTROL y PIRENO, se realizó una electroforesis en gel de agarosa al 1% con las cuatro réplicas biológicas por tratamiento. En todas las muestras se observaron bandas bien definidas correspondientes a los ARN ribosomales 23S, 16S y 5S (Fig. 5), lo que indica una alta calidad del ARN. La nitidez de estas bandas, sin evidencia de degradación, particularmente de la banda 23S, reflejó la ausencia de endonucleasas y confirmó la eficacia del método TRIzol25 para la extracción de ARN total íntegro en esta bacteria Gram-positiva.

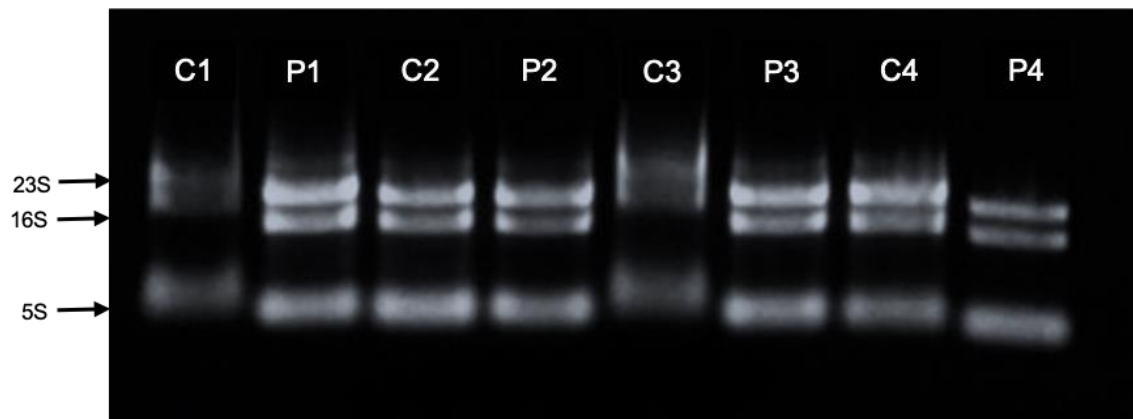


Figura 5. Electroforesis en gel de agarosa [1%]. Condiciones experimentales = Tratamiento CONTROL: C1, C2, C3, C4; tratamiento PIRENO: P1, P2, P3, P4.

CONTROL DE CALIDAD Y SECUENCIACIÓN MASIVA (RNASeq)

El análisis de calidad de los datos crudos de secuenciación correspondientes a las cuatro réplicas por tratamiento reveló que, si bien se esperaba una salida de datos de 2 GB por muestra, la réplica P4 generó únicamente 1.175 GB, representando el 58.7 % del volumen esperado. En total, las ocho muestras generaron una salida combinada de 24.1 GB, con un almacenamiento final de 8.82 GB.

La calidad de las secuencias obtenidas fue alta, con valores Qphred > 30 (Q30) en todas las muestras, lo que implica un error estimado de una base por cada 1000 nucleótidos. El porcentaje de lecturas de alta calidad (>Q30) osciló entre 93 % y 95 %, mientras que la tasa de error promedio a lo largo de las lecturas fue ≤ 0.03 %, lo que confirma la integridad y confiabilidad de los datos.

Tras la limpieza de las secuencias crudas con FastQC, más del 98 % de las lecturas se clasificaron como limpias, con excepción de la muestra C3, que presentó un 96.66 %. Aunque no se detectaron lecturas de nucleótidos

indeterminados (N), se identificó la presencia de adaptadores de secuenciación en todas las muestras, cuyas secuencias se muestran en la Tabla V.

Tabla V. Secuencias de los adaptadores de secuenciación utilizados por Novogene Co. Ltd.

Adaptador de 5':
5'-AGATCGGAAGAGCGTCGTGTAGGGAAAGGTAGATCTCGGTGGTCCGCGTATCATT-3'
Adaptador de 3':
5'-GATCGGAAGAGCACACGTCTGAACTCCAGTCACGGATGACTATCTCGTATGCCGTCTTCTGCTTG-3'

Las lecturas crudas obtenidas correspondieron a 40.4 millones de pares de lecturas *paired-end* para las muestras del tratamiento CONTROL y 39.8 millones para las muestras del tratamiento PIRENO. Entre las ocho muestras analizadas, C3 y P4 registraron las menores cantidades totales de lecturas, con más de 15 y 8 millones, respectivamente; mientras que las seis muestras restantes superaron consistentemente los 21 millones de lecturas por muestra. En cuanto al contenido de GC, todas las muestras presentaron valores dentro del rango de 60 al 66%, destacando P3 con el porcentaje más bajo (60.42 %). En el análisis de calidad del ARN, los valores de RIN (RNA Integrity Number) oscilaron entre 8.9 (el más bajo correspondiente a C3) y 9.9 (el más alto correspondiente a P4) (Tabla VI), considerando como referencia que un RIN de 10 representa un ARN íntegro y un valor de 1 indica degradación total.

Tabla VI. Número de Integridad de ARN (RIN) de muestras experimentales. Tratamiento CONTROL: C1, C2, C3, C4; tratamiento PIRENO: P1, P2, P3, P4.

Muestra	C1	P1	C2	P2	C3	P3	C4	P4
RIN	9.4	9.5	9.2	9.4	8.9	9.6	9.6	9.9

ANÁLISIS DE CALIDAD DE DATOS DE SECUENCIACIÓN Y ALINEAMIENTO DE SECUENCIAS.

La secuenciación masiva RNA-seq de ocho muestras (con y sin pireno) generó un total de 160,454,618 secuencias crudas, con un tamaño de lectura de 150 nt. El análisis de calidad de los datos crudos (raw reads) mediante FastQC reveló la presencia de adaptadores de secuenciación y lecturas de baja calidad ($Q < 30$, Fig. 6a), los cuales fueron eliminados mediante Trimmomatic. Los resultados arrojaron un total de 80,227,309 lecturas limpias y de buena calidad. En la tabla VII se muestra el número de secuencias de buena calidad (% SA), el porcentaje de lecturas no pareadas tanto Forward como Reverse (% SHF y % SHR) y el porcentaje de lecturas que fueron descartadas (% SB).

La eficiencia del filtrado se validó mediante un segundo análisis con FastQC , que mostró una mejora significativa en la calidad de las lecturas, con valores superiores a Q30 (Fig. 6b).

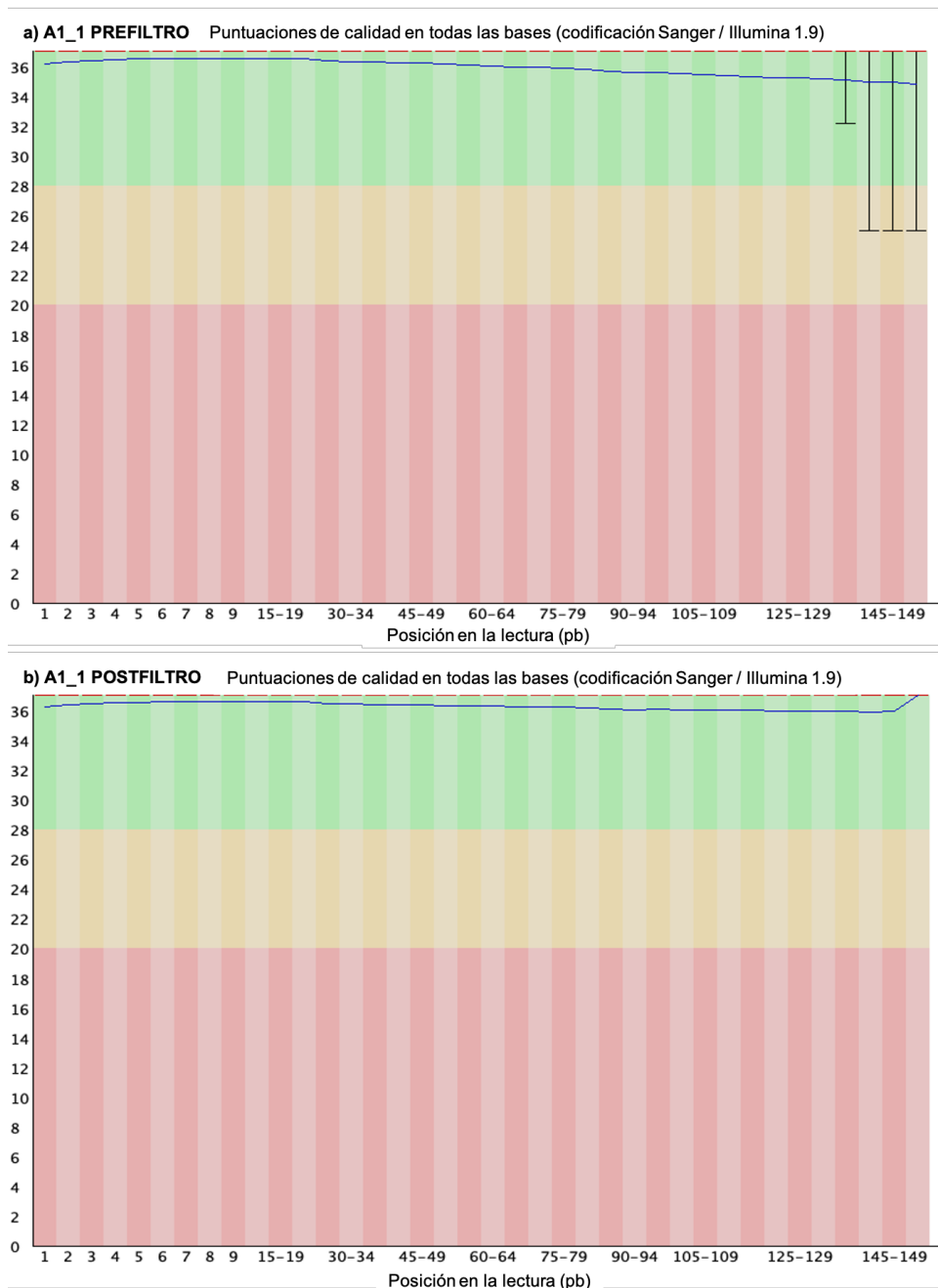


Figura 6. Ejemplo de la evaluación de la librería A1_1 con FastQC; A) Reporte FastQC de la calidad de la secuencia por base (Prefiltrado); B) Reporte FastQC de la calidad de la secuencia por base (Postfiltrado con Trimomatic). Eje y= Qphred, eje x= posición de par de bases "bp" en secuencia de 150 pares de bases de longitud (Resto de limpieza de lecturas consultar en ANEXOS 2).

Tabla VII. Recorte de lecturas de secuencias Pair-End en formato FastQC proporcionadas por la plataforma Illumina, empleando TRIMMOMATIC v0.39. Secuencias rescatadas de buena calidad (% SA); lecturas no pareadas (F / R) (% SHF y % SHR, respectivamente) y lecturas descartadas (% SB).

MUESTRA	Lecturas crudas	Lecturas limpias	% SA	% SHF	% SHR	%SB
CONTROL1	21293292	10646646	98.45	0.85	0.60	0.09
PIRENO1	23617288	11808644	98.51	0.80	0.58	0.10
CONTROL2	22224680	11112340	98.26	0.96	0.66	0.11
PIRENO2	21871260	10935630	98.30	0.95	0.63	0.11
CONTROL3	26152768	7991408	98.37	0.93	0.57	0.12
PIRENO3	21309066	13076384	98.32	0.90	0.69	0.10
CONTROL4	15982816	10654533	98.55	0.80	0.55	0.10
PIRENO4	8003448	4001724	98.21	0.99	0.65	0.15

ALINEAMIENTO DE LECTURAS Y ENSAMBLE DEL TRANSCRIPTOMA CON GENOMA GUIADO.

Se realizó el alineamiento de las lecturas de los tratamientos con y sin pireno al genoma de referencia de *Rhodococcus ruber* MSA14 (ID de ensamblaje del GenBank del NCBI: GCF_038853135.1). Seguidamente, utilizando el software StringTie y empleando gffcompare (vO.12.8), se comparó el transcriptoma con la anotación genómica de referencia, lo cual reveló una alta fidelidad en la reconstrucción. Para ello se ensambló un total de 6,063 transcritos, organizados en 4,439 loci genómicos, de los cuales 237 son transcritos multi-exón; por su parte el archivo de referencia contiene 5,820 ARNm en 4,809 loci (sin transcritos multi-exón) de los cuales 4,31 loci ensamblados se superponen con loci de referencia.

Las métricas de sensibilidad y precisión mostraron un rendimiento robusto del ensamblaje, alcanzando una sensibilidad del 100 % a nivel de base y del 96.5 % a nivel de locus, lo que indicó que prácticamente todas las regiones de referencia fueron cubiertas. La precisión del 95.2 % a nivel de base y del 97.4% a nivel de locus, valida la exactitud posicional de los ensamblajes. A nivel de transcrito, se obtuvo una sensibilidad del 82.4 % y una precisión del 79.1 %, con valores similares para el nivel de exón (77.4 % de sensibilidad y 71.7 % de precisión). A su vez, el ensamblaje reveló 4,797 transcritos y 4,640 loci que coinciden con el genoma de *R. ruber* MSA14. Estas métricas confirman que el ensamblaje realizado generó un conjunto de transcritos de alta calidad, lo que permitió un análisis confiable de la expresión génica de *R. ruber* MSA14 durante la fase de tiempo que se tomó la muestra.

ANÁLISIS DE EXPRESIÓN GÉNICA MEDIANTE BALLGOWN Y STRINGTIE

Para el procesamiento inicial de los datos, se utilizó el paquete Ballgown en el entorno de programación R. En primer lugar, se cargaron los metadatos de las muestras mediante un archivo en formato CSV (valores separados por comas), en el cual cada fila correspondió al nombre de la muestra y cada columna a una variable (condición experimental). Paralelamente, se importaron los datos de expresión génica generados por StringTie de cada transcrito. Como paso crítico de control de calidad, se filtraron los transcritos de baja abundancia ($FPKM < 1$), los datos analizados revelaron patrones inconsistentes entre réplicas P1 y C2, lo

que llevó a su exclusión y redujo el conjunto de datos de 6,063 a 5,218 transcritos en las seis muestras restantes analizadas. Posteriormente, se identificaron los transcritos con diferencias estadísticamente significativas entre los grupos experimentales, utilizando un umbral de $q\text{-value} < 0.05$. Este análisis reveló un total de 5,113 transcritos significativamente expresados y 105 transcritos no significativos.

Para visualizar la relación de expresión génica absoluta de los 5,113 transcritos entre las condiciones experimentales se emplearon los valores de FPKM (*Fragments Per Kilobase of transcript per Million mapped reads*). Estos valores se transformaron a escala logarítmica base 2, lo que permitió una distribución simétrica de los datos y redujo el sesgo asociado a la variabilidad en la longitud de los transcritos, facilitando la interpretación de los perfiles de expresión entre las réplicas como se observa en el diagrama de cajas (Figura 7), el cual, permite comparar, resumir y visualizar la dispersión de datos numéricos entre grupos basado en los cuartiles (Q1, Q2 o mediana, Q3), mediante los valores máximo y mínimo (o valores "bigote", que pueden ser valores fuera de los cuartiles), permitió identificar tendencias y patrones en los datos.

Posteriormente, la matriz de expresión se evaluó y corroboró entre las condiciones experimentales (sin y con pireno) mediante un gráfico de calor (Figura 8). Para ello, se promediaron los valores de FPKM de las réplicas para cada transcrito en cada tratamiento, del mismo modo los valores se convirtieron a escala logarítmica base dos.

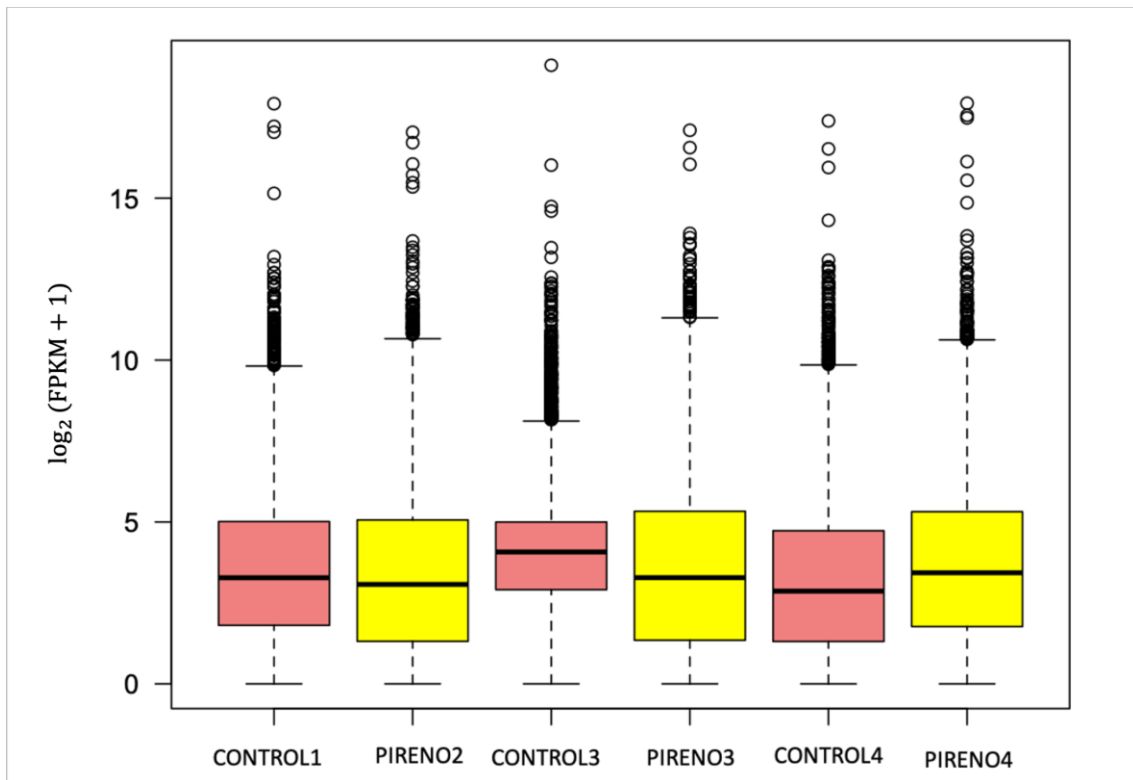


Figura 7. Gráfico de la distribución de los valores de expresión génica transformados a $\log_2(\text{FPKM} + 1)$. Muestras experimentales del tratamiento control en rosa (C1, C3, C4) y tratamiento con pireno en amarillo (P2, P3, P4). Eje y = distribución numérica de expresión génica (expresada como logaritmo base 2 de los FPKM, medida normalizada de la abundancia de transcritos), en base a cada muestra experimental (eje x). Ilustrado mediante cajas (que representan el rango intercuartílico-IQR) con bigotes (muestran el rango de los datos) y una línea central en el medio (mediana). Los puntos representan valores atípicos candidatos (de naturaleza indeterminada).

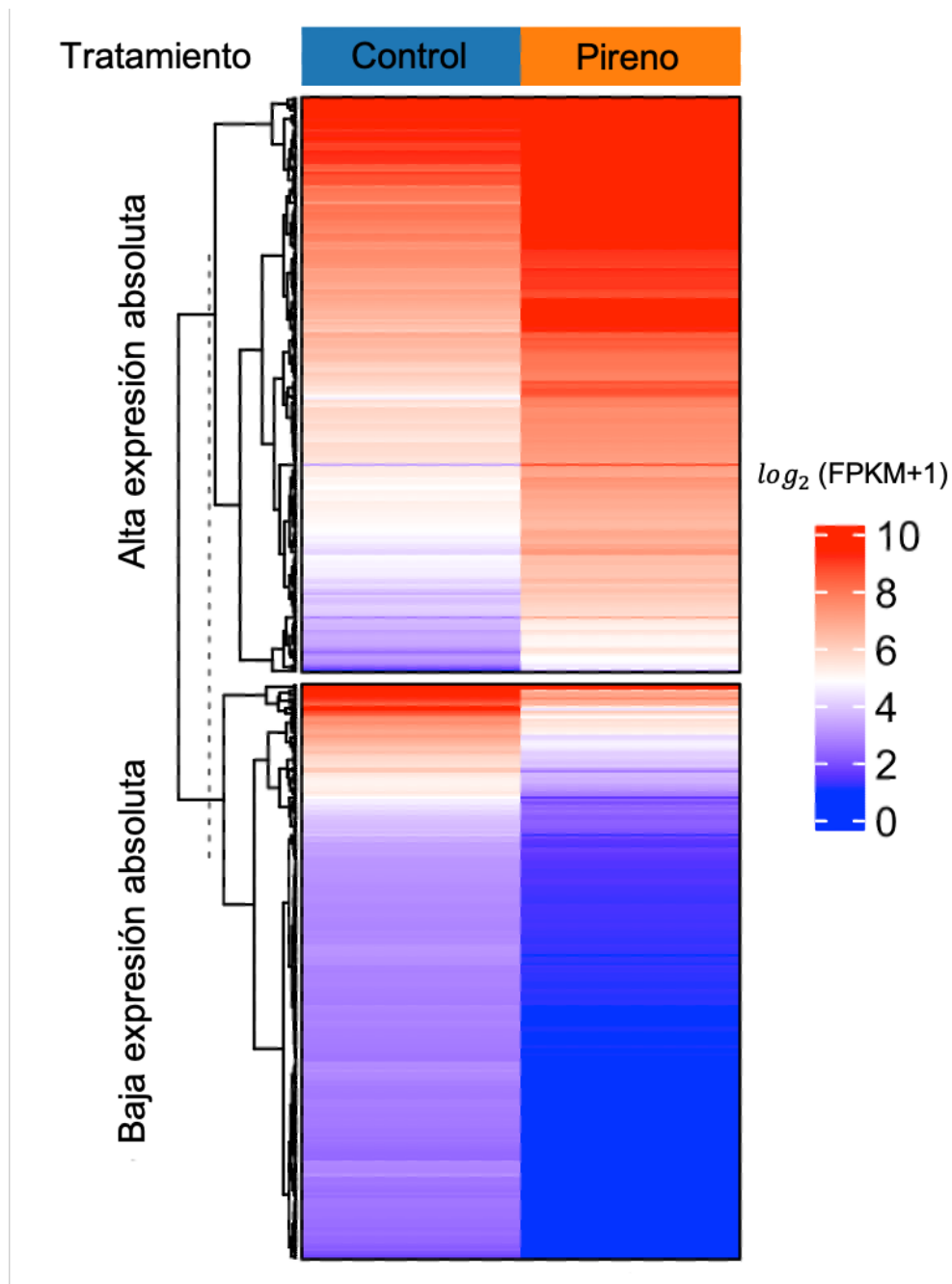


Figura 8. Gráfico de calor de la expresión diferencial de genes con alta y baja expresión en tratamientos control (sin pireno) y con pireno involucrados en el crecimiento de *Rhodococcus ruber* MSA14, empleando valores de expresión transformados a logaritmo base 2 (FPKM+1). La normalización de datos logarítmica con +1 evita valores cero. Escala adaptativa, en rojo se indican valores altos de expresión y en azul, valores bajos de expresión.

IDENTIFICACIÓN DE GENES EXPRESADOS DIFERENCIALMENTE (DEGs) EN PRESENCIA DE PIRENO.

Un gráfico de volcán (Fig. 9) ilustra la relación entre la magnitud del cambio y la significancia estadística de los transcritos, utilizando el logaritmo base 2 del cambio en el pliegue entre dos condiciones ($\log_2 fc$) y el logaritmo negativo base 10 del valor p , respectivamente. Este análisis resaltó inherentemente regiones críticas con expresión diferencial significativa. Del total de transcritos evaluados (anotados y no anotados), se identificaron 88 sobreexpresados y 17 subexpresados (Método D descrito en Materiales y Metodología); el resto correspondió a transcritos con expresión marginal o no significativa. Al asociar los 105 transcritos diferencialmente expresados con los símbolos oficiales de los genes de *R. ruber* MSA14, sólo se logró identificar con anotación nueve genes sobreexpresados y dos subexpresados.

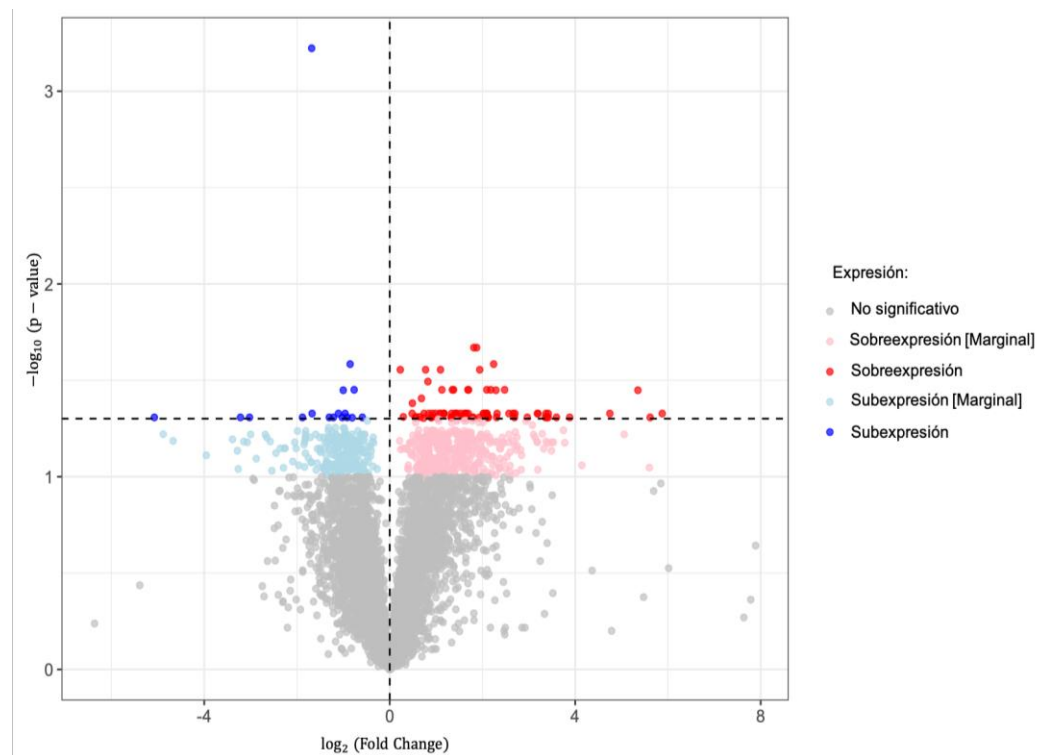


Figura 9. Gráfico de volcán de los cambios en la expresión génica entre el tratamiento con pireno y la condición control (sin pireno). Los genes sobreexpresados significativamente se representan en rojo, los subexpresados significativamente en azul y aquellos que no representan cambios significativos en gris. Los expresados marginalmente se representan en color rojo claro y azul claro, respectivamente. Cada punto representa un gen, cuya posición refleja dos aspectos clave: en el eje x, la magnitud del cambio en la expresión génica entre ambas condiciones expresada como el logaritmo base 2 de la tasa de cambio (\log_2 fold change) y en el eje y, la significancia estadística de dicho cambio representada como el logaritmo negativo en base 10 del valor q ($-\log_{10} q\text{-value}$). Un valor positivo en el eje x indica que el gen está más expresado en la condición experimental (con pireno) en comparación con el control. La línea discontinua horizontal en $-\log_{10} q\text{val}(0.05) = 1.3$ marca el umbral de significancia estadística, considerándose significativos aquellos valores superiores a esta línea. A mayor valor en el eje y, mayor es la significancia del cambio en la expresión génica.

Para obtener resultados de enriquecimiento fiables y estadísticamente significativos, es sugerible usar una lista de interés mayor a veinte genes, menor a esta cantidad los resultados de enriquecimiento funcional serán muy poco fiables o inexistentes en términos de significancia estadística y su interpretación biológica será extremadamente limitada (Huang *et al.*, 2009; Khatri *et al.*, 2012; Ge *et al.*, 2020).

De un total de 5,113 genes estadísticamente significativos identificados en este estudio, se observó que 4,470 carecían de anotación funcional, mientras que 643 contaban con identificación oficial. De este subconjunto anotado, solo 11´ genes mostraron un cambio diferencial significativo (sobreexpresión o subexpresión) al aplicar los valores $\log_2 fc$ y $-\log_{10} p\text{-value}$. Dada esta limitación en el número de genes con anotación diferencial y con el fin de obtener una perspectiva más amplia de los patrones de expresión, se amplió el análisis mediante la evaluación del enriquecimiento funcional. Para ello se utilizaron los valores de FPKM y se aplicaron los métodos A, B y C (descritos en la sección de Materiales y Metodología). Como resultado de estas cuatro estrategias metodológicas aplicadas, la Tabla VIII resume el número de genes anotados sobreexpresados (GSA-P) y subexpresados (GsA-P) obtenidos en presencia de pireno por cada método.

Tabla VIII. Cantidad de genes sobreexpresados y subexpresados anotados en *R. ruber* MSA14 en presencia de pireno, identificados mediante cuatro diferentes estrategias metodológicas. A: Método A (Li et al., 2015; Kuijjer et al., 2020); B: Método B (Li et al., 2015; Kuijjer et al., 2020); C: Método C (Li et al., 2015; Kuijjer et al., 2020); D: Método D (Pertea et al. 2016); GSA-P: Genes Sobreexpresados y Anotados en presencia de Pireno; GsA-P: Genes subexpresados y Anotados en presencia de Pireno.

MÉTODO	VALORES BASE	NÚMERO DE GSA-P	NÚMERO DE GsA-P
A	FPKM	163 genes	79 genes
B	FPKM	44 genes	112 genes
C	FPKM	85 genes	36 genes
D	fc & q-val	9 S	2 s
		44 SM.	49 sM

S= Sobreexpresado; s=subexpresado; SM= Sobreexpresado [Marginal]; sM= Subexpresado [Marginal].

Consecutivamente, los listados de genes obtenidos se sometieron a comparación con ayuda de la herramienta bioinformática Venny 2.1. Mediante diagramas de Venn se identificaron aquellos genes únicos y compartidos entre estrategias en presencia de pireno, eliminando genes duplicados de cada conjunto de datos en caso de ser identificados. Como resultado, se identificaron 15 genes sobreexpresados en común en los cuatro conjuntos de datos; 66, 5 y 20 genes incluidos exclusivamente en los conjuntos de los métodos A, B y D, respectivamente; cuatro genes en común en el conjunto A y B; 56 genes incluidos en los conjuntos A y C; tres genes compartidos entre el conjunto A y D; ocho genes incluidos en el conjunto B y D; ocho genes incluidos en los conjuntos A, B y C; tres genes en común entre los conjuntos A, B y D; y 3 genes en común en los conjuntos A, C y D (ANEXO 3a).

Por su parte, entre los genes subexpresados se identificaron dos genes en común en los cuatro conjuntos de datos de los diferentes métodos; 14, 24 y 20 exclusivos de los conjuntos A, B y D, respectivamente; 30 genes incluidos en los conjuntos A, B y F; 7 genes presentes en los conjuntos A, B y D; 21 genes incluidos en los conjuntos A y B; un gen en común entre el conjunto A y F; y 22 genes incluidos en los conjuntos B y D (ANEXO 3b).

En este sentido, al comparar los genes sobreexpresados en presencia de pireno respecto al tratamiento control, se identificaron 44 y 122 genes únicos de cada respectivo tratamiento, con un común de 147 genes. De estos últimos 51 genes fueron de relevancia biológica para ser anexados al concentrado de GSA-P (búsqueda manual), mientras que para los genes subexpresados en presencia

de pireno respecto al control, se identificaron 88 y 75 genes respectivamente, teniendo en común 53 genes, de los cuales 21 genes se anexaron a la lista de subexpresados por su relevancia biológica (búsqueda manual). Por lo tanto, de la expresión diferencial se obtuvo un total de 95 GSA-P y 109 GsA-P.

Para corroborar la significancia de la expresión de cada gen se consideraron los siguientes criterios: un $fc > 1$ para GSA-P y < 1 para GsA-P, un $p \leq 0.05$ y un $q \leq 0.1$, para ambos grupos de genes. Aplicando los criterios establecidos, se obtuvieron 61 GSA-P y 71 GsA-P (Tabla IX y X) (para consultar los genes que no cumplieron los umbrales, ver el ANEXO 4).

Sorprendentemente, entre los genes anotados diferencialmente expresados bajo las condiciones de crecimiento y en el tiempo en el que se tomó la muestra de la cepa, no se identificó ningún gen que codifique para HQ y RR de los 22 SDC completos previamente identificados en *R. ruber* MSA14, ni tampoco para 6 RR huérfanos y 2 HQ huérfanas reportados en esta cepa (Kenedy-Fregoso, tesis en proceso, comunicación personal). Lo anterior sugiere que estos genes pueden formar parte del grupo de 4,470 genes con expresión significativa sin anotación asignada.

Es relevante mencionar que genes implicados en la degradación de hidrocarburos aromáticos como *pcaG*, *catA*, *bphC* y *dmpG* (Tabla 16-ANEXO 4) no tuvieron una expresión diferencial significativa a una temprana exposición de *R. ruber* MSA14 al pireno. Sin embargo, es importante enfatizar que se detectaron genes diferencialmente expresados con patrones distintivos, que en

conjunto delimitan las prioridades metabólicas y las respuestas de estrés del microorganismo (Tabla IX y X).

El análisis integrado de los genes sobreexpresados y subexpresados en *Rhodococcus ruber* MSA14, sugiere una estrategia adaptativa dominante de "priorizar la adaptación y tolerancia sobre el catabolismo inmediato" frente al pireno.

Por ejemplo, la sobreexpresión de genes como *steA* (proteína putativa del anillo citocinético SteA, GO:0006629, GO:0006635) y *acpM* (proteína acarreadora de acilo para la extensión de meromicolato, GO:0071768), con un fc de 17 y 10, respectivamente, involucrados en la integridad de la pared y de la membrana celular; así como *pgeF* (Factor de edición del peptidoglicano PgeF, GO:0009251) con un fc > 10, que permite la correcta modulación de la membrana y del metabolismo lipídico, mecanismos que son necesarios para remodelar la membrana celular y reorientar los flujos de carbono y energía asegurando un suministro de energía futura (Álvarez, 2003; Eberlein *et al.*, 2018). Asimismo, *lipA* (Lipoil sintasa, GO:0009107) con un fc 1.3 y *lipB* con un fc 2.5 (Lipoil(octanoil) transferasa LipB, GO:0009249), participan ajustando la fluidez y permeabilidad de la membrana para proteger del daño oxidativo a la célula (Alvarez & Steinbüchel, 2019; de Carvalho, 2012). Dicha protección se potencia con el uso de reservas de glucógeno con *glgX* (Proteína de desramificación de glucógeno *GlgX*, GO:0005980) y *glgC* (Glucosa-1-fosfato adenililtransferasa, GO:0005978) con un fc de 2.1 y 3.2, correspondientemente, así como de polifosfato por *ppk2* expresada con un fc 2.7 (Polifosfato quinasa 2, GO:0006793), y se complementa

como mecanismo antioxidante con la expresión de *bfr* (Bacterioferritina, GO:0006879) con un cambio de 2.3 (Hernández *et al.*, 2008; Typas *et al.*, 2011; Urbano , *et al.*, 2014; Juarez, *et al.*, 2017; Rivera, 2017; Arosio, *et al.*, 2017; Cereijo, *et al.*, 2024). Aunque *ftsZ* (Proteína de división celular FtsZ, GO:0000910) es central para la división celular, su sobreexpresión junto con *ftsW* (Peptidoglicano Glicosiltransferasa, GO:0009252, GO:0051301) podría indicar un esfuerzo de la célula para mantener la integridad de la pared celular durante o después de eventos de daño (De Benedetti, *et al.*, 2021; Angeles & Scheffers, 2020) (Tabla IX).

La sobreexpresión de genes como *soxR* (Activador transcripcional sensible a redox SoxR, GO:0006355, GO:0006979) que regula la respuesta al estrés oxidativo, su inducción con un fc de 4.2 sugiere una respuesta directa y fundamental a la potencial formación de especies reactivas de oxígeno (ROS) generadas por la presencia del pireno. Como parte de la resiliencia de *R. ruber* MSA14, genes como *htpG* (Chaperona molecular HtpG, GO:0006457), le permiten mantener la calidad y el plegamiento correcto de sus proteínas, mientras que *groES* y *groL* (Co-chaperonina GroES, GO:0006457 y Chaperonina GroEL, GO:0042026; respectivamente) actúan en la reparación del ADN, siendo un mecanismo vital para la supervivencia celular en un entorno tóxico (Kim *et al.*, 2013; Gupta *et al.*, 2019; Elyamine *et al.*, 2021; Fedorov, 2022; Horovitz *et al.*, 2022)(Tabla IX).

Por otro lado, *R. ruber* MSA14 refleja un esfuerzo concertado de protección del ADN con la sobreexpresión de los genes *dps* (Proteína de Protección de fase

estacionaria/ Inanición de ADN *Dps*, GO:0006950, GO:0000723, GO:0009266, GO:0006979) y *recX*; (Regulador de recombinación *RecX*, GO:0000018), sugiriendo una preparación para un estado de menor crecimiento o latencia. Por su parte, genes como *nuoI* con fc 4,2 y *nuoK* con fc 4.4 (Subunidad de la NADH-quinona oxidoreductasa *NuoI*, GO:0015981, GO:0009061; y Subunidad de la NADH-quinona oxidoreductasa *NuoK*, GO:0042773, respectivamente) sugieren una potenciación de la cadena respiratoria y la producción de ATP, fundamental para llevar a cabo los procesos de adaptación y detoxificación (Karas *et al.*, 2015; Calisto *et al.*, 2021) (Tabla IX).

Tabla IX. Genes sobreexpresados en respuesta a pireno. *fc*: tasa de cambio; *pval*: valor de probabilidad; *qval*: *p*- valor ajustado (*qval* > 0.1 pero < 0.2 se consideran aceptables si *fc* > 1 y si *pval* ≤ 0.05).

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc > 1	pval ≤ 0.05	qval ≤ 0.1
<i>steA</i>	Proteína putativa del anillo citocinético <i>SteA</i>	Proceso metabólico de lípidos(GO:0006629); desaturación de ácidos grasos (GO:0006635)	17.67614	0.00961	0.08731
<i>acpM</i>	Proteína acarreadora de acilo para la extensión de meromicolato <i>AcpM</i>	Proceso biosintético del ácido micólico (GO:0071768)	10.62216	0.00425	0.06684
<i>pgeF</i>	Factor de edición del peptidoglicano <i>PgeF</i>	Metabolismo del peptidoglicano (GO:0009251)	10.48076	0.00094	0.04924
<i>ftsZ</i>	Proteína de división celular <i>FtsZ</i>	Citocinesis (GO:0000910)	9.08834	0.00055	0.04703
<i>sephH</i>	Proteína de septación H (<i>SepH</i>)	División celular (GO:0051301)	5.44380	0.00410	0.06628
<i>gvpJ</i>	Proteína de vesícula de gas <i>GvpJ</i>	Biogénesis de vesículas de gas (GO:0009313)	4.86071	0.00878	0.08405
<i>nuoK</i>	Subunidad de la NADH-quinona oxidoreductasa de <i>NuoK</i>	Transporte de electrones acoplados a la síntesis de ATP (GO:0042773)	4.46278	0.00424	0.06684
<i>soxR</i>	Activador transcripcional sensible a redox <i>SoxR</i>	Regulación de la transcripción con plantilla de ADN (GO:0006355); Respuesta al estrés oxidativo (GO:0006979)	4.29816	0.00078	0.04924
<i>aroB</i>	3-Deshidroquinato sintasa	Proceso biosintético de la familia de aminoácidos aromáticos (GO:0009073); Proceso biosintético de corismato (GO:00:00423)	4.27762	0.00074	0.04924
<i>nuoI</i>	Subunidad de la NADH-quinona oxidoreductasa de <i>NuoI</i>	Proceso metabólico de ATP (GO:0015981), cadena de transporte de electrones respiratoria (GO:0009061)	4.24611	0.00044	0.04703
<i>mihF</i>	Factor de integración al huésped, actinobacteriano	Regulación de la transcripción, molde ADN (GO:0006355).	3.93193	0.02204	0.11919

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc > 1	pval ≤ 0.05	qval ≤ 0.1
ugpC	Proteína de Unión a ATP del Transportador ABC de sn-Glicerol-3-Fosfato UgpC	Transporte de fosfato (GO:0006816)	3.90926	0.00244	0.06254
pyrE	Orotato fosforribosiltransferasa	Proceso biosintético de ribonucleótido de pirimidina (GO:0009220)	3.79597	0.00117	0.05233
murG	Undecaprenildifosfo-muramoilpentapéptido Beta-N-acetilglucosaminiltransferasa	Proceso biosintético de peptidoglicano (GO:0009252)	3.66592	0.00001	0.02140
queD	6-carboxitetrahydropterina sintasa QueD	Proceso biosintético de queosina (GO:0008616)	3.54320	0.01236	0.09553
pyk	Piruvato quinasa	Proceso glicolítico (GO:0006096)	3.53830	0.02851	0.13132
glp	Molibdotransferasa Glp similar a la gefirina	Proceso biosintético del cofactor de mo-molibdopterina (GO:0006777)	3.51404	0.00001	0.02140
glgC	Glucosa-1-fosfato adenililtransferasa	Proceso biosintético de glucógeno (GO:0005978)	3.21361	0.00603	0.07330
ftsW	Peptidoglicano Glicosiltransferasae FtsW	Proceso biosintético de peptidoglicano (GO:0009252); división celular (GO:0051301)	3.12442	0.00216	0.06043
lysX	Lisilfosfatidilglicerol Sintetasa Bifuncional/Lisina--tARN Ligasa LysX	Aminoacilación de lisil-tRNA (GO:00:006430)	3.09902	0.01896	0.11256
egtA	Ergotioneína biosíntesis glutamato--cisteína ligasa EgtA	Proceso biosintético de ergotioneína (GO:0052699)	2.87363	0.00447	0.06856
dnaE	Subunidad alfa de la ADN polimerasa III	Proceso biosintético NAD (GO:0009435)	2.87054	0.00376	0.06628
ppk2	Polifosfato quinasa 2	Proceso metabólico del fósforo (GO:0006793)	2.77845	0.04168	0.15700
bioB	Biotina sintasa BioB	Proceso biosintético de biotina (GO:0009102)	2.58884	0.05069	0.17454
prcB	Subunidad beta del proteasoma	Proceso catabólico de la proteína proteasómica (GO:0010498)	2.58830	0.01097	0.09119
nadE	NAD(+) Sintetasa Dependiente de Amoníaco	Proceso biosintético de NAD (GO:0009435)	2.52360	0.01158	0.09226
lipB	Lipoil(octanoil) transferasa LipB	Lipoilación de proteínas (GO:0009249)	2.50690	0.00076	0.04924
mftG	Micofactocina deshidrogenasa MftG	Proceso biosintético de la glicina betaína a partir de la colina (GO:0019285)	2.49550	0.02902	0.13223
hisN	Histidinol-Fosfatasa	Proceso biosintético de L-histidina (GO:0000105)	2.38287	0.00519	0.07135
bfr	Bacterioferritina	Homeostasis de iones de hierro intracelular (GO:0006879)	2.38256	0.04402	0.16217
dps	Proteína de Protección de fase estacionaria/Inanición de ADN Dps	Respuesta al estrés (GO:0006950), mantenimiento del genoma (GO:0000723), respuesta a la inanición (GO:0009266) y respuesta a estrés oxidativo (GO:0006979)	2.36149	0.01629	0.10599
glpK	Glicerol quinasa GlpK	Proceso metabólico del glicerol-3-fosfato (GO:0006072)	2.34387	0.00801	0.08149
mraZ	Represor transcripcional del Clúster de división/pared Celular MraZ	Regulación de la transcripción con plantilla de ADN (GO:0006355)	2.33107	0.01187	0.09338
dapF	Diaminopimelato epimerasa	Proceso biosintético de lisina a través de diaminopimelato (GO:0009089)	2.32099	0.01953	0.11351
gcvT	Sistema de escisión de glicina aminometiltransferasa GcvT	Descarboxilación de glicina a través del sistema de escisión de glicina (GO:0019464)	2.31218	0.02051	0.11596

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc > 1	pval ≤ 0.05	qval ≤ 0.1
htpG	Chaperona molecular HtpG	Plegamiento de proteínas (GO:0006457)	2.30504	0.00518	0.07135
recX	Regulador de recombinación RecX	Regulación de la recombinación del ADN (GO:0000018)	2.27490	0.01458	0.10101
glgX	Proteína de desramificación de glucógeno GlgX	Proceso catabólico de glucógeno (GO:0005980)	2.18283	0.00420	0.06684
hemQ	Hemo sintasa dependiente de peróxido de hidrógeno	Biosíntesis de hemo (GO:0006783)	2.17298	0.01833	0.11083
trpD	Fosforibosiltransferasa de antranilato	Proceso biosintético de L-triptófano (GO:0000162)	2.14467	0.00316	0.06511
eno	Fosfoenolpiruvato hidratasa	Proceso glicolítico (GO:0006096)	2.13084	0.00433	0.06739
tsaE	TRNA (adenosina (37)-N6)-treonilcarbamoiltransferasa compleja ATPasa subunidad tipo 1 TsaE	Modificación del ARNt de treonilcarbamoiladenosina (GO:0002949)	2.06038	0.00482	0.06999
arcA	Arginina deiminasa	Vía de la arginina deiminasa (GO:0019546)	2.01584	0.01201	0.09395
dbpB	DpdB de proteína que contiene el dominio DGQHR	Reparación del ADN (GO:0006281)	1.99038	0.00765	0.08015
bcp	Tiol peroxidasa dependiente de tiorredoxina	Respuesta al estrés oxidativo (GO:0006979)	1.93174	0.01218	0.09471
bluB	5,6-Dimetilbencimidazol sintasa	Proceso biosintético de cobalamina (GO:0009236); Proceso catabólico FMN (GO:0032363)	1.88903	0.00832	0.08329
dcm	ADN (citosina-5)-metiltransferasa	Modificación del ADN (GO:0006304)	1.88678	0.00140	0.05742
groES	Co-chaperonina GroES	Plegamiento de proteínas (GO:0006457)	1.87212	0.01619	0.10585
dpdA	tARN-Guanina transglicosilasa DpdA	Modificación del ARNt (GO:0006400)	1.81803	0.00924	0.08626
folP	Dihidropteroato sintasa	Proceso biosintético compuesto que contiene ácido fólico (GO:0009396)	1.77785	0.02765	0.12985
corA	Transportador de Magnesio/Cobalto CorA	Transporte de iones de cobalto (GO:0006824), transporte de iones de magnesio (GO:0015693)	1.68310	0.03832	0.15115
glgP	Fosforilasa de la Familia Alfa-Glucano	Proceso metabólico de carbohidratos (GO:0005975)	1.67944	0.00816	0.08249
pgi	Glucosa-6-Fosfato Isomerasa	Gluconeogénesis (GO:0006094); Proceso glicolítico (GO:0006096)	1.66140	0.00284	0.06254
groL	Chaperonina GroEL	Replegamiento de proteínas (GO:0042026)	1.64956	0.00826	0.08321
dacB	D-alanil-D-alanina carboxipeptidasa/D-alanil-D-alanina endopeptidasa	Proceso biosintético de peptidoglicano (GO:0009252)	1.62594	0.00872	0.08404
ftsH	Metaloproteasa de zinc dependiente de ATP FtsH	Proteólisis (GO:0006508); división celular (GO:0051301)	1.56737	0.00936	0.08686
nudC	NAD(+) difosfatasa	División (GO:0051301) y ciclo celular (GO:0007049)	1.52112	0.01025	0.08903
hemE	Decarboxilasa uroporfirinógena	Proceso biosintético compuesto que contiene porfirina (GO:0006779)	1.50869	0.00348	0.06570
aroC	Corismato sintasa	Proceso biosintético de la familia de aminoácidos aromáticos (GO:0009073)	1.43957	0.01835	0.11083
lipA	Lipoil sintasa	Proceso biosintético de lipoato (GO:0009107)	1.35685	0.04411	0.16221
pknB	Familia PASTA que contiene dominio Ser/Thr kinasa de la familia Stk1	Fosforilación de proteínas (GO:0006468)	1.21983	0.00502	0.07127

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc > 1	pval ≤ 0.05	qval ≤ 0.1
steA	Proteína putativa del anillo citocinético SteA	Proceso metabólico de lípidos(GO:0006629); desaturación de ácidos grasos (GO:0006635)	17.67614	0.00961	0.08731
acpM	Proteína acarreadora de acilo para la extensión de meromicolato AcpM	Proceso biosintético del ácido micólico (GO:0071768)	10.62216	0.00425	0.06684
pgeF	Factor de edición del peptidoglicano PgeF	Metabolismo del peptidoglicano (GO:0009251)	10.48076	0.00094	0.04924
ftsZ	Proteína de división celular FtsZ	Citocinesis (GO:0000910)	9.08834	0.00055	0.04703
sepH	Proteína de septación H (SepH)	División celular (GO:0051301)	5.44380	0.00410	0.06628
gvpJ	Proteína de vesícula de gas GvpJ	Biogénesis de vesículas de gas (GO:0009313)	4.86071	0.00878	0.08405
nuoK	Subunidad de la NADH-quinona oxidoreductasa de NuoK	Transporte de electrones acoplados a la síntesis de ATP (GO:0042773)	4.46278	0.00424	0.06684
soxR	Activador transcripcional sensible a redox SoxR	Regulación de la transcripción con plantilla de ADN (GO:0006355); Respuesta al estrés oxidativo (GO:0006979)	4.29816	0.00078	0.04924
aroB	3-Deshidroquinato sintasa	Proceso biosintético de la familia de aminoácidos aromáticos (GO:0009073); Proceso biosintético de corismato (GO:00:00423)	4.27762	0.00074	0.04924
nuoI	Subunidad de la NADH-quinona oxidoreductasa de NuoI	Proceso metabólico de ATP (GO:0015981), cadena de transporte de electrones respiratoria (GO:0009061)	4.24611	0.00044	0.04703
mihF	Factor de integración al huésped, tipo actinobacteriano	Regulación de la transcripción, molde ADN (GO:0006355).	3.93193	0.02204	0.11919
ugpC	Proteína de Unión a ATP del Transportador ABC de sn-Glicerol-3-Fosfato UgpC	Transporte de fosfato (GO:0006816)	3.90926	0.00244	0.06254
pyrE	Orotato fosforribosiltransferasa	Proceso biosintético de ribonucleótido de pirimidina (GO:0009220)	3.79597	0.00117	0.05233
murG	Undecaprenilidifosfo-muramoilpentapéptido Beta-N-acetilglucosaminiltransferasa	Proceso biosintético de peptidoglicano (GO:0009252)	3.66592	0.00001	0.02140
queD	6-carboxitetrahydropterina sintasa QueD	Proceso biosintético de queosina (GO:0008616)	3.54320	0.01236	0.09553
pyk	Piruvato quinasa	Proceso glicolítico (GO:0006096)	3.53830	0.02851	0.13132
glp	Molibdotransferasa Glp similar a la gefirina	Proceso biosintético del cofactor de mo-molibdoterina (GO:0006777)	3.51404	0.00001	0.02140
glgC	Glucosa-1-fosfato adenililtransferasa	Proceso biosintético de glucógeno (GO:0005978)	3.21361	0.00603	0.07330
ftsW	Peptidoglicano Glicosiltransferasae FtsW	Proceso biosintético de peptidoglicano (GO:0009252); división celular (GO:0051301)	3.12442	0.00216	0.06043
lysX	Lisilfosfatidilglicerol Sintetasa Bifuncional/Lisina--tARN Ligasa LysX	Aminoacilación de lisil-tRNA (GO:00:006430)	3.09902	0.01896	0.11256
egtA	Ergotioneína biosíntesis glutamato--cisteína ligasa EgtA	Proceso biosintético de ergotioneína (GO:0052699)	2.87363	0.00447	0.06856
dnaE	Subunidad alfa de la ADN polimerasa III	Proceso biosintético NAD (GO:0009435)	2.87054	0.00376	0.06628
ppk2	Polifosfato quinasa 2	Proceso metabólico del fósforo (GO:0006793)	2.77845	0.04168	0.15700

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc > 1	pval ≤ 0.05	qval ≤ 0.1
bioB	Biotina sintasa BioB	Proceso biosintético de biotina (GO:0009102)	2.58884	0.05069	0.17454
prcB	Subunidad beta del proteasoma	Proceso catabólico de la proteína proteasómica (GO:0010498)	2.58830	0.01097	0.09119
nadE	NAD(+) Sintetasa Dependiente de Amoníaco	Proceso biosintético de NAD (GO:0009435)	2.52360	0.01158	0.09226
lipB	Lipoil(octanoil) transferasa LipB	Lipoilación de proteínas (GO:0009249)	2.50690	0.00076	0.04924
mftG	Micofactocina deshidrogenasa MftG	Proceso biosintético de la glicina betaina a partir de la colina (GO:0019285)	2.49550	0.02902	0.13223
hisN	Histidinol-Fosfatasa	Proceso biosintético de L-histidina (GO:0000105)	2.38287	0.00519	0.07135
bfr	Bacterioferritina	Homeostasis de iones de hierro intracelular (GO:0006879)	2.38256	0.04402	0.16217
dps	Proteína de Protección de fase estacionaria/Inanición de ADN Dps	Respuesta al estrés (GO:0006950), mantenimiento del genoma (GO:0000723), respuesta a la inanición (GO:0009266) y respuesta a estrés oxidativo (GO:0006979)	2.36149	0.01629	0.10599
glpK	Glicerol quinasa GlpK	Proceso metabólico del glicerol-3-fosfato (GO:0006072)	2.34387	0.00801	0.08149
mraZ	Represor transcripcional del Clúster de división/pared Celular MraZ	Regulación de la transcripción con plantilla de ADN (GO:0006355)	2.33107	0.01187	0.09338
dapF	Diaminopimelato epimerasa	Proceso biosintético de lisina a través de diaminopimelato (GO:0009089)	2.32099	0.01953	0.11351
gcvT	Sistema de escisión de glicina aminometiltransferasa GcvT	Descarboxilación de glicina a través del sistema de escisión de glicina (GO:0019464)	2.31218	0.02051	0.11596
htpG	Chaperona molecular HtpG	Plegamiento de proteínas (GO:0006457)	2.30504	0.00518	0.07135
recX	Regulador de recombinación RecX	Regulación de la recombinación del ADN (GO:0000018)	2.27490	0.01458	0.10101
glgX	Proteína de desramificación de glucógeno GlgX	Proceso catabólico de glucógeno (GO:0005980)	2.18283	0.00420	0.06684
hemQ	Hemo sintasa dependiente de peróxido de hidrógeno	Biosíntesis de hemo (GO:0006783)	2.17298	0.01833	0.11083
trpD	Fosforibosiltransferasa de antranilato	Proceso biosintético de L-triptófano (GO:0000162)	2.14467	0.00316	0.06511
eno	Fosfoenolpiruvato hidratasa	Proceso glicolítico (GO:0006096)	2.13084	0.00433	0.06739
tsaE	TRNA (adenosina (37)-N6)-treonilcarbamoiltransferasa compleja ATPasa subunidad tipo 1 TsaE	Modificación del ARNt de treonilcarbamoiladenosina (GO:0002949)	2.06038	0.00482	0.06999
arcA	Arginina deiminasa	Vía de la arginina deiminasa (GO:0019546)	2.01584	0.01201	0.09395
dbpB	DpdB de proteína que contiene el dominio DGQHR	Reparación del ADN (GO:0006281)	1.99038	0.00765	0.08015
bcp	Tiol peroxidasa dependiente de tiorredoxina	Respuesta al estrés oxidativo (GO:0006979)	1.93174	0.01218	0.09471
bluB	5,6-Dimetilbencimidazol sintasa	Proceso biosintético de cobalamina (GO:0009236); Proceso catabólico FMN (GO:0032363)	1.88903	0.00832	0.08329
dcm	ADN (citosina-5)-metiltransferasa	Modificación del ADN (GO:0006304)	1.88678	0.00140	0.05742
groES	Co-chaperonina GroES	Plegamiento de proteínas (GO:0006457)	1.87212	0.01619	0.10585
dpdA	tARN-Guanina transglicosilasa DpdA	Modificación del ARNt (GO:00:006400)	1.81803	0.00924	0.08626

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc > 1	pval ≤ 0.05	qval ≤ 0.1
<i>folP</i>	Dihidropteroato sintasa	Proceso biosintético compuesto que contiene ácido fólico (GO:0009396)	1.77785	0.02765	0.12985
<i>corA</i>	Transportador de Magnesio/Cobalto CorA	Transporte de iones de cobalto (GO:0006824), transporte de iones de magnesio (GO:0015693)	1.68310	0.03832	0.15115
<i>glgP</i>	Fosforilasa de la Familia Alfa-Glucano	Proceso metabólico de carbohidratos (GO:0005975)	1.67944	0.00816	0.08249
<i>pgi</i>	Glucosa-6-Fosfato Isomerasa	Gluconeogénesis (GO:0006094); Proceso glicolítico (GO:0006096)	1.66140	0.00284	0.06254
<i>groL</i>	Chaperonina GroEL	Replegamiento de proteínas (GO:0042026)	1.64956	0.00826	0.08321
<i>dacB</i>	D-alanil-D-alanina carboxipeptidasa/D-alanil-D-alanina endopeptidasa	Proceso biosintético de peptidoglicano (GO:0009252)	1.62594	0.00872	0.08404
<i>ftsH</i>	Metaloproteasa de zinc dependiente de ATP FtsH	Proteólisis (GO:0006508); división celular (GO:0051301)	1.56737	0.00936	0.08686
<i>nudC</i>	NAD(+) difosfatasa	División (GO:0051301) y ciclo celular (GO:0007049)	1.52112	0.01025	0.08903
<i>hemE</i>	Decarboxilasa uroporfirinógena	Proceso biosintético compuesto que contiene porfirina (GO:0006779)	1.50869	0.00348	0.06570
<i>aroC</i>	Corismato sintasa	Proceso biosintético de la familia de aminoácidos aromáticos (GO:0009073)	1.43957	0.01835	0.11083
<i>lipA</i>	Lipoil sintasa	Proceso biosintético de lipoato (GO:0009107)	1.35685	0.04411	0.16221
<i>pknB</i>	Familia PASTA que contiene dominio Ser/Thr kinasa de la familia Stk1	Fosforilación de proteínas (GO:0006468)	1.21983	0.00502	0.07127

En cuanto a la subexpresión generalizada de la maquinaria de síntesis de proteínas (Tabla X) *rpl-ABEFRW* y *rps-BDKS* (proteínas ribosomales de las subunidades 50S y 30S respectivamente, GO:0006412) con un fc en un rango de 0.45 a 0.064 y *fusA* con 0.35 de cambio, sugieren una desaceleración de los procesos de alto costo energético asociados con el crecimiento y la proliferación, y a una reorientación metabólica para ahorrar energía y recursos en condiciones de estrés del pireno (Shajani *et al.*, 2011; Starosta *et al.*, 2014; Borkowski *et al.*, 2016; Bennison *et al.*, 2021).

La subexpresión de *aceA* (Proceso metabólico del ácido carboxílico, GO:0019752), con un cambio de 0.1 y *ppc* (Fijación de carbono, GO:0015977)

con un fc 0.56 apoyan la idea de una menor demanda de intermediarios para la biosíntesis de biomasa en un escenario de crecimiento lento para ahorro de energía (Lemuth *et al.*, 2008; Ahn *et al.*, 2016; Yang *et al.*, 2022).

Por su parte, los fc de *manA* (Manosa-6-fosfato isomerasa, GO:0005975) y *zwf* (lucosa-6-fosfato deshidrogenasa, GO:0006006, GO:0006098) podrían indicar una reorientación y/o reducción en ciertos flujos de carbono, lo que puede ser consistente con una menor prioridad en el uso de fuentes de carbono complejas (Zhao *et al.*, 2004; Sun & Altenbuchner, 2010; Papagianni, 2012). La amplia gama de genes involucrados en la biosíntesis de diversos aminoácidos y precursores como *asnB* (Asparagina sintasa, GO:0006529), *argH* (Argininosuccinato liasa, GO:0042450), *thrB* (Homoserina quinasa, GO:0009088), *serA* (Fosfoglicerato deshidrogenasa, GO:0006564), *glyA* (Serina hidroximetiltransferasa, GO:0019264), *panD* (Aspartato 1-descarboxilasa, GO:0006523), *dapC* (Succildiaminopimelato transaminasa, GO:0009058), *ilvC* (Cetoácido reductoisomerasa, GO:0009083); apoyan la idea de una reducción en la síntesis de macromoléculas y con ello la ralentización del crecimiento celular.

La subexpresión de *fadD1* (fc 0.45) (Ácido graso-CoA ligasa *FadD1*, GO:0006631) sugiere una disminución en la activación de ácidos grasos para la beta-oxidación, implicando una menor degradación de lípidos como fuente de energía, lo que contrasta con la sobreexpresión de genes de biosíntesis de lípidos. Esto indica una priorización de la síntesis de lípidos estructurales o de almacenamiento sobre su catabolismo completo (Beld *et al.*, 2015; Wang, 2024).

En cuanto a la respuesta a estrés y homeostasis, *trxB* (Tiorredoxina-disulfuro

reductasa, GO:0019430), *nph* (Flavina reductasa dependiente de NADH, GO:0051186), *ribB* (3, 4-dihidroxi-2-butanona-4-fosfato sintasa, GO:0009230), coincidiendo con un cambio de 0.61, contrasta con la sobreexpresión de *soxR* (fc 4.2) (Activador transcripcional sensible a redox SoxR, GO:0006355, GO:0006979) y otras chaperonas. Esto puede indicar una reorganización selectiva de los mecanismos antioxidantes, priorizándose algunos de ellos, mientras que otros se regulan negativamente si no son la estrategia principal (Imlay, 2008).

La subexpresión, con fc 0.47, de *otsB* (Trehalosa-fosfatasa, GO:0005992), sugiere que la bacteria podría estar reevaluando su estrategia de protección osmótica o energética en un ambiente con pireno (Altendorf, K., *et al.*, 2009). Los genes *efeB* (Subunidad deferroquelatasa/peroxidasa del transportador de captación de hierro, GO:0033212), *phoU* (Proteína *PhoU* del complejo de señalización de fosfato, GO:0006817, GO:0030643, GO:0045936), *pstA*, *pstB*, *pstC* (Permeasa *Pst-AC* del transportador ABC de fosfato, Proteína de unión a ATP *PstB* del transportador ABC de fosfato; GO:0006817), con un rango de cambio de 0.20 a 0.50, sugieren una menor demanda de hierro y fósforo, lo cual es coherente con una baja actividad metabólica y por ende, un menor crecimiento celular (Santos-Beneit, 2015; Arosio *et al.*, 2024). Por otro lado, los fc de los genes *secD* y *secF* (Subunidad de translocasa de proteína SecDF, GO:0006886, GO:0043952) de 0.59 y 0.38, respectivamente, indican una reducción en la cantidad de proteínas que se exportan a través de la membrana o que se insertan en ella. Esto concuerda con la subexpresión en los genes que codifican para las

proteínas ribosomales de las subunidades 50S y 30S (Tabla X), lo que conlleva a la posible disminución en la síntesis de proteínas y la desaceleración del crecimiento (Driessen & Nouwen, 2008).

La subexpresión de *benB* (Subunidad pequeña de la benzoato 1,2-dioxigenasa, GO:0008152) con cambio de 0.33 , junto con la falta de sobreexpresión de otras enzimas clave en las vías de degradación de HAP (como *pcaG*, *catA*, *bphC*; Tabla XIV, ANEXO 4), sugiere que *R. ruber* MSA14, en las condiciones y tiempo de exposición de este estudio, no está activando significativamente las vías del catabolismo de hidrocarburos aromáticos, sino que prioriza la tolerancia y la adaptación fisiológica (Guevara *et al.*, 2019; Dangar *et al.*, 2022).

Tabla X. Genes subexpresados en respuesta a pireno. *fc*: tasa de cambio; *pval*: valor de probabilidad; *qval*: *p*-valor ajustado (*qval* > 0.1 pero < 0.2 se consideran aceptables si *fc* < 1 y si *pval* ≤ 0.05).

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	<i>fc</i> < 1	<i>pval</i> ≤ 0.05	<i>qval</i> ≤ 0.1
<i>thrS</i>	Treonina--tARN ligasa	Aminoacilación de treonil-ARNt (GO:0006435)	0.838444	0.028210	0.130482
<i>truA</i>	tRNA pseudouridina(38-40) sintasa TruA	Síntesis de pseudouridina (GO:0001522), modificación del ARN (GO:0009451)	0.779990	0.010084	0.088438
<i>rfbA</i>	RfbA (glucosa-1-fosfato timidililtransferasa)	Proceso biosintético de polisacáridos (GO:0000271), proceso biosintético dTDP-rhamnosa (GO:0019305)	0.688986	0.008611	0.084045
<i>asnB</i>	Asparagina sintasa (glutamina-hidrolizante)	Proceso biosintético de asparagina (GO:0006529)	0.682480	0.002124	0.060428
<i>argH</i>	Argininosuccinato liasa	Proceso biosintético de L-arginina a través de ornitina (GO:0042450)	0.676874	0.011403	0.091961
<i>cobT</i>	Nicotinato-nucleótido--dimetilbencimidazol fosforribosiltransferasa	Proceso biosintético de cobalamina (GO:0009236)	0.639254	0.049765	0.173811
<i>hrpA</i>	ARN helicasa dependiente de ATP HrpA	Modificación del ARN (GO:0016070), metabolismo del ARN (GO:0016070)	0.635069	0.021739	0.118284
<i>hemL</i>	Glutamato-1-semialdehído 2,1-aminomutasa	Proceso biosintético de tetrapirrol (GO:0033014)	0.633554	0.004042	0.066276
<i>argS</i>	Arginina--ARNt ligasa	Aminoacilación de arginil-ARNt (GO:0006420)	0.626818	0.006872	0.076954
<i>trxB</i>	Tiorredoxina-disulfuro reductasa	Eliminación de radicales superóxido (GO:0019430)	0.618014	0.011741	0.092871
<i>ribB</i>	3,4-dihidroxi-2-butanona-4-fosfato sintasa	Biosíntesis de riboflavina (vitamina B2) (GO:0009230)	0.616231	0.038371	0.151157

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc < 1	pval ≤ 0.05	qval ≤ 0.1
glyA	Serina hidroximetiltransferasa	Proceso biosintético de glicina a partir de serina (GO:0019264)	0.612195	0.005847	0.072818
nphA	Flavina reductasa dependiente de NADH	Metabolismo de cofactores (GO:0051186)	0.611749	0.049807	0.173842
ppk2	Polifosfato quinasa 2	Proceso metabólico del fósforo (GO:0006793)	0.598527	0.009487	0.087310
secD	Subunidad de translocasa de proteína SecD	Transporte de proteínas intracelulares (GO:0006886), transporte de proteínas por el complejo Sec (GO:0043952)	0.594668	0.003117	0.064681
kstR	Regulador transcripcional del catabolismo del colesterol KstR	Catabolismo de esteroides (GO:0006695)	0.570201	0.009648	0.087310
ppc	Fosfoenolpiruvato carboxilasa	Ciclo del ácido tricarbóxico (GO:0006099), proceso biológico: fijación de carbono (GO:0015977)	0.569835	0.007830	0.080547
ilvC	Cetoácido reductoisomerasa	Biosíntesis de aminoácidos ramificados (GO:0009083)	0.543719	0.018006	0.110142
serA	Fosfoglicerato deshidrogenasa	Proceso biosintético de L-serina (GO:0006564)	0.528602	0.009496	0.087310
mtr	Micotiona reductasa	Proceso metabólico de micotiol (GO:00010126)	0.517796	0.009415	0.087258
efeB	Subunidad de ferroquelatasa/peroxidasa del transportador de captación de hierro	Importación de hierro en la célula (GO:0033212)	0.505593	0.014363	0.101007
hisB	Imidazolglicerol-fosfato deshidratasa HisB	Proceso biosintético de L-histidina (GO:0000105)	0.499802	0.007008	0.077434
hemB	Porfobilinógeno sintasa	Proceso biosintético de tetrapirrol (GO:0033014)	0.499200	0.000150	0.035603
glbB	Subunidad grande de glutamato sintasa	Proceso biosintético de glutamato (GO:0006537), proceso metabólico (GO:0008152)	0.496139	0.011244	0.091529
argJ	ArgJ (glutamato N-acetiltransferasa/aminoácido acetiltransferasa bifuncional)	Proceso biosintético de L-arginina (GO:0006526)	0.495762	0.002914	0.063082
carA	Subunidad pequeña de la carbamoil-fosfato sintetasa hidrolizante de glutamina	Proceso biosintético de nucleobase de pirimidina "de novo" (GO:0006207), Proceso metabólico de la glutamina (GO:0006541)	0.480522	0.047104	0.168117
obgE	GTPasa ObgE		0.476301	0.009297	0.086476
ybeY	RNasa YbeY de maduración del ARNr	Biogénesis del ribosoma (GO:0042254)	0.475462	0.006441	0.075471
otsB	Trehalosa-fosfatasa	Proceso biosintético de trehalosa (GO:0005992)	0.474429	0.050471	0.174541
nadC	Nicotinato-nucleótido difosforilasa carboxilante	Proceso biosintético NAD (GO:0009435)	0.473349	0.024721	0.124842
phoU	Proteína PhoU del complejo de señalización de fosfato	Transporte de iones fosfato (GO:0006817), homeostasis iónica de fosfato intracelular (GO:0030643), regulación negativa del proceso metabólico de fosfato (GO:0045936)	0.468017	0.012727	0.095967
thrB	Homoserina quinasa	Proceso biosintético de treonina (GO:0009088)	0.463458	0.016228	0.105847
rpsK	Proteína ribosomal 30S S11	Traducción (GO:0006412)	0.461340	0.011458	0.091961
fadD1	Ácido graso--CoA ligasa FadD1	Metabolismo de ácidos grasos (GO:0006631)	0.459191	0.003379	0.065209
rpsB	Proteína ribosomal 30S S2	Traducción (GO:0006412)	0.458197	0.034497	0.144121
ychF	ATPasa YchF regulada por redox	Respuesta al estrés oxidativo (GO:0006979)	0.451662	0.018415	0.110833

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc < 1	pval ≤ 0.05	qval ≤ 0.1
nusG	Proteína de terminación de transcripción/antiterminación NusG	Regulación del alargamiento de la transcripción con plantilla de ADN (GO:0032784)	0.448751	0.002121	0.060428
era	GTPasa Era	Biogénesis de ribosomas (GO:0042254)	0.440468	0.005091	0.071350
leuC	Subunidad grande de la 3-isopropilmalato deshidratasa	Proceso biosintético de L-leucina (GO:0009098)	0.434412	0.037204	0.150087
tsf	Factor de elongación de la traducción Ts	Alargamiento traslacional (GO:0006414)	0.431025	0.009690	0.087310
rlmN	RlmN (adenina(2503)-C(2))-metiltransferasa del ARNr 23S	Metilación de ARNt (GO:0030488), proceso biológico: metilación de base de ARNr (GO:0070475)	0.426866	0.014998	0.102644
rpIA	Proteína ribosomal 50S L1	Traducción (GO:0006412)	0.410166	0.019336	0.113113
pstC	Subunidad permeasa PstC del transportador ABC de fosfato	Transporte de iones fosfato (GO:0006817)	0.399522	0.001461	0.057918
rpIE	Proteína ribosomal 50S L5	Traducción (GO:0006412)	0.396239	0.008862	0.084383
secF	Subunidad de translocasa de proteína SecF	Transporte de proteínas intracelulares (GO:0006886), transporte de proteínas por el complejo Sec (GO:0043952)	0.382881	0.041391	0.156671
panD	Aspartato 1-descarboxilasa	Proceso biosintético de alanina (GO:0006523)	0.361520	0.013659	0.100033
fusA	Factor de elongación G	Elongación traslacional (GO:0006414)	0.355468	0.006335	0.074619
dapC	Succildiaminopimelato transaminasa	Proceso biosintético (GO:0009058)	0.351279	0.002679	0.062541
pyrR	Regulador transcripcional del operón pir bifuncional/uracilo fosforribosiltransferasa PyrR	Proceso metabólico de nucleósidos (GO:0009116)	0.344372	0.026085	0.126262
rpsD	Proteína ribosomal 30S S4	Traducción (GO:0006412)	0.338301	0.009002	0.085095
benB	Subunidad pequeña de la benzoato 1,2-dioxigenasa	Proceso metabólico (GO:0008152)	0.332977	0.011573	0.092256
rpIR	Proteína ribosomal 50S L18	Traducción (GO:0006412)	0.314280	0.047380	0.168469
pstB	Proteína de unión a ATP PstB del transportador ABC de fosfato	Transporte de iones fosfato (GO:0006817)	0.308231	0.008569	0.083893
hisI	Fosforribosil-AMP ciclohrolasa	Proceso biosintético de L-histidina (GO:0000105)	0.284781	0.002658	0.062541
rpIF	Proteína ribosomal 50S L6	Traducción (GO:0006412)	0.283624	0.018088	0.110261
ilvN	Subunidad pequeña de la acetolactato sintasa	Proceso biosintético de aminoácidos de cadena ramificada (GO:0009082)	0.283562	0.006513	0.075611
manA	Manosa-6-fosfato isomerasa, clase I	Proceso metabólico de carbohidratos (GO:0005975)	0.280250	0.010218	0.089031
rpmC	Proteína ribosomal 50S L29	Traducción (GO:0006412)	0.246844	0.005319	0.071350
rpmD	Proteína ribosomal 50S L30	Traducción (GO:0006412)	0.238217	0.024391	0.124842
zwf	Glucosa-6-fosfato deshidrogenasa	Proceso metabólico de la glucosa (GO:0006006), derivación de pentosa-fosfato (GO:0006098)	0.236628	0.013911	0.100553
pstA	Permeasa PstA del transportador ABC de fosfato	Transporte de iones fosfato (GO:0006817)	0.209551	0.004965	0.071170
map	Metionil aminopeptidasa tipo I	Proteólisis (GO:0006508)	0.194640	0.005088	0.071350
rpsS	Proteína ribosomal 30S S19	Traducción (GO:0006412)	0.171770	0.011799	0.093132

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc < 1	pval ≤ 0.05	qval ≤ 0.1
ahpF	Subunidad de alquilo hidropéroxido reductasa F	Respuesta a especies reactivas de oxígeno (GO:0000302)	0.159147	0.002386	0.062541
ahpC	subunidad de la reductasa de hidropéroxido de alquilo C	Respuesta al estrés oxidativo (GO:0006979)	0.125048	0.001943	0.060336
thiM	Hidroxietiltiazol quinasa	Proceso biosintético de tiamina (GO:0009228)	0.118878	0.003890	0.066276
recO	Proteína de reparación de ADN RecO	Reparación del ADN (GO:0006281), recombinación de ADN (GO:0006310)	0.107389	0.000909	0.049245
rplB	Proteína ribosomal 50S L2	Traducción (GO:0006412)	0.103883	0.011149	0.091186
aceA	Isocitrato liasa	Proceso metabólico del ácido carboxílico (GO:0019752)	0.102614	0.006006	0.073302
rplW	Proteína ribosomal 50S L23	Traducción (GO:0006412)	0.064492	0.007039	0.077434
metE	5-metiltetrahidropteroiltri-glutamato--homocisteína S-metiltransferasa	Proceso biosintético de aminoácidos (GO:0008652), proceso biosintético de metionina (GO:0009086)	0.039248	0.003203	0.065209
thrS	Treonina--tARN ligasa	Aminoacilación de treonil-ARNt (GO:0006435)	0.838444	0.028210	0.130482
truA	tRNA pseudouridina(38-40) sintasa TruA	Síntesis de pseudouridina (GO:0001522), modificación del ARN (GO:0009451)	0.779990	0.010084	0.088438
rfaA	RfbA (glucosa-1-fosfato timidililtransferasa)	Proceso biosintético de polisacáridos (GO:0000271), proceso biosintético dTDP-rhamnosa (GO:0019305)	0.688986	0.008611	0.084045
asnB	Asparagina sintasa (glutamina-hidrolizante)	Proceso biosintético de asparagina (GO:0006529)	0.682480	0.002124	0.060428
argH	Argininosuccinato liasa	Proceso biosintético de L-arginina a través de ornitina (GO:0042450)	0.676874	0.011403	0.091961
cobT	Nicotinato-nucleótido--dimetilbencimidazol fosforribosiltransferasa	Proceso biosintético de cobalamina (GO:0009236)	0.639254	0.049765	0.173811
hrpA	ARN helicasa dependiente de ATP HrpA	Modificación del ARN (GO:0016070), metabolismo del ARN (GO:0016070)	0.635069	0.021739	0.118284
hemL	Glutamato-1-semialdehído 2,1-aminomutasa	Proceso biosintético de tetrapirrol (GO:0033014)	0.633554	0.004042	0.066276
argS	Arginina--ARNt ligasa	Aminoacilación de arginil-ARNt (GO:0006420)	0.626818	0.006872	0.076954
trxB	Tiorredoxina-disulfuro reductasa	Eliminación de radicales superóxido (GO:0019430)	0.618014	0.011741	0.092871
ribB	3,4-dihidroxi-2-butanona-4-fosfato sintasa	Biosíntesis de riboflavina (vitamina B2) (GO:0009230)	0.616231	0.038371	0.151157
glyA	Serina hidroximetiltransferasa	Proceso biosintético de glicina a partir de serina (GO:0019264)	0.612195	0.005847	0.072818
nphA	Flavina reductasa dependiente de NADH	Metabolismo de cofactores (GO:0051186)	0.611749	0.049807	0.173842
ppk2	Polifosfato quinasa 2	Proceso metabólico del fósforo (GO:0006793)	0.598527	0.009487	0.087310
secD	Subunidad de translocasa de proteína SecD	Transporte de proteínas intracelulares (GO:0006886), transporte de proteínas por el complejo Sec (GO:0043952)	0.594668	0.003117	0.064681
kstR	Regulador transcripcional del catabolismo del colesterol KstR	Catabolismo de esteroides (GO:0006695)	0.570201	0.009648	0.087310
ppc	Fosfoenolpiruvato carboxilasa	Ciclo del ácido tricarbóxico (GO:0006099), proceso biológico: fijación de carbono (GO:0015977)	0.569835	0.007830	0.080547
ilvC	Cetoácido reductoisomerasa	Biosíntesis de aminoácidos ramificados (GO:0009083)	0.543719	0.018006	0.110142

GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc < 1	pval ≤ 0.05	qval ≤ 0.1
serA	Fosfoglicerato deshidrogenasa	Proceso biosintético de L-serina (GO:0006564)	0.528602	0.009496	0.087310
mtr	Micotiona reductasa	Proceso metabólico de micotiol (GO:00010126)	0.517796	0.009415	0.087258
efeB	Subunidad deferroquelatasa/peroxidasa del transportador de captación de hierro	Importación de hierro en la célula (GO:0033212)	0.505593	0.014363	0.101007
hisB	Imidazolglicerol-fosfato deshidratasa HisB	Proceso biosintético de L-histidina (GO:0000105)	0.499802	0.007008	0.077434
hemB	Porfobilinógeno sintasa	Proceso biosintético de tetrapirrol (GO:0033014)	0.499200	0.000150	0.035603
glbB	Subunidad grande de glutamato sintasa	Proceso biosintético de glutamato (GO:0006537), proceso metabólico (GO:0008152)	0.496139	0.011244	0.091529
argJ	ArgJ (glutamato N-acetiltransferasa/aminoácido acetiltransferasa bifuncional)	Proceso biosintético de L-arginina (GO:0006526)	0.495762	0.002914	0.063082
carA	Subunidad pequeña de la carbamoil-fosfato sintetasa hidrolizante de glutamina	Proceso biosintético de nucleobase de pirimidina "de novo" (GO:0006207), Proceso metabólico de la glutamina (GO:0006541)	0.480522	0.047104	0.168117
obgE	GTPasa ObgE	Biogénesis del ribosoma (GO:0042254)	0.476301	0.009297	0.086476
ybeY	RNasa YbeY de maduración del ARNr	Proceso biosintético de trehalosa (GO:0005992)	0.475462	0.006441	0.075471
otsB	Trehalosa-fosfatasa	Proceso biosintético NAD (GO:0009435)	0.474429	0.050471	0.174541
nadC	Nicotinato-nucleótido difosforilasa carboxilante	Transporte de iones fosfato (GO:0006817), homeostasis iónica de fosfato intracelular (GO:0030643), regulación negativa del proceso metabólico de fosfato (GO:0045936)	0.473349	0.024721	0.124842
phoU	Proteína PhoU del complejo de señalización de fosfato	Proceso biosintético de treonina (GO:0009088)	0.468017	0.012727	0.095967
thrB	Homoserina quinasa	Traducción (GO:0006412)	0.461340	0.011458	0.091961
rpsK	Proteína ribosomal 30S S11	Metabolismo de ácidos grasos (GO:0006631)	0.459191	0.003379	0.065209
fadD1	Ácido graso--CoA ligasa FadD1	Traducción (GO:0006412)	0.458197	0.034497	0.144121
rpsB	Proteína ribosomal 30S S2	Respuesta al estrés oxidativo (GO:0006979)	0.451662	0.018415	0.110833
yhcF	ATPasa YhcF regulada por redox	Regulación del alargamiento de la transcripción con plantilla de ADN (GO:0032784)	0.448751	0.002121	0.060428
nusG	Proteína de terminación de transcripción/antiterminación NusG	Biogénesis de ribosomas (GO:0042254)	0.440468	0.005091	0.071350
era	GTPasa Era	Proceso biosintético de L-leucina (GO:0009098)	0.434412	0.037204	0.150087
leuC	Subunidad grande de la 3-isopropilmalato deshidratasa	Alargamiento traslacional (GO:0006414)	0.431025	0.009690	0.087310
tsf	Factor de elongación de la traducción Ts	Metilación de ARNt (GO:0030488), proceso biológico: metilación de base de ARNr (GO:0070475)	0.426866	0.014998	0.102644
rlmN	RlmN (adenina(2503)-C(2))-metiltransferasa del ARNr 23S	Traducción (GO:0006412)	0.410166	0.019336	0.113113
rpIA	Proteína ribosomal 50S L1	Transporte de iones fosfato (GO:0006817)	0.399522	0.001461	0.057918
pstC	Subunidad permeasa PstC del transportador ABC de fosfato				

<i>rplE</i>	Proteína ribosomal 50S L5	Traducción (GO:0006412)	0.396239	0.008862	0.084383
GEN	PROTEÍNA	ONTOLOGÍA GENICA. AmiGO 2	fc < 1	pval ≤ 0.05	qval ≤ 0.1
<i>secF</i>	Subunidad de translocasa de proteína SecF	Transporte de proteínas intracelulares (GO:0006886), transporte de proteínas por el complejo Sec (GO:0043952)	0.382881	0.041391	0.156671
<i>panD</i>	Aspartato 1-descarboxilasa	Proceso biosintético de alanina (GO:0006523)	0.361520	0.013659	0.100033
<i>fusA</i>	Factor de elongación G	Elongación traslacional (GO:0006414)	0.355468	0.006335	0.074619
<i>dapC</i>	Succildiaminopimelato transaminasa	Proceso biosintético (GO:0009058)	0.351279	0.002679	0.062541
<i>pyrR</i>	Regulador transcripcional del operón pir bifuncional/uracilo fosforribosiltransferasa PyrR	Proceso metabólico de nucleósidos (GO:0009116)	0.344372	0.026085	0.126262
<i>rpsD</i>	Proteína ribosomal 30S S4	Traducción (GO:0006412)	0.338301	0.009002	0.085095
<i>benB</i>	Subunidad pequeña de la benzoato 1,2-dioxigenasa	Proceso metabólico (GO:0008152)	0.332977	0.011573	0.092256
<i>rplR</i>	Proteína ribosomal 50S L18	Traducción (GO:0006412)	0.314280	0.047380	0.168469
<i>pstB</i>	Proteína de unión a ATP PstB del transportador ABC de fosfato	Transporte de iones fosfato (GO:0006817)	0.308231	0.008569	0.083893
<i>hisI</i>	Fosforribosil-AMP ciclohrolasa	Proceso biosintético de L-histidina (GO:0000105)	0.284781	0.002658	0.062541
<i>rplF</i>	Proteína ribosomal 50S L6	Traducción (GO:0006412)	0.283624	0.018088	0.110261
<i>ilvN</i>	Subunidad pequeña de la acetolactato sintasa	Proceso biosintético de aminoácidos de cadena ramificada (GO:0009082)	0.283562	0.006513	0.075611
<i>manA</i>	Manosa-6-fosfato isomerasa, clase I	Proceso metabólico de carbohidratos (GO:0005975)	0.280250	0.010218	0.089031
<i>rpmC</i>	Proteína ribosomal 50S L29	Traducción (GO:0006412)	0.246844	0.005319	0.071350
<i>rpmD</i>	Proteína ribosomal 50S L30	Traducción (GO:0006412)	0.238217	0.024391	0.124842
<i>zwf</i>	Glucosa-6-fosfato deshidrogenasa	Proceso metabólico de la glucosa (GO:0006006), derivación de pentosa-fosfato (GO:0006098)	0.236628	0.013911	0.100553
<i>pstA</i>	Permeasa PstA del transportador ABC de fosfato	Transporte de iones fosfato (GO:0006817)	0.209551	0.004965	0.071170
<i>map</i>	Metionil aminopeptidasa tipo I	Proteólisis (GO:0006508)	0.194640	0.005088	0.071350
<i>rpsS</i>	Proteína ribosomal 30S S19	Traducción (GO:0006412)	0.171770	0.011799	0.093132
<i>ahpF</i>	Subunidad de alquilo hidropéroxido reductasa F	Respuesta a especies reactivas de oxígeno (GO:0000302)	0.159147	0.002386	0.062541
<i>ahpC</i>	subunidad de la reductasa de hidropéroxido de alquilo C	Respuesta al estrés oxidativo (GO:0006979)	0.125048	0.001943	0.060336
<i>thiM</i>	Hidroxiethyliazol quinasa	Proceso biosintético de tiamina (GO:0009228)	0.118878	0.003890	0.066276
<i>recO</i>	Proteína de reparación de ADN RecO	Reparación del ADN (GO:0006281), recombinación de ADN (GO:0006310)	0.107389	0.000909	0.049245
<i>rplB</i>	Proteína ribosomal 50S L2	Traducción (GO:0006412)	0.103883	0.011149	0.091186
<i>aceA</i>	Isocitrato liasa	Proceso metabólico del ácido carboxílico (GO:0019752)	0.102614	0.006006	0.073302
<i>rplW</i>	Proteína ribosomal 50S L23	Traducción (GO:0006412)	0.064492	0.007039	0.077434

<i>metE</i>	5-metiltetrahidropteriltri-glutamato--homocisteína metiltransferasa	S-	Proceso biosintético de aminoácidos (GO:0008652), proceso biosintético de metionina (GO:0009086)	0.039248	0.003203	0.065209
-------------	---	----	--	----------	----------	----------

El análisis de los genes anotados y expresados diferencialmente sugieren que *R. ruber* MSA14, ante el pireno, adapta una estrategia inicial que prioriza la resiliencia y la adaptación fisiológica sobre la degradación rápida del contaminante mediante la reorientación masiva del metabolismo, la fortificación de las barreras celulares, la gestión estratégica de la energía y las reservas lipídicas, así como la activación de robustos sistemas de respuesta al estrés. La ralentización del crecimiento y de la síntesis de macromoléculas refleja un estado de preparación y conservación de recursos, lo que le permite a *R. ruber* MSA14 tolerar la presencia del pireno en etapas tempranas para su eventual metabolización una vez que las condiciones sean más propicias.

Una visión global de cómo el pireno influye en la expresión génica en *Rhodococcus ruber* MSA14 se puede observar en la figura 10. Cada barra representa la respuesta de un gen individual anotado a la exposición al pireno. Los genes con barras rojas son aquellos con $fc > 1$ en respuesta al pireno; cuanto mayor es el valor positivo del $\log_2 fc$ mayor es la sobreexpresión. A su vez, su contraparte, los genes subexpresados se indican en barras azules que se extienden hacia la izquierda del cero; cuanto más negativo es el valor del $\log_2 fc$, mayor es la subexpresión; (se adjunta ANEXO 5 con la magnitud de cambio de cada gen respecto al tratamiento control, corroborando el estado de expresión diferencial).

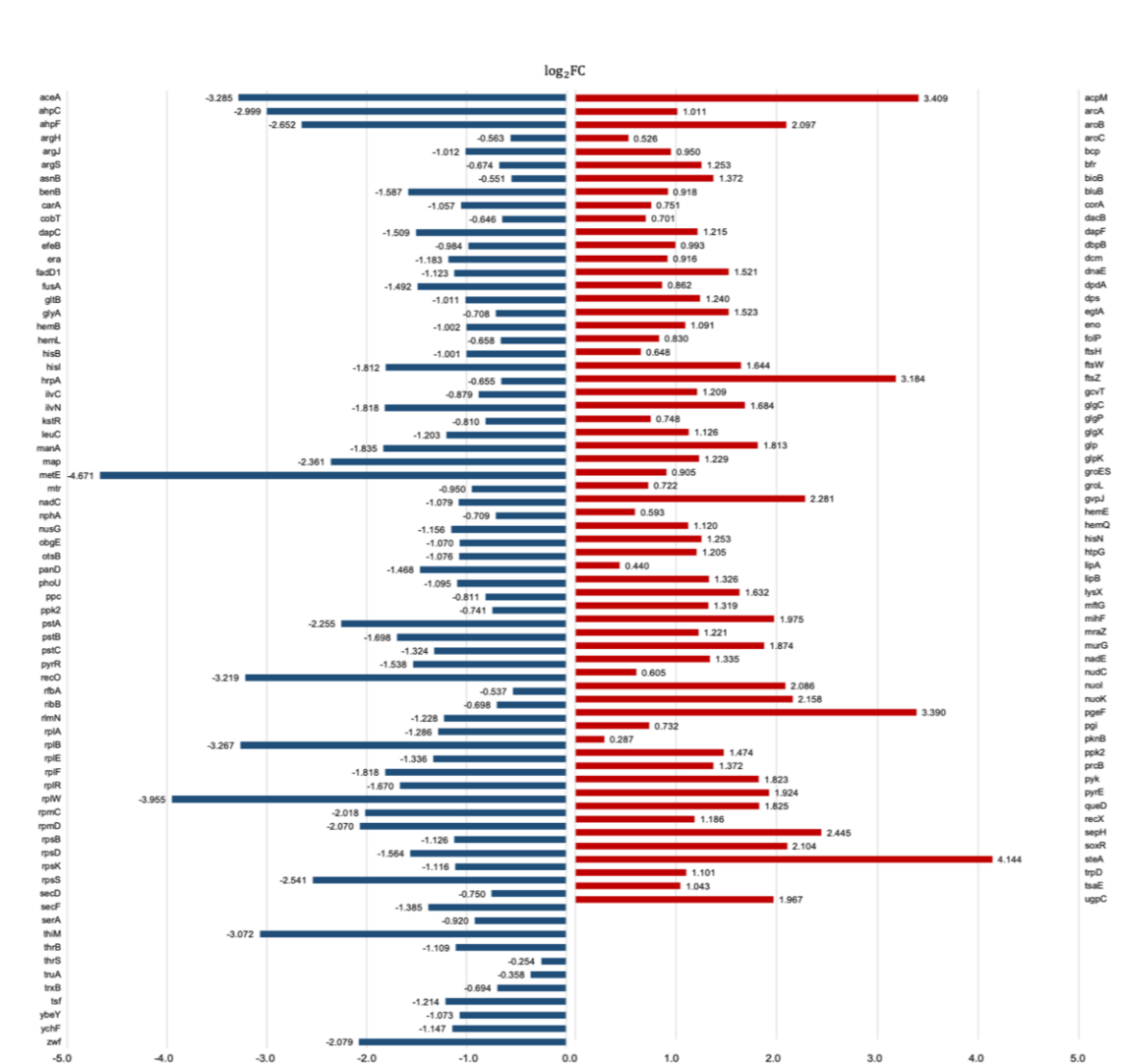


Figura 10. Relación de expresión de GsA-P (azul) y GSA-P (rojo). Eje y: Símbolo de genes, eje x: log₂ fc.

ENRIQUECIMIENTO DE GENES EXPRESADOS DIFERENCIALMENTE

Tras la identificación de los genes diferencialmente expresados (GSA-P y GsA-P) en *Rhodococcus ruber* MSA14 en respuesta al tratamiento con pireno, se llevó a cabo un análisis de enriquecimiento funcional. Este análisis, realizado mediante

las herramientas bioinformáticas DAVID y ShinyGO (versión 0.82), permitió la identificación de categorías funcionales y vías biológicas significativamente asociadas a los cambios en la expresión génica inducidos por el pireno.

DAVID (GO)

Empleando la herramienta DAVID con los módulos Gene Ontology term (GOTERM) y Uniprot Keyword (UP_KW), la configuración de análisis de enriquecimiento funcional incluyó un mínimo de 20 genes por grupo para la concordancia, un umbral de 0.85 (donde al menos dos términos deben compartir el 85 % de sus genes para ser considerados similares) y un umbral de enlace múltiple de 0.50 para fusionar de grupos con una similitud mínima del 50 %. La integración de GOTERM y UP_KW permitió una cobertura con mayor profundidad y amplitud en la anotación y clasificación de los genes con expresión diferencial. GOTERM, al ser una ontología estructurada facilitó la asignación de funciones a genes y productos génicos en diversos niveles de especificidad. Por su parte, UP_KW, como conjunto de palabras clave controladas, proporcionó una visión concisa de los atributos funcionales más relevantes de las proteínas asociadas.

La herramienta GOTERM reveló un 83.80 % de cobertura para procesos biológicos (BP), 86.80 % de composición celular (CC) y con el 94.70 % correspondiente a su función molecular (MF). Mientras que UP_KW, para el dominio BP presentó la mayor cobertura con el 60.30 % respecto a la CC con el 39.40 % y la MF con el 35.10 % (Fig. 11, gráfica de barras). Las barras más largas

indican categorías con una asociación más fuerte y estadísticamente significativa con los genes sobreexpresados en presencia de pireno.

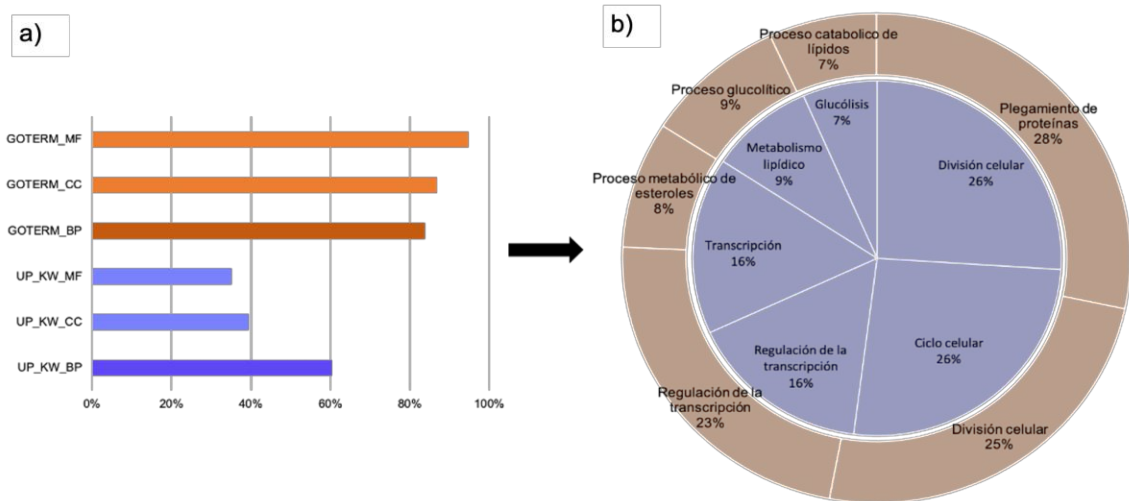


Figura 11. Análisis de enriquecimiento funcional de GSA-P en *R. ruber* MSA14 en respuesta al pireno empleando la base de datos DAVID. La anotación de los procesos biológicos (BP) se muestran en azul usando UP_KW y en naranja usando GOTERM. MF: Función Molecular; BP: Proceso biológico; CC: Composición celular. Gráfico circular: grupos de términos con mayor cobertura funcional relacionados a los GSA-P.

Si bien, ninguno de los procesos biológicos enriquecidos corresponde directamente con el metabolismo de HAP, algunos de ellos, asociados a procesos esenciales y a la supervivencia, podrían estar indirectamente relacionados con este proceso (Fig. 11b). El plegamiento de proteínas, representando el 28% del 83.80% en GOTERM_BP, sugiere la necesidad de replegar o estabilizar la correcta estructura de las proteínas en respuesta al posible estrés inducido por el pireno. La división celular, con el 26% de cobertura, podría indicar una alteración del ciclo celular de *R. ruber* MSA14 para proliferar en presencia del contaminante. Por otro lado, la regulación de la transcripción, proceso que también fue detectado por UP_KB_BP, es significativa para la

adaptación y supervivencia del microorganismo a este hidrocarburo. Una cobertura del 7% fue enfocada para el proceso catabólico de lípidos, mientras que el 8% para el metabolismo de esteroides, lo que sugiere cambios en la síntesis o modificación de lípidos y esteroides que pueden estar relacionados con la integridad de la membrana y/o mecanismos de defensa. En cuanto al gráfico circular en UP_KW_BP (Fig. 11b), sobresale la transcripción y la regulación de la transcripción, ambos con un 16% de representación que infieren un desempeño para la síntesis de ARN. El metabolismo lipídico tiene un 9% de cobertura que sugiere una alteración en las vías de síntesis y/o degradación de lípidos. La sobrerrepresentación de 7% de la glucólisis indica un posible aumento en la vía principal de obtención de energía.

Estos resultados, en conjunto, infieren que la respuesta de *Rhodococcus ruber* MSA14 al tratamiento con pireno implica una reorganización celular integral para sobrevivir a los efectos tóxicos de este compuesto. En el contexto de la degradación de pireno, el enriquecimiento funcional de los genes sobreexpresados sugieren que la bacteria activa mecanismos de respuesta al estrés (plegamiento de proteínas), ajusta su metabolismo energético (glucólisis, metabolismo lipídico), y reprograma su expresión génica (regulación de la transcripción, transcripción) para hacer frente a la presencia del pireno y potencialmente facilitar su metabolismo.

Por otro lado, el enriquecimiento funcional en DAVID de los GsA-P de *Rhodococcus ruber* MSA14 reveló una cobertura del 98%, 97.70% y 55.60% para los terminos GOTERM_BP; GOTERM_FM y GOTERM_CC, respectivamente

(Fig. 12a). En cuanto al módulo UP_KW, la FM representó el dominio con mayor cobertura con el 79.30%, seguido por CC y PB con el 40.10% y 19.10%, respectivamente. En este caso, las bases de datos GOTERM y UP_KW mostraron una consistencia regulatoria de los GsA-P con procesos biológicos significativamente asociados a la biosíntesis de aminoácidos, el catabolismo de hidrocarburos aromáticos (representando el 37% del 19.10% de UP_KW_BP), la diferenciación celular, el procesamiento de ARN, la metilación, el catabolismo del 3-fenilpropionato, el metabolismo de la homocisteína, el proceso metabólico de compuestos que contienen pteridina y el proceso biosintético de metionina.

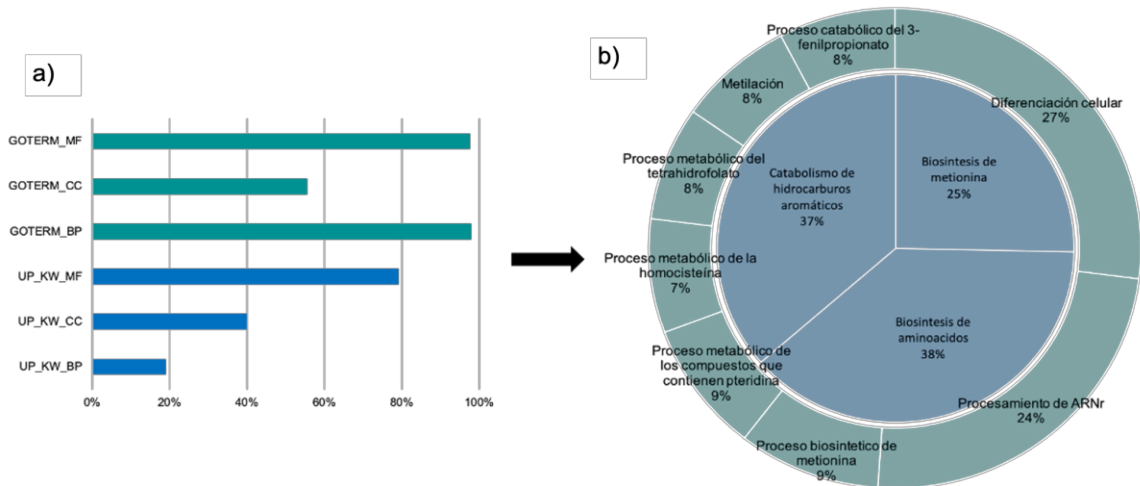


Figura 12. Enriquecimiento funcional de GsA-P empleando la base de datos DAVID. Gráfico de barras: anotación de los procesos biológicos (BP) se muestran en color azul usando UP_KW y en color turquesa usando GOTERM_BP. MF: Función Molecular; BP: Proceso biológico; CC: Composición celular; Gráfico circular: grupos de términos con mayor cobertura funcional relacionados a los GsA-P.

En contraposición a la expectativa de una inducción de vías relacionadas con el catabolismo de pireno en *R. ruber* MSA14 bajo la presencia de este hidrocarburo, la subexpresión significativa de genes involucrados en el catabolismo de

hidrocarburos aromáticos sugirió que en respuesta a las condiciones experimentales empleadas, la bacteria podría estar reprimiendo la expresión de los genes que codifican parte de su maquinaria genética relacionada con la degradación de HAP, como parte de una estrategia de adaptación, a largo plazo, al pireno.

ShinyGO 0.81

Al analizar los genes de interés en la plataforma ShinyGO 0.81, el grupo de GSA-P y GsA-P indicaron un 100% de los IDs emparejados con genes presentes en *Rhodococcus ruber* Chol-4, mientras que el conjunto de genes anotados de fondo proporcionado de *R. ruber* MSA14 permitió la reducción de sesgos para el análisis. Los datos obtenidos para los grupos de GSA-P (A) y GsA-P (B) generaron agrupamientos (Tabla XI) que resumen la correlación entre las 20 vías más significativas del enriquecimiento, considerando la base de datos de vías “Proceso Biológico GO” (BP) de ShinyGO, aplicando un valor de corte predeterminado de FDR (False Discovery Rate) < 0,05 mediante el método Benjamini-Hochberg como filtro para un control estricto, eliminando redundancias de las vías que comparten el 95% de sus genes. Aunado a lo anterior, se consideraron las vías dentro de los límites de tamaño específicos con un mínimo de 10 genes hasta un máximo de 1000 genes.

Tabla XI. Enriquecimiento BP GSA-P (a) y GsA-P (b). FDR bajo indica mayor confianza; Núm. de Genes: cantidad de genes asociados con el término GO o vía KEGG correspondiente

a) DESCRIPCIÓN BP	TÉRMINO	p_value	FDR	GENES
Proceso metabólico de compuestos organonitrogenados	GO:1901564	2.55E-07	0.00046	<i>pgi, arcA, egtA, eno, nudC, ftsH, folP, lipA, lipB, gvT, trpD, dapF, pyrE, hisN, pknB, lysX, hemE, ftsW, bioB, FolP, dacB, FtsW, murG, prcB, aroB, aroC, pyk.</i>
Proceso celular	GO:0009987	1.25E-06	0.0011	<i>pgi, arcA, egtA, eno, nudC, glpK, dnaE, groL, tsaE, bfr, glgC, folP, ugpC, acpM, lipA, lipB, gcvT, trpD, ppk2, dapF, pyrE, corA, glgX, hisN, glgP, bcp, pknB, lysX, dps, GroL, htpG, nuoK, hemE, ftsW, GlgX, bioB, FolP, dacB, FtsW, murG, ftsZ, prcB, CorA, aroB, aroC, pyk.</i>
Proceso metabólico de los compuestos nitrogenados	GO:0006807	4.13E-06	0.0025	<i>pgi, arcA, egtA, eno, nudC, dnaE, tsaE, ftsH, folP, lipA, lipB, gcvT, trpD, nadE, dapF, pyrE, hisN, pknB, lysX, hemE, ftsW, bioB, FolP, dacB, FtsW, murG, prcB, aroB, aroC, pyk.</i>
Proceso metabólico del glucógeno	GO:0005977	1.05E-05	0.0047	<i>glgC, glgX, glgP, GlgX.</i>
Proceso catabólico de los carbohidratos	GO:0016052	1.97E-05	0.006	<i>pgi, eno, glpK, glgX, GlgX, pyk.</i>
Proceso metabólico de los carbohidratos	GO:0005975	0.00017	0.0334	<i>pgi, eno, glpK, glgC, glgX, hisN, glgP, GlgX, murG, pyk.</i>
Proceso biosintético de compuestos cíclicos orgánicos	GO:1901362	0.00018	0.0334	<i>nudC, dnaE, folP, lipA, trpD, pyrE, hisN, hemE, bioB, FolP, aroB, aroC.</i>
Proceso biosintético del ácido carboxílico	GO:0046394	0.00022	0.0362	<i>egtA, folP, acpM, lipA, trpD, dapF, hisN, bioB, FolP, aroB, aroC.</i>
Proceso metabólico del ácido carboxílico	GO:0019752	0.00027	0.0378	<i>pgi, arcA, egtA, eno, folP, acpM, lipA, gcvT, trpD, dapF, hisN, lysX, bioB, FolP, aroB, aroC, pyk.</i>
Replegamiento de proteínas	GO:0042026	0.00029	0.0378	<i>groL, GroL.</i>
Proceso biosintético de moléculas pequeñas	GO:0044283	3.00E-04	0.0378	<i>pgi, egtA, nudC, folP, acpM, lipA, trpD, dapF, hisN, bioB, FolP, aroB, aroC.</i>
Proceso catabólico de sustancias orgánicas	GO:1901575	3.00E-04	0.0378	<i>pgi, arcA, eno, nudC, glpK, ftsH, gcvT, glgX, hisN, GlgX, prcB, pyk.</i>
Proceso biosintético de compuestos aromáticos	GO:0019438	0.00043	0.0408	<i>nudC, dnaE, folP, trpD, pyrE, hisN, hemE, FolP, aroB, aroC.</i>
b) DESCRIPCIÓN BP	TÉRMINO	p_value	FDR	GENES
Proceso biosintético de compuestos organonitrogenados	GO:1901566	2.21E-17	3.99E-14	<i>hemL, rpsD, rpsK, rpmD, rplR, fusA, rplA, HemL, hemB, glyA, serA, leuC, thiM, cobT, asnB, ilvC, rpsB, tsf, gltB, metE, thrS, nadC, hisB, hisI, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ, panD, argS, thrB, carA, ThrS.</i>
Expresión genética	GO:0010467	1.57E-14	1.42E-11	<i>truA, rpsD, rpsK, map, rpmD, rplR, fusA, rplA, nusG, ybeY, rpsB, tsf, rlmN, Map, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, argS, ThrS.</i>
Proceso metabólico de los compuestos organonitrogenados	GO:1901564	5.69E-14	3.44E-11	<i>hemL, rpsD, rpsK, map, rpmD, rplR, fusA, rplA, HemL, hemB, glyA, serA, leuC, ybeY, thiM, cobT, asnB, ilvC, rpsB, tsf, Map, gltB, metE, thrS, nadC, hisB, hisI, rplF, rplE, rpmC, rpsS, rplB, rplW, argH,</i>

argJ, panD, argS, thrB, carA, pyrR, ThrS.

b) DESCRIPCIÓN BP	TÉRMINO	p_value	FDR	GENES
Proceso metabólico celular	GO:0044237	6.79E-13	3.07E-10	<i>hemL, truA, rpsD, rpsK, map, rpmD, rplR, nphA, benB, fusA, rplA, nusG, HemL, hemB, zwf, glyA, serA, leuC, ybeY, recO, thiM, cobT, asnB, ilvC, ppk2, rpsB, tsf, rlmN, Map, mtr, ahpF, hrpA, otsB, trxB, gltB, OtsB, metE, aceA, thrS, nadC, hisB, hisl, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ, panD, argS, thrB, ppc, carA, pyrR, ThrS. hemL, truA, rpsD, rpsK, map, rpmD, rplR, nphA, benB, fusA, rplA, nusG, HemL, hemB, zwf, glyA, serA, leuC, ybeY, era, recO, thiM, cobT, asnB, ilvC, ppk2, rpsB, tsf, rlmN, Map, mtr, ahpF, hrpA, phoU, pstB, pstA, pstC, otsB, trxB, gltB, OtsB, metE, aceA, secF, secD, thrS, nadC, hisB, hisl, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ, panD, argS, thrB, ppc, carA, pyrR, ThrS.</i>
Proceso celular	GO:0009987	2.24E-12	8.12E-10	<i>hemL, truA, rpsD, rpsK, map, rpmD, rplR, fusA, rplA, nusG, HemL, hemB, glyA, serA, leuC, ybeY, recO, thiM, cobT, asnB, ilvC, rpsB, tsf, rlmN, Map, hrpA, gltB, metE, thrS, nadC, hisB, hisl, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ, panD, argS, thrB, carA, pyrR, ThrS.</i>
Proceso metabólico de los compuestos nitrogenados	GO:0006807	5.21E-12	1.57E-09	<i>hemL, truA, rpsD, rpsK, map, rpmD, rplR, fusA, rplA, nusG, HemL, hemB, glyA, serA, leuC, ybeY, recO, thiM, cobT, asnB, ilvC, rpsB, tsf, rlmN, Map, hrpA, gltB, metE, thrS, nadC, hisB, hisl, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ, panD, argS, thrB, carA, pyrR, ThrS.</i>
Traducción	GO:0006412	6.48E-12	1.67E-09	<i>rpsD, rpsK, rpmD, rplR, fusA, rplA, rpsB, tsf, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, argS, ThrS.</i>
Proceso biosintético de amidas	GO:0043604	8.68E-12	1.96E-09	<i>rpsD, rpsK, rpmD, rplR, fusA, rplA, asnB, rpsB, tsf, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, panD, argS, ThrS.</i>
Proceso metabólico de la amida celular	GO:0043603	1.52E-11	3.02E-09	<i>rpsD, rpsK, rpmD, rplR, fusA, rplA, glyA, asnB, rpsB, tsf, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, panD, argS, ThrS.</i>
Proceso biosintético celular	GO:0044249	3.87E-11	5.83E-09	<i>hemL, rpsD, rpsK, rpmD, rplR, fusA, rplA, nusG, HemL, hemB, glyA, serA, leuC, thiM, cobT, asnB, ilvC, rpsB, tsf, gltB, OtsB, metE, thrS, nadC, hisB, hisl, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ, panD, argS, thrB, carA, ThrS.</i>
Proceso biosintético de sustancias orgánicas	GO:1901576	4.91E-11	6.83E-09	<i>hemL, rpsD, rpsK, rpmD, rplR, fusA, rplA, nusG, HemL, hemB, glyA, serA, leuC, thiM, cobT, asnB, ilvC, rpsB, tsf, gltB, OtsB, metE, thrS, nadC, hisB, hisl, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ, panD, argS, thrB, carA, ThrS.</i>
Proceso biosintético	GO:0009058	8.99E-11	1.16E-08	<i>hemL, rpsD, rpsK, rpmD, rplR, fusA, rplA, nusG, HemL, hemB, glyA, serA, leuC, thiM, cobT, asnB, ilvC, rpsB, tsf, gltB, OtsB, metE, thrS, nadC, hisB, hisl, rplF, rplE, rpmC, rpsS, rplB, rplW, argH, argJ,</i>

panD, argS, thrB, carA, ThrS, dapC.

b) DESCRIPCIÓN BP	TÉRMINO	p_value	FDR	GENES
Proceso metabólico de los aminoácidos celulares	GO:0006520	1.17E-10	1.41E-08	<i>map, glyA, serA, leuC, asnB, ilvC, Map, gltB, metE, thrS, hisB, hisI, argH, argJ, panD, argS, thrB, carA, ThrS.</i>
Proceso biosintético de compuestos nitrogenados celulares	GO:0044271	1.23E-10	1.41E-08	<i>hemL, rpsD, rpsK, rpmD, rplR, fusA, rplA, nusG, HemL, hemB, thiM, cobT, asnB, rpsB, tsf, thrS, nadC, rplF, rplE, rpmC, rpsS, rplB, rplW, panD, argS, carA, ThrS.</i>
Proceso metabólico de los compuestos nitrogenados celulares	GO:0034641	2.37E-10	2.52E-08	<i>hemL, truA, rpsD, rpsK, rpmD, rplR, fusA, rplA, nusG, HemL, hemB, glyA, ybeY, recO, thiM, cobT, rpsB, tsf, rlmN, hrpA, thrS, nadC, hisB, hisI, rplF, rplE, rpmC, rpsS, rplB, rplW, panD, argS, carA, pyrR, ThrS.</i>
Proceso metabólico de sustancias orgánicas	GO:0071704	3.99E-10	4.01E-08	<i>hemL, truA, rpsD, rpsK, map, rpmD, rplR, nphA, fusA, rplA, nusG, HemL, hemB, zwf, glyA, serA, leuC, ybeY, recO, thiM, cobT, asnB, ilvC, rpsB, tsf, rlmN, Map, mtr, hrpA, otsB, gltB, OtsB, metE, thrS, nadC, hisB, hisI, rplF, rplE, rpmC, rpsS, rplB, rplW, manA, argH, argJ, panD, argS, thrB, ppc, carA, pyrR, ThrS.</i>
Proceso metabólico primario	GO:0044238	9.60E-10	9.14E-08	<i>truA, rpsD, rpsK, map, rpmD, rplR, fusA, rplA, nusG, zwf, glyA, serA, leuC, ybeY, recO, cobT, asnB, ilvC, rpsB, tsf, rlmN, Map, hrpA, otsB, gltB, OtsB, metE, aceA, thrS, nadC, hisB, hisI, rplF, rplE, rpmC, rpsS, rplB, rplW, manA, argH, argJ, panD, argS, thrB, ppc, carA, pyrR, ThrS.</i>
Proceso metabólico de los alfa-aminoácidos	GO:1901605	1.10E-09	9.93E-08	<i>map, glyA, serA, leuC, asnB, ilvC, Map, gltB, metE, hisB, hisI, argH, argJ, panD, thrB, carA.</i>
Proceso metabólico de las proteínas celulares	GO:0044267	2.04E-09	1.76E-07	<i>rpsD, rpsK, map, rpmD, rplR, fusA, rplA, rpsB, tsf, Map, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, argS, ThrS.</i>
Proceso biosintético de alfa-aminoácidos	GO:1901607	4.80E-09	3.95E-07	<i>glyA, serA, leuC, asnB, ilvC, gltB, metE, hisB, hisI, argH, argJ, panD, thrB, carA.</i>
Proceso metabólico	GO:0008152	2.50E-08	1.97E-06	<i>hemL, truA, rpsD, rpsK, map, rpmD, rplR, nphA, benB, fusA, rplA, nusG, HemL, hemB, zwf, glyA, serA, leuC, ybeY, recO, thiM, cobT, asnB, ilvC, ppk2, rpsB, tsf, rlmN, Map, mtr, ahpF, ahpC, hrpA, otsB, trxB, gltB, OtsB, metE, aceA, thrS, nadC, hisB, hisI, rplF, rplE, rpmC, rpsS, rplB, rplW, manA, argH, argJ, panD, argS, thrB, ppc, carA, pyrR, ThrS, dapC.</i>
Proceso metabólico de las proteínas	GO:0019538	3.93E-08	2.84E-06	<i>rpsD, rpsK, map, rpmD, rplR, fusA, rplA, ybeY, rpsB, tsf, Map, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, argS, ThrS.</i>
Proceso metabólico del ácido carboxílico	GO:0019752	5.30E-08	3.69E-06	<i>map, nphA, glyA, serA, leuC, asnB, ilvC, Map, gltB, metE, aceA, thrS, hisB, hisI, argH, argJ, panD, argS, thrB, ppc, carA, ThrS.</i>

b) DESCRIPCIÓN BP	TÉRMINO	p_value	FDR	GENES
Proceso biosintético de macromoléculas celulares	GO:0034645	1.91E-07	1.19E-05	<i>rpsD, rpsK, rpmD, rplR, fusA, rplA, nusG, rpsB, tsf, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, argS, ThrS.</i>
Proceso metabólico de moléculas pequeñas	GO:0044281	1.84E-06	0.00011	<i>map, nphA, zwf, glyA, serA, leuC, thiM, cobT, asnB, ilvC, Map, gltB, metE, aceA, thrS, nadC, hisB, hisI, argH, argJ, panD, argS, thrB, ppc, carA, pyrR, ThrS.</i>
Proceso biosintético de moléculas pequeñas	GO:0044283	1.31E-05	7.00E-04	<i>glyA, serA, leuC, thiM, cobT, asnB, ilvC, gltB, metE, hisB, hisI, argH, argJ, panD, thrB, carA.</i>
Proceso metabólico de las macromoléculas celulares	GO:0044260	2.82E-05	0.0015	<i>rpsD, rpsK, map, rpmD, rplR, fusA, rplA, nusG, recO, rpsB, tsf, rlmN, Map, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, argS, ThrS.</i>
Transporte de iones fosfato	GO:0006817	3.13E-05	0.0016	<i>phoU, pstB, pstA, pstC.</i>
Proceso metabólico de las macromoléculas	GO:0043170	5.43E-05	0.0027	<i>truA, rpsD, rpsK, map, rpmD, rplR, fusA, rplA, nusG, ybeY, recO, rpsB, tsf, rlmN, Map, hrpA, thrS, rplF, rplE, rpmC, rpsS, rplB, rplW, argS, ThrS</i>
La eliminación de metionina, un iniciador de proteínas, está implicada en la maduración de las proteínas.	GO:0035551	6.46E-05	0.0031	<i>map.</i>

Entre los GSA-P, los términos GO significativamente enriquecidos (menor valor FDR) sugieren una respuesta metabólica activa en presencia de pireno como, por ejemplo, el proceso metabólico de los compuestos nitrogenados (GO:0006807), indicando una posible alteración en el metabolismo de moléculas como aminoácidos y bases nitrogenadas, lo que podría reflejar la necesidad de sintetizar proteínas específicas de respuesta al estrés y/o para reparar daños. O bien el proceso celular (GO:0009987), un término muy general infiere que una amplia gama de procesos celulares se ven afectados, como el proceso metabólico de carbohidratos (GO:0005975) y el proceso metabólico del ácido carboxílico (GO:0019752) que sugieren ajustes en el metabolismo energético y de intermediarios centrales. Finalmente, el proceso biosintético de compuestos

cíclicos orgánicos (GO:1901362) y el proceso biosintético de compuestos aromáticos (GO:0019438) deducen estrategias de defensa y de manejo de recursos cuando la degradación de pireno no es la prioridad inmediata.

Por otro lado, los términos GO significativamente enriquecidos en el conjunto de GsA-P apuntan a una posible represión de procesos biosintéticos y metabólicos fundamentales como los términos centrales GO:0010467 y GO:0006412 correspondientes a la expresión genética y traducción, respectivamente, que sugieren una disminución general en la producción de proteínas, lo que podría ser una respuesta de ahorro de energía o una reorientación de los recursos celulares hacia la respuesta al estrés.

A su vez, el proceso metabólico celular (GO:0044237) y el proceso celular (GO:0009987), también significativo en el conjunto sobreexpresado, resalta la magnitud del impacto del pireno en la célula. La represión generalizada de procesos relacionados con aminoácidos y proteínas (GO:0043604, GO:0043603, GO:0006520, GO:1901605, GO:0044267, GO:1901607, GO:0019538) sugiere una disminución en la síntesis de componentes celulares fundamentales.

En cuanto a los términos GO:0044249 (proceso biosintético celular), GO:1901576 (proceso biosintético de sustancias orgánicas), GO:0009058 (proceso biosintético), GO:0044271 (proceso biosintético de compuestos nitrogenados celulares), GO:0034641 (proceso metabólico de los compuestos nitrogenados celulares), GO:0071704 (proceso metabólico de sustancias orgánicas), GO:0044238 (proceso metabólico primario), GO:0034645 (proceso biosintético de macromoléculas celulares), GO:0044281 (proceso metabólico de

moléculas pequeñas), GO:0044283 (proceso biosintético de moléculas pequeñas, también significativo en el conjunto sobreexpresado), GO:0044260 (proceso metabólico de las macromoléculas celulares) y GO:0043170 (proceso metabólico de las macromoléculas) sugieren que la supresión significativa de estas vías biosintéticas y metabólicas esenciales permiten el ahorro de energía bajo la presencia del pireno en el medio. Por su parte el proceso metabólico del ácido carboxílico (también significativo en el conjunto sobreexpresado GO:0019752) sugiere una modulación precisa y diferenciada de las diversas rutas catabólicas y anabólicas (síntesis, degradación o modificación de diversas biomacromoléculas) en las que participa esta ruta anfibólica del metabolismo básico.

El análisis infiere una gran versatilidad metabólica en *R. ruber* MSA14 que le permite adaptarse a condiciones ambientales y nutricionales cambiantes para mantener funciones metabólicas esenciales. Por otro lado, la subexpresión generalizada de procesos biosintéticos y metabólicos primarios podrían indicar una re-priorización de los recursos celulares hacia la detoxificación a expensas del crecimiento y el metabolismo basal. Por lo tanto, estos resultados sugieren que *Rhodococcus ruber* MSA14 activa vías metabólicas para manejar el estrés asociado y desarrollarse en presencia del hidrocarburo. Simultáneamente, la represión significativa de procesos biosintéticos fundamentales infiere una estrategia de ahorro de energía y recursos como consecuencia del estrés causado por este compuesto tóxico. Por otro lado, la ausencia de un enriquecimiento altamente significativo de términos directamente relacionados

con la degradación de hidrocarburos aromáticos en los genes sobreexpresados infiere que la respuesta inicial se centra más en el metabolismo general y la respuesta al estrés antes de una inducción masiva de las vías catabólicas de pireno.

Mediante la base de datos KEEG (Kyoto Encyclopedia of Genes and Genomes) en ShinyGO, se asoció a los conjuntos de GSA-P y GsA-P en *Rhodococcus ruber* MSA14 vías y procesos biológicos (Tabla XII) significativamente enriquecidos que sugieren respuestas metabólicas activas (mapa01100) que convergen en una posible alteración en las reservas de carbono y las fuentes de energía (mapa00500) para regular genes específicos en presencia del contaminante (mapa03018) que permitan su defensa y/o adaptación al estrés (mapa01110) regulando la demanda energética de la respuesta al pireno.

Si bien, en ninguno de los conjuntos de genes se observó un enriquecimiento directo de una vía KEEG específica y altamente significativa para la degradación de pireno.

Tabla XII. Genes agrupados por categorías funcionales definidas por términos GO de alto nivel

	FDR	p_value	Términos o vías GO	Descripción	Genes
GSA-P	4.49E-06	3.87E-08	mapa01100	Vías metabólicas	<i>pgi, arcA, egtA, eno, nudC, glpK, glgC, folP, acpM, lipA, lipB, gcvT, trpD, nadE, bluB, dapF, pyrE, glgX, hisN, glgP, nuoK, nuol, hemE, GlgX, bioB, FolP, murG, aroB, aroC, pyk.</i>
	0.00021	3.61E-06	mapa03018	Degradación del ARN	<i>eno, groL, ppk2, GroL.</i>
	0.0012	3.08E-05	mapa01110	Biosíntesis de metabolitos secundarios	<i>pgi, arcA, eno, glgC, acpM, gcvT, trpD, dapF, glgX, hisN, glgP, hemE, GlgX, aroB, aroC, pyk.</i>
	0.0022	7.51E-05	mapa00500	Metabolismo del almidón y la sacarosa	<i>pgi, glgC, glgX, glgP, GlgX.</i>
	0.0237	0.001	mapa01230	Biosíntesis de aminoácidos	<i>eno, trpD, dapF, hisN, aroB, aroC, pyk.</i>
	0.0294	0.0015	mapa00785	Metabolismo del ácido lipoico	<i>lipA, lipB.</i>

	FDR	p_value	Términos o vías GO	Descripción	Genes
GsA-P	2.56E-08	2.21E-10	mapa03010	Ribosoma	<i>rpsD, rpsK, rpmD, rplR, rplA, rpsB, rplF, rplE, rpmC, rpsS, rplB, rplW.</i>
	5.09E-07	8.77E-09	mapa01230	Biosíntesis de aminoácidos	<i>glyA, serA, leuC, asnB, ilvC, gltB, metE, hisB, hisI, argH, argJ, thrB, dapC.</i>
	1.76E-05	4.54E-07	mapa01110	Biosíntesis de metabolitos secundarios	<i>hemL, hemB, zwf, glyA, serA, leuC, asnB, ilvC, gltB, OtsB, metE, aceA, hisB, hisI, manA, argH, argJ, panD, thrB.</i>
	5.63E-05	1.94E-06	mapa01100	Vías metabólicas	<i>hemL, nphA, benB, HemL, hemB, zwf, glyA, serA, leuC, thiM, cobT, asnB, ilvC, gltB, OtsB, metE, aceA, nadC, hisB, hisI, manA, argH, argJ, panD, thrB, ppc, carA, pyrR, dapC.</i>
	0.0013	5.41E-05	mapa01120	Metabolismo microbiano en diversos ambientes	<i>hemL, nphA, benB, hemB, zwf, glyA, serA, gltB, aceA, thrB, ppc, dapC.</i>
	0.0035	0.00018	mapa00250	Metabolismo de la alanina, aspartato y glutamato	<i>asnB, gltB, argH, carA.</i>

DISCUSIÓN

ESTRATEGIAS DE ADAPTACIÓN DE *R. ruber* AL PIRENO A NIVEL TRANSCRIPCIONAL.

La exposición bacteriana a hidrocarburos ha favorecido el desarrollo de mecanismos adaptativos que les permiten sobrevivir y crecer en presencia de diferentes tipos de estrés ambiental (Lăzăroaie, 2010). Estos mecanismos incluyen modificaciones en la estructura de la membrana o alteraciones en el estado energético general, cambios en la morfología celular y en las propiedades de la superficie celular (Eberlein *et al.*, 2018). Estos mecanismos, a su vez, están regulados por la heterogeneidad poblacional y la regulación transcripcional, factores que les permiten tolerar el estrés a largo plazo (Booth, 2002; Tropis *et al.*, 2005; Bhamidi *et al.*, 2008). Sin embargo, frente a cambios ambientales repentinos, las bacterias activan respuestas a corto plazo para mantener la viabilidad celular ante la exposición a compuestos orgánicos tóxicos u otros factores de estrés ambiental. Entre estas respuestas se incluye una rápida isomerización de ácidos grasos insaturados que genera rigidez de la membrana celular o la liberación de vesículas de membrana externa, que provoca un rápido aumento de la hidrofobicidad de la superficie celular. La respuesta inmediata permite a las bacterias mantenerse fisiológicamente activas y emplear respuestas a largo plazo para asegurar su viabilidad ante condiciones cambiantes (Eberlein *et al.*, 2018).

Entre las bacterias Gram-positivas hidrocarbonoclastas, se ha reportado al género *Rhodococcus* por su capacidad para catabolizar eficazmente compuestos xenobióticos y prosperar en ambientes estresantes (Kanaly and Harayama, 2000; Alvarez, 2003; McLeod *et al.*, 2006; Martínková *et al.*, 2009; Yang *et al.*, 2011; Zhang *et al.*, 2012, Zampolli *et al.*, 2020). Un ejemplo notable es la cepa marina *R. ruber* MSA14, aislada a partir de sedimentos marinos contaminados con pireno (Silva-Jiménez *et al.*, 2018; Embarcadero-Jiménez *et al.*, 2024).

En este sentido, la exposición a corto plazo de *R. ruber* MSA14 (seis horas) a pireno, utilizado como única fuente de carbono y energía, comparado con los 30 días de exposición probado por Embarcadero-Jiménez *et al.* (2024), se reflejó en la expresión diferencial de mecanismos relacionados con el metabolismo central, la biosíntesis de moléculas esenciales (aminoácidos, cofactores y nucleótidos), el metabolismo de lípidos y la dinámica de las membranas, así como el metabolismo de ácidos nucleicos. Estos mecanismos aseguran la supervivencia hasta que los mecanismos dependientes del crecimiento a pireno completan la respuesta adaptativa (Hernández *et al.*, 2008).

Para preservar sus funciones, las células bacterianas adaptan la composición de su membrana, principalmente mediante la modificación de su contenido lipídico, así como de la proporción de proteínas, esteroides, hopanoides y carotenoides (Sardesai y Bhosle, 2002; de Carvalho, 2010; Yam *et al.*, 2010; Álvarez, 2010; Nuñez *et al.*, 2014; Guevara *et al.*, 2019). Estos cambios estructurales permiten a las células sobrevivir y adaptarse metabólicamente a condiciones ambientales fluctuantes.

El análisis transcriptómico de *R. ruber* MSA14 en presencia de pireno, reflejó una adaptación de protección y supervivencia con la sobreexpresión del conjunto de GSA-P asociados principalmente con procesos metabólicos fundamentales, incluyendo el metabolismo de compuestos organonitrogenados y nitrogenados, procesos celulares generales, metabolismo del glucógeno y catabolismo de carbohidratos, sugiriendo una reorientación del flujo de carbono hacia la producción de energía y de metabolitos esenciales para la adaptación y el crecimiento en condiciones adversas (Hernández *et al.*, 2008). Para asegurar su supervivencia, la acumulación de lípidos de almacenamiento por actinomicetos, como los miembros de *Mycobacterium*, *Rhodococcus*, *Nocardia* y *Streptomyces* (Alvarez, 2006) pueden ser utilizados como fuentes endógenas de carbono y donantes de electrones durante períodos de escasez nutricional (Hernández *et al.*, 2008), aunado a que son cruciales en la tolerancia bacteriana a compuestos orgánicos tóxicos (Kuyukina *et al.*, 2000; Sokolovska *et al.*, 2003; Alvarez *et al.*, 2004; Nguyen *et al.*, 2005).

Los miembros de este género producen cantidades variables de triacilgliceroles (TAG) durante el crecimiento en diferentes fuentes de carbono, llegando a acumular niveles muy altos en sus células (McLeod, 2006). *Rhodococcus jostii* RHA1 es considerado un organismo modelo (Hernández *et al.*, 2008) para comprender la genética y la fisiología del metabolismo de compuestos de almacenamiento dada su versatilidad metabólica (McLeod, 2006). La acumulación de TAG en *Rhodococcus* sugiere que estos sirven como fuentes endógenas de carbono y energía durante periodos de inanición, posiblemente

como forma de almacenamiento de energía para la latencia a largo plazo (Alvarez y Steinbüchel, 2002). En *R. ruber* MSA14 la sobreexpresión de *lipA* y *lipB*, genes relacionados con el metabolismo del ácido lipoico de los triacilglicerolos (TAG), se podría inferir que, bajo condiciones adversas provocadas por el pireno en etapas tempranas de exposición, *R. ruber* MSA14 biosintetiza TGA como material de reserva energética para activar posiblemente un periodo de latencia.

Otros compuestos de almacenamiento producidos por bacterias, además de los lípidos, son los polifosfatos y carbohidratos como la trehalosa o el glucógeno, este último generalmente considerado una fuente de energía y carbono almacenada. A menudo, la producción de compuestos de almacenamiento ocurre durante la fase estacionaria, aunque algunas especies bacterianas sintetizan glucógeno principalmente durante la fase de crecimiento exponencial en un ambiente de estrés hiperosmótico (Seibold & Eikmanns, 2007; Hernández *et al.*, 2008).

Los genes clave que codifican enzimas involucradas en el catabolismo del glucógeno en varias bacterias incluyen *glgC* (Martínková *et al.*, 2009; Maarel Van Der *et al.*, 2003), *glgB* (McLeod *et al.*, 2006), *glgX* (Alvarez, 2006), *glgP* (Alonso *et al.*, 2006); *glgA* (Alvarez y Steinbüchel, 2002) y *glgE* (Kalscheuer *et al.*, 2010) de los cuales, en *R. ruber* MSA14 se identificaron los genes *glgP*, *glgX* y *glgC* sobreexpresados, lo que sugiere que este microorganismo obtiene energía a partir del glucógeno en condiciones de estrés provocado por el pireno.

Por otro lado, la acumulación de polifosfatos (poliP) se ha reportado previamente para *R. opacus* PD630, que es un microorganismo taxonómicamente relacionado

con *R. jostii* RHA1 (Jahid *et al.*, 2006; Fraley *et al.*, 2007; Hernández *et al.*, 2008). El poliP puede tener múltiples funciones en las células bacterianas, mejorando potencialmente su capacidad para responder al estrés oxidativo, choque térmico, estrés osmótico, desecación y entornos con bajo contenido de fosfato. El poliP también puede servir como fuente de energía para reemplazar el ATP y puede estar involucrado en la regulación de las actividades enzimáticas. En ese sentido, *R. ruber* MSA14 es una cepa marina que se aisló de una zona impactada por HAP (Silva-Jiménez *et al.*, 2018). Las condiciones de ambientes contaminados con compuestos orgánicos tóxicos suelen ser oligotróficos, y esto pudo promover que este microorganismo desarrollara adaptaciones para el almacenamiento de poliP para poder obtener energía. El crecimiento en medio mínimo de sales suplementado con pireno podría estar replicando una condición oligotrófica dando lugar a la sobreexpresión del gen *ppk2*, lo que permitiría a *R. ruber* MSA14 utilizar poliP como fuente de energía.

La capacidad de las bacterias, para sobrevivir en condiciones estresantes se refleja en la expresión de genes especializados de mediación del estrés (Imlay, 2003) como *dps*, *bfr*, *soxR*, *groES* y *groL*, los cuales fueron identificados en *R. ruber* MSA14. En *Escherichia coli*, *dps* actúa como un componente de varias vías de respuesta al estrés (McCord & Fridovich, 1969) incluyendo la inanición, el choque térmico, la sobreexposición al hierro y el estrés oxidativo (Berend Snel *et al.*, 2001).

Los organismos mantienen la homeostasis del hierro equilibrando su captación, almacenamiento y utilización, por ello las ferritinas son esenciales para llevar a

cabo la homeostasis del hierro (Rivera, 2014; Andrews, 2010). La bacterioferritina (*bfr*), se considera la principal proteína de almacenamiento de hierro, ya que mantiene un gran núcleo de hierro en su cavidad y posee actividad ferroxidasa que protege el entorno celular reductor del ciclo redox no deseado (Rivera, 2017) lo que garantiza la suficiencia de hierro y la prevención de toxicidad inducida por éste (Arosio *et al.*, 2017). Por ejemplo, en bacterias como *Pseudomonas aeruginosa* y *Bradyrhizobium japonicum* la ferritina se expresa en respuesta a altos niveles de hierro y de peróxido de hidrógeno (Palma *et al.*, 2003; Arosio *et al.*, 2017). Se sugiere que la regulación de *bfr* en *R. ruber* MSA14 debe ser crucial para que la célula mantenga una cantidad suficiente de hierro disponible y se proteja ante la posible formación de especies reactivas de oxígeno (ROS) durante el metabolismo del pireno.

Entre la activación de genes de detoxificación, *Escherichia coli* cuenta varios reguladores principales que se activan durante el estrés oxidativo, incluyendo *OxyR*, *SoxRS* y *RpoS* (Chiang & Schellhorn, 2012), de los cuales *R. ruber* MSA14 sobreexpresa *soxR* sugiriendo una posible exposición a ROS generados como consecuencia de la exposición a pireno (Chiang & Schellhorn, 2012). A su vez, *R. ruber* MSA14 para asegurar su supervivencia a los efectos nocivos del pireno, expresa las chaperonas *groES* y *groL*, posiblemente para mantener la estructura tridimensional de las proteínas y la homeostasis celular (Kim *et al.*, 2013; Horovitz *et al.*, 2022). En cepas de *Pseudomonas* se puede observar a nivel de la envoltura celular, producción de lípidos y respuesta al estrés (Pinkart *et al.*, 1996; Domínguez-Cuevas *et al.*, 2006; Baumgarten *et al.*, 2012) como se observa con

la expresión de genes *ftsW* y *murG* en *R. ruber* MSA14, los cuales sugieren estar reforzando la pared celular como mecanismo clave de protección, además de su necesidad como proteína esencial en la división celular (Mohammadi *et al.*, 2011; Zhang *et al.*, 2021). Por lo tanto, el acoplamiento de ambos mecanismos, la capacidad de acumular TAG y de soportar estrés oxidativo, parecen ser características habituales en rodococos extremófilos (Ordoñez *et al.*, 2009; Urbano *et al.* 2013; Urbano *et al.*, 2014)

Para equilibrar la asignación de sus recursos *R. ruber* MSA14 subexpresó genes asociados a procesos únicos de traducción y maquinaria ribosomal; transporte y homeostasis; respuesta a estrés y supervivencia; degradación y reciclaje; así como secreción y translocación de proteínas, lo que sugiere que *R. ruber* MSA14 opta por una estrategia de ralentización general de los procesos de crecimiento y la priorización de la respuesta al estrés sobre la proliferación celular.

Con el fin de sobrevivir al estrés nutricional las bacterias producen una red de señalización de estrés denominada respuesta rigurosa, siendo una red de señalización bacteriana para reprogramar las células (Bennison *et al.*, 2021). Durante los períodos de inanición, las bacterias producen en exceso las alarmonas guanosina penta y tetrafosfato (denominadas colectivamente (p)ppGpp), que funcionan como mediadoras del sistema de señalización de respuesta rigurosa (Irving *et al.*, 2020) para mantener la homeostasis y la estabilidad de los componentes celulares (Fernandez-Coll & Cashel, 2020). Esto da como resultado efectos posteriores de lo que se cree que es un proceso altamente priorizado como es observado en *Staphylococcus aureus*, cepa en la

cual se identificaron las cuatro RA-GTPasas (GTPasas asociadas a ribosomas) *RsgA*, *RbgA*, *Era* y *HflX* como enzimas que pueden unirse a (p)ppGpp y que son inhibidas por el estrés nutricional, lo que resulta en un impacto negativo en el ensamblaje y crecimiento del ribosoma 70S (Corrigan et al., 2016). Lo anterior podría explicar la subexpresión del gen *era*, así como de la subexpresión de componentes ribosomales *rpl-ABEFRW* y *rps-BDKS* en *R. ruber* MSA14.

La resistencia al estrés relacionada con la presencia de xenobióticos orgánicos genera inestabilidad celular que consiste en una ralentización de la síntesis de enzimas del ciclo del ácido tricarboxílico (Domínguez-Cuevas et al., 2006). Por ejemplo en el transcriptoma de *Rhodococcus opacus* R7 crecida en o-xileno (Zampolli et al., 2020) reveló varias enzimas reprimidas tales como tres permeasas de glutamato, un transportador de dicarboxilato para fumarato, L-malato, D-malato, succinato, un supuesto simportador de sodio:dicarboxilato TctB, y transportadores de citrato TctC, una cadena beta de citrato liasa (*aceA*), una probable piruvato carboxilasa (*ppc*) y una descarboxilante de 6-fosfogluconato deshidrogenasa (Zampolli et al., 2020); lo que concuerda con lo encontrado en el análisis transcripcional de *R. ruber* MSA14 expuesta a pireno en donde se inhibió la expresión de los genes *gltB* (GO:0006537); *leuC* (GO:0009098); *argH* (GO:0042450); *aceA* (GO:0019752) y *ppc* (GO:0006099). Por su parte, en el transcriptoma de *Pseudomonas putida* KT2440 también se observa la subexpresión de los genes que codifican para las enzimas del ciclo de Krebs en presencia de tolueno y o-xileno, así como del gen *zwf* (componente clave de la vía de las pentosas fosfato) (Domínguez-Cuevas et al., 2006), el cual

también fue subexpresado en *R. ruber* MSA14 en respuesta a pireno. Ambas rutas metabólicas (ciclo de Krebs y la vía de las pentosas) están interconectadas y contribuyen al equilibrio energético y biosintético de la célula (Zampolli *et al.*, 2020).

La subexpresión del gen *aceA* en *R. ruber* MSA14 le puede conferir una capacidad limitada o nula para activar el ciclo del glioxilato, ocasionando una vulnerabilidad metabólica y de supervivencia (Ahn *et al.*, 2016). Por otro lado, la subexpresión del gen *manA* en *R. ruber* MSA14 sugiere posibles ajustes en el metabolismo de carbohidratos que podrían indirectamente influir en la disponibilidad de precursores para la síntesis lipídica. En *B. subtilis*, la inactivación de *manA* detuvo su crecimiento en medio líquido mínimo (Reizer *et al.*, 1999), confirmando que este gen es necesario para su desarrollo (Sun & Altenbuchner, 2010).

En contraste con la degradación de hidrocarburos aromáticos, donde participan enzimas tales como *bphC*, *nphA1*, *phbH*, *npcB* y *pcaG* (Embarcadero-Jiménez *et al.*, 2024), el análisis transcriptómico de este estudio no reveló sobreexpresión en estos genes; sin embargo, se detectó la subexpresión del gen *benB*. Esto podría indicar que, en las condiciones de este estudio, las etapas de degradación de compuestos aromáticos podrían estar temporalmente reprimidas, sugiriendo que *R. ruber* MSA14 prioriza mecanismos de tolerancia y adaptación antes de la degradación activa del pireno.

La compleja correulación de genes catabólicos, incluso los dispersos en el genoma (Larkin *et al.*, 2010), donde unidades transcripcionales no vinculadas

genéticamente pueden expresarse, como se ha visto con *bphA* y *etbA/ebd* en *R. jostii* RHA1 (Iwasaki *et al.*, 2007), podría explicar la aparente falta de expresión de algunas vías canónicas de degradación de compuestos aromáticos en este punto temporal en la cepa *R. ruber* MSA14.

En *R. ruber* MSA14 la respuesta a pireno induce una profunda reorganización transcripcional mediante la sobreexpresión de genes relacionados con el metabolismo energético, la síntesis de lípidos y la respuesta al estrés, lo que reafirma una estrategia de adaptación basada en la priorización de la supervivencia y la preparación para el uso del hidrocarburo, más que en su degradación inmediata. La flexibilidad metabólica y la capacidad de acumular y movilizar lípidos de reserva, así como el reforzamiento de las barreras celulares, parecen ser mecanismos clave que permiten a *R. ruber* MSA14 prosperar en ambientes contaminados con pireno.

Por otro lado, el análisis transcriptómico de *R. ruber* MSA14 frente a pireno no permitió observar la expresión diferencial de genes anotados que codifican para proteínas que conforman los SDC. Este hecho pudo haber sido por la forma en que se llevó a cabo el ensamble y la anotación del genoma, el cual se basó en una estrategia de genoma guiado. De los datos obtenidos del análisis transcriptómico de *R. ruber* MSA14, aún se tienen 4,470 genes significativos sin anotación por analizar. Entre los genes no anotados expresados diferencialmente pueden estar HQ y RR con una posible implicación en la regulación de rutas de degradación de HAP. Tal es el caso de *Lysinibacillus fusiformis* la cepa 15-4, en

cuyo análisis transcriptómico en presencia de petróleo se identificaron la sobreexpresión de 120 genes relacionados con SDC (Li *et al.*, 2020).

CONCLUSIÓN

Rhodococcus ruber MSA14, exhibe una respuesta integrada ante el pireno, la modulación de genes diferencialmente expresados sugiere un reajuste dinámico de vías reguladoras para optimizar el uso de recursos coordinando sistemas de detoxificación, reconfiguración metabólica y protección celular. *R. ruber* MSA14 activa vías para metabolizar el pireno, se protege contra el estrés oxidativo, mantiene la integridad de sus macromoléculas, ajusta su metabolismo central y la biosíntesis de componentes esenciales, a su vez, modula su crecimiento y supervivencia para adaptarse a la presencia del contaminante.

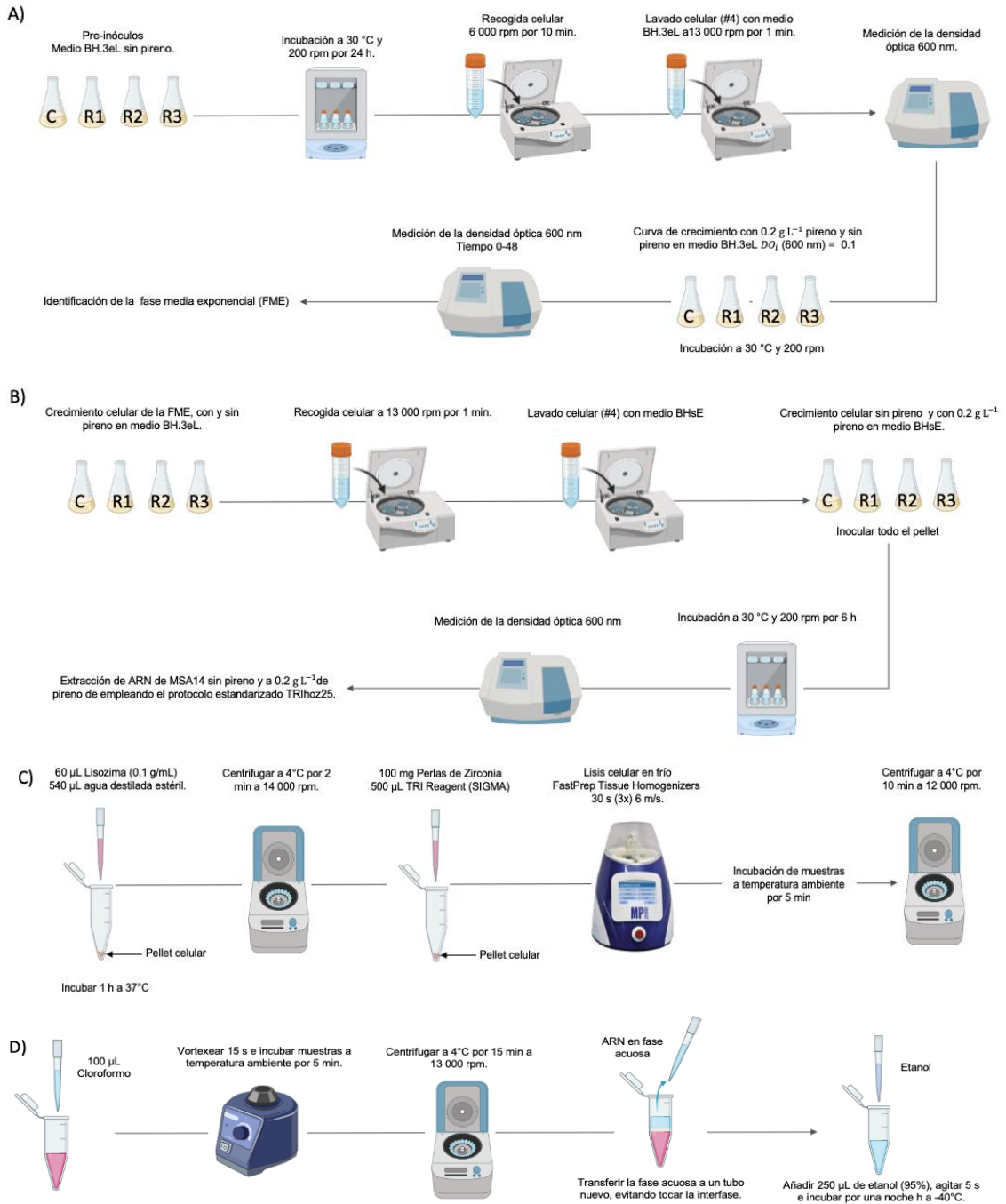
A pesar de que se tiene conocimiento de la presencia de genes relevantes para la degradación de HAP en el genoma de *R. ruber* MSA14 y factores regulatorios como sistemas de dos componentes, las condiciones específicas evaluadas en este estudio podrían limitar su expresión. El presente trabajo establece las bases para el estudio del transcriptoma de *Rhodococcus ruber* MSA14, proporcionando un marco para futuras investigaciones sobre la caracterización genética, las redes reguladoras de la transcripción tales como los sistemas de dos componentes (SDC) implicados en la regulación de las rutas de degradación de HAP en este organismo marino hidrocarbonoclasto.

PERPECTIVAS

- Análisis del transcriptoma de *R. ruber* MSA14 enfrentado a pireno basado en ensamblaje *de novo* (*de novo* RNA-seq) para superar la dependencia del tamaño de los datos del genoma de referencia.
- Análisis y validación funcional de los datos transcriptómicos (RNA-seq) por qRT-PCR.
- Análisis del transcriptoma de *R. ruber* MSA14 enfrentado a mayor tiempo de exposición a pireno.

ANEXOS

[1] PROTOCOLO ESTANDARIZADO DE EXTRACCIÓN DE ARN "TRIhoz25" en *R. ruber* MSA14.



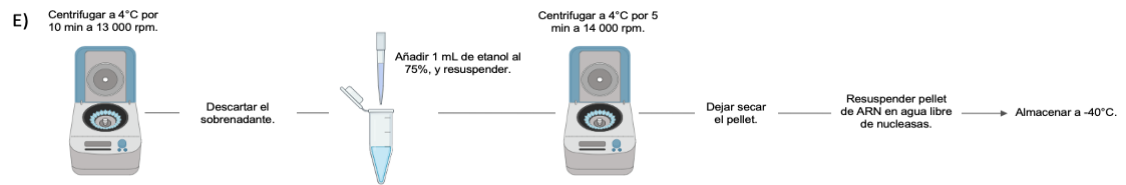


Figura 13. Obtención de ARN de *Rhodococcus ruber* MSA14 empleando el método TRIhoz25 (elaboración propia). A) Cinética de crecimiento en medio BH.3eL sin pireno y con 0.2 g L de pireno; B) Cinética de crecimiento en medio BHsE con 0.2 de pireno y sin pireno; C) Homogeneización de células de MSA14, inicio de extracción de ARN de MSA14 con TRIhoz25; D) Extracción y precipitación de ARN de MSA14, continuación de protocolo de extracción de ARN con TRIhoz25; E) Limpieza de ARN de MSA14, finalización de protocolo de extracción de ARN con TRIhoz25.

[2] LIMPIEZA DE LECTURAS (PAIR-END) CON FASTQ “PREFILTRO y POSTFILTRO”.



Figura 14. Análisis cualitativo de limpieza de lecturas empleando FASTQ; Muestras de A1_2 – B1_2.



Figura 15. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: C1_1 – D1_1.

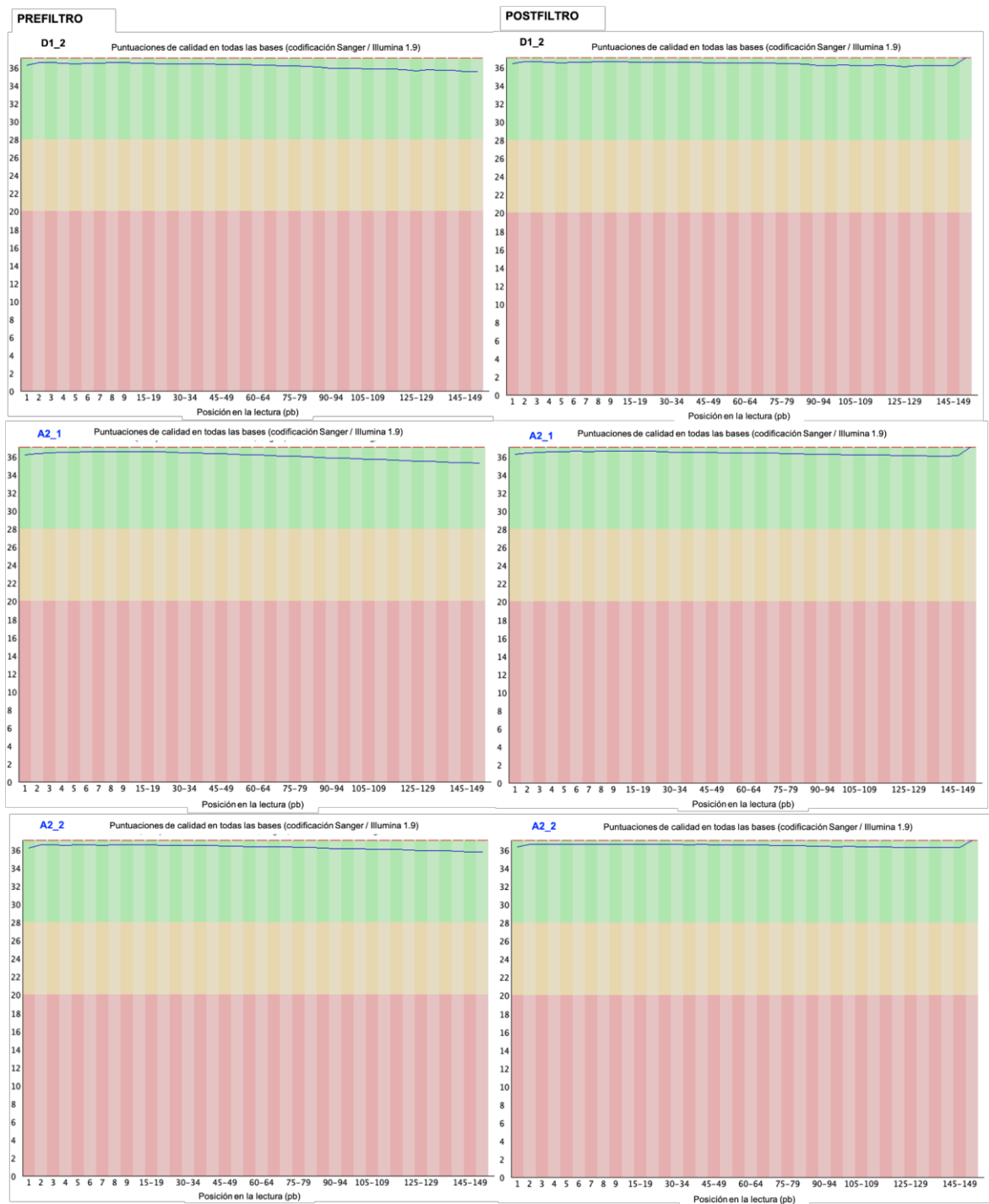


Figura 16. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: D1_2 – A2_2.



Figura 17. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: B2_1 – C2_1.

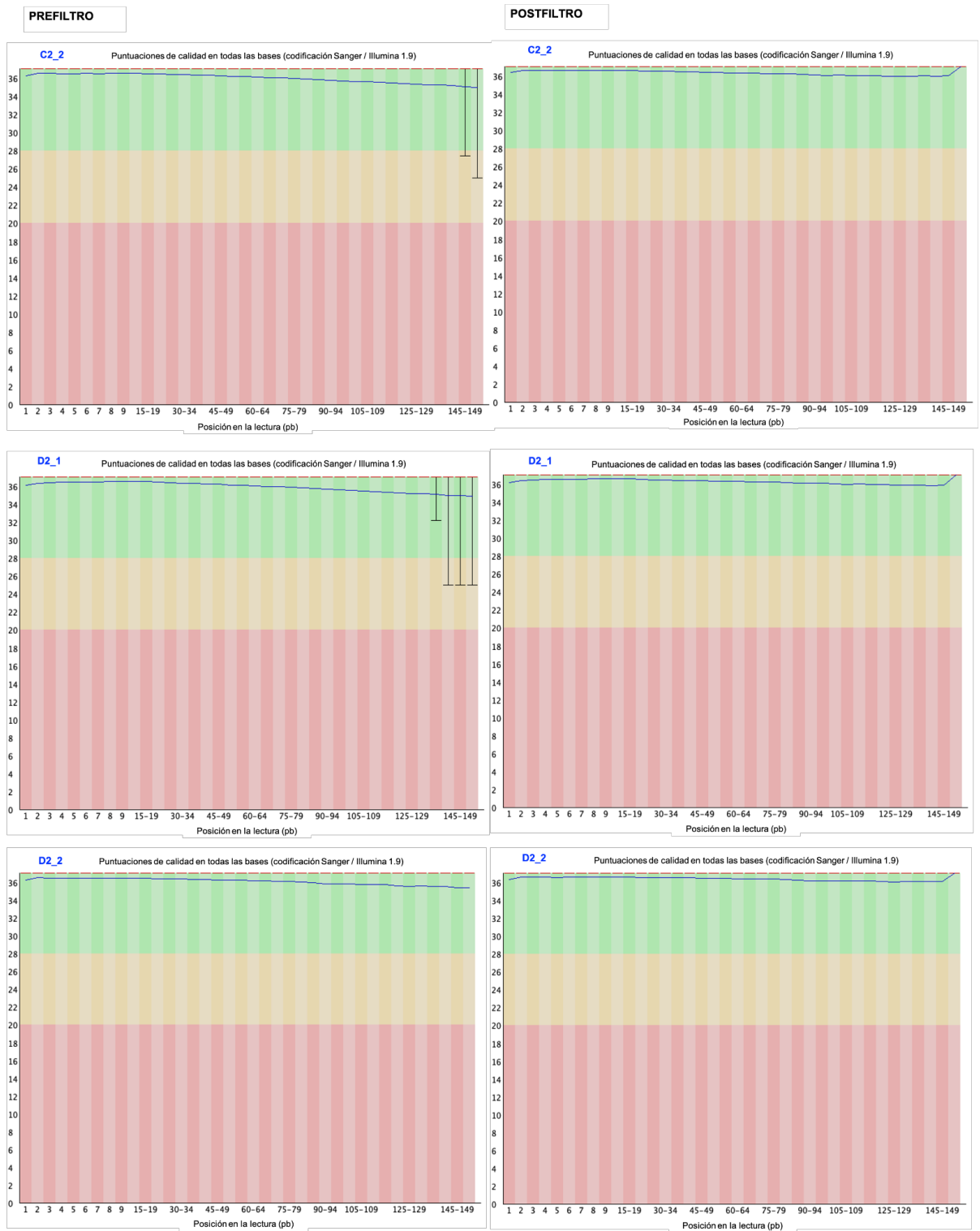
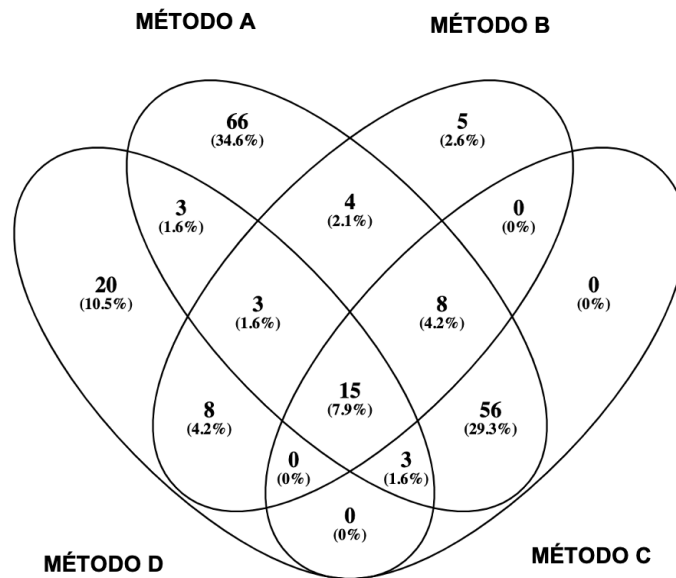


Figura 18. Análisis cualitativo de limpieza de lecturas empleando FASTQ. Muestras: C2_2 - D2_2.

[3] COMPARACIÓN DE CONJUNTOS METODOLÓGICOS DE GSA-P MEDIANTE DIAGRAMAS DE VENN GENÉTICO.

a) GSA-P



b) GsA-P

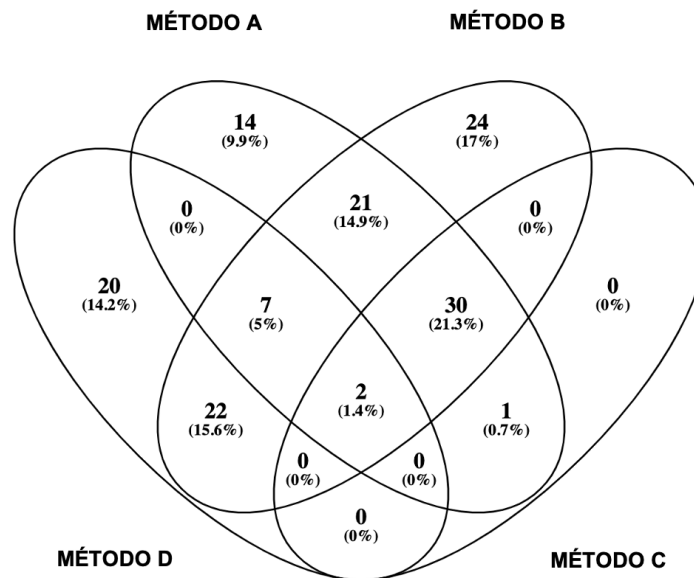


Figura 19. Proporción de GSA-P (a) y GsA-P (b) únicos y compartidos, obtenidos de las cuatro estrategias metodológicas implementadas.

[4] GSA-P y GsA-P no significativos.

Tabla XIII. Genes no significativos, inicialmente identificados como GSA-P, sin embargo, no alcanzaron con los parámetros establecidos (valores fuera del umbral): cambio de expresión superior a uno ($fc > 1$), un valor $pval \leq 0.05$ o un $qval \leq 0.1$.

GEN	PROTEÍNA	fc > 1	pval \leq 0.05	qval \leq 0.1
<i>yhjD</i>	Proteína de membrana interna YhjD	2.77139664	0.06819153	0.202509685
<i>lexA</i>	Represor transcripcional LexA	1.78466884	0.06661514	0.199995641
<i>ilvD</i>	Dihidroxiácido deshidratasa	1.78322233	0.12410252	0.276222283
<i>nrdR</i>	Regulador transcripcional NrdR	1.68599418	0.1225357	0.274369084
<i>arc</i>	ATPasa del proteasoma	1.49651447	0.0793946	0.21802899
<i>gcvH</i>	Proteína del sistema de escisión de glicina GcvH	1.48866183	0.2079433	0.372080488
<i>ndk</i>	Nucleósido-difosfato quinasa	1.43984097	0.30009954	0.468978551
<i>ctaD</i>	Subunidad I de la citocromo c oxidasa	1.37927581	0.14193876	0.296960695
<i>catA</i>	Catecol 1,2-dioxigenasa	1.28709824	0.47953821	0.622076636
<i>glgB</i>	Proteína ramificadora de 1,4-alfa-glucano GlgB	1.26034123	0.47734059	0.620097809
<i>sdhC</i>	Succinato deshidrogenasa, subunidad del citocromo b556	1.21548165	0.37946834	0.539086801
<i>mscL</i>	Proteína de canal mecanosensible de gran conductancia MscL	1.20235828	0.58058222	0.705842968
<i>mca</i>	Amidasa conjugada de micotiol Mca	1.17656177	0.14895034	0.30466906
<i>sucC</i>	Subunidad beta de la ligasa de succinato-CoA formadora de ADP	1.16726824	0.61734481	0.736250222
<i>ispG</i>	(E)-4-hidroxi-3-metilbut-2-enil-difosfato sintasa dependiente de flavodoxina	1.13248242	0.46992312	0.614089362
<i>purM</i>	Fosforribosilformilglicinamida ciclólisis	1.12217883	0.33781761	0.500491841
<i>thpD</i>	Ectoína hidroxilasa	1.11355292	0.25976158	0.427447468
<i>trxA</i>	Tiorredoxina	1.11051384	0.50687227	0.645318771
<i>map</i>	Metionil aminopeptidasa tipo I	1.10631561	0.63575413	0.749944176
<i>hflX</i>	GTPasa HflX	1.04062458	0.78668783	0.857338571
<i>clpB</i>	Chaperona dependiente de ATP ClpB	1.03431011	0.83991266	0.896800547
<i>sucD</i>	Subunidad alfa de la ligasa de succinato-CoA	1.03363912	0.88235105	0.923407098
<i>ispD</i>	2-C-metil-D-eritritol 4-fosfato citidililtransferasa	0.9963572	0.98848299	0.991332738
<i>otsB</i>	Trehalosa-fosfatasa	0.86981738	0.52340158	0.659051507
<i>clpS</i>	Adaptador de proteasa Clp dependiente de ATP ClpS	0.82750595	0.36557253	0.526804046
<i>lpdA</i>	Dihidrolipoil deshidrogenasa	0.81820998	0.28208351	0.449988308
<i>crcB</i>	Transportador de eflujo de fluoruro CrcB	0.6418019	0.01648174	0.106108198
<i>mftB</i>	Chaperona de biosíntesis de micofactocina MftB	0.60027782	0.3818279	0.541259978
<i>kstR</i>	Regulador transcripcional del catabolismo del colesterol KstR	0.57020108	0.00964818	0.087309825
<i>acnA</i>	Aconitato hidratasa AcnA	0.54285864	0.04150842	0.156722836
<i>mftC</i>	Micofactocina radical SAM madurasa	0.36131737	0.24821806	0.415395073
<i>mftD</i>	Pre-micofactocina sintasa MftD	0.32034271	0.27686868	0.444660131

<i>mdo</i>	Metanol deshidrogenasa dependiente de NDMA	0.18905692	0.24372299	0.410240831
<i>adhE</i>	Acetaldehído-CoA/alcohol deshidrogenasa bifuncional	0.18874381	0.05294681	0.178734946

Tabla XIV. Genes no significativos, inicialmente identificados como *GsA-P*, sin embargo, no alcanzaron con los umbrales establecidos: un valor $pval \leq 0.05$ o un $qval \leq 0.1$.

GEN	PROTEÍNA	fc < 1	pval \leq 0.05	qval \leq 0.1
<i>pcaG</i>	Subunidad alfa de la protocatecuato 3,4-dioxigenasa	0.9347926	0.68452859	0.786250504
<i>rox</i>	Rifampina monooxigenasa	0.88727862	0.53166333	0.66607905
<i>catA</i>	Catecol 1,2-dioxigenasa	0.87488833	0.42375016	0.575964668
<i>folP</i>	Dihidropteroato sintasa	0.85356892	0.26300467	0.431830829
<i>mftG</i>	Micofactocina deshidrogenasa MftG	0.84411033	0.444328	0.592531211
<i>aztD</i>	Metalochaperona de zinc AztD	0.80780381	0.56988515	0.697878596
<i>nirB</i>	Subunidad grande NirB de la nitrito reductasa	0.80396235	0.21219363	0.376992286
<i>coaA</i>	Pantotenato quinasa tipo I	0.79318747	0.30238869	0.471708282
<i>rfbD</i>	dTDP-4-deshidroramnosaa reductasa	0.78386588	0.06281242	0.193480065
<i>kdpA</i>	Subunidad de la ATPasa transportadora de potasio KdpA	0.75493053	0.23325133	0.400100413
<i>cofC</i>	2-fosfo-L-lactato guanililtransferasa	0.75036387	0.18566208	0.347630516
<i>ctaD</i>	Subunidad I de la citocromo c oxidasa	0.73030905	0.06072291	0.189926521
<i>scpB</i>	Subunidad ScpB del complejo SMC-Scp	0.72990903	0.06391727	0.195301015
<i>pheT</i>	Subunidad beta de la ligasa de fenilalanina-ARNt	0.72837443	0.11043029	0.258847639
<i>arsA</i>	ATPasa impulsora de la bomba de arsénico	0.71449809	0.39316492	0.551814737
<i>kdpB</i>	Subunidad KdpB de la ATPasa transportadora de potasio	0.71424887	0.33165167	0.495294342
<i>arsD</i>	Transportador de eflujo de arsenito metalochaperona ArsD	0.68097125	0.26454363	0.433226534
<i>hsaA</i>	Subunidad oxigenasa de la 3-hidroxi-9,10-secoandrosta-1,3,5(10)-trieno-9,17-diona monooxigenasa	0.65925777	0.14623742	0.302085062
<i>mutM</i>	ADN-formamidopirimidina glicosilasa bifuncional/ADN-(sitio apurínico o apirimidínico) liasa	0.64883423	0.22473873	0.390895565
<i>menB</i>	1,4-dihidroxi-2-naftoil-CoA sintasa	0.6400836	0.41051403	0.566533253
<i>lpdA</i>	Dihidrolipoil deshidrogenasa	0.62006003	0.23958052	0.406150477
<i>urtA</i>	Proteína de unión al sustrato del transportador ABC de urea	0.61643962	0.07913034	0.21802899
<i>glmM</i>	Fosfoglucosamina mutasa	0.5813555	0.05794016	0.185707461
<i>bphC</i>	Bifenil-2,3-diol 1,2-dioxigenasa	0.57049449	0.20724487	0.371233689
<i>fgd</i>	Glucosa-6-fosfato deshidrogenasa (coenzima-F420)	0.55170324	0.09843441	0.243903509
<i>eda</i>	4-hidroxi-2-oxoglutarato aldolasa/2-deshidro-3-desoxifosfluconato aldolasa bifuncional	0.51400819	0.17930795	0.340476303
<i>fdxA</i>	Ferredoxina	0.50817869	0.10115169	0.247151043

<i>efeB</i>	Subunidad deferroquelatasa/peroxidasa del transportador de captación de hierro	0.50559323	0.01436262	0.101007187
<i>trmD</i>	ARNt (guanosina(37)-N1)-metiltransferasa TrmD	0.49282102	0.09647631	0.240910127
<i>fdhA</i>	Formaldehído deshidrogenasa, independiente del glutatión	0.47587637	0.09586912	0.240271407
<i>rpsG</i>	Proteína ribosomal 30S S7	0.47053031	0.07647865	0.21558661
<i>rplM</i>	Proteína ribosomal 50S L35	0.45541517	0.05872946	0.187363947
<i>rplK</i>	Proteína ribosomal 50S L11	0.40119434	0.05028673	0.174541244
<i>dmpG</i>	4-hidroxi-2-oxovalerato aldolasa	0.38423103	0.16741336	0.327788745
<i>fdhD</i>	Formato deshidrogenasa sulfurtransferasa accesoria FdhD	0.36879211	0.10154479	0.247598469
<i>mftC</i>	Micofactocina radical SAM madurasa	0.36131737	0.24821806	0.415395073
<i>rrf</i>	ARN ribosómico 5S	0.35215288	0.23866891	0.405377312
<i>mftD</i>	Pre-micofactocina sintasa MftD	0.32034271	0.27686868	0.444660131
<i>mdo</i>	Metanol deshidrogenasa dependiente de NDMA	0.18905692	0.24372299	0.410240831

[5] GSA-P y GsA-P

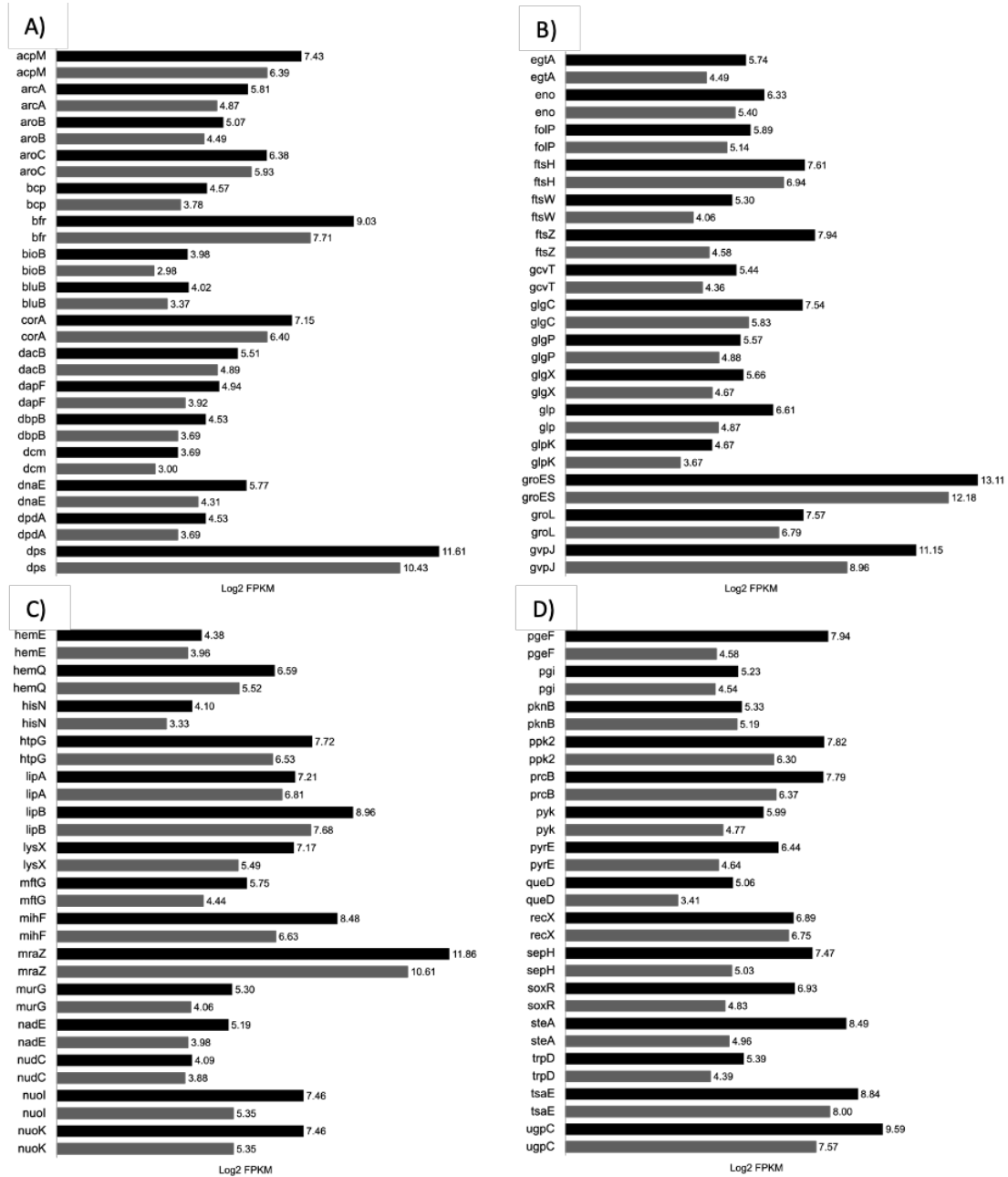


Figura 20. Validación del conjunto de GSA-P, empleando la abundancia de transcritos (FPKM) de cada gen respecto al tratamiento control. Ejemplos: *nuoI*, *acpM*, *arcA*, *steA*, *pgeF* y *soxR*; sobreexpresión aproximada de 1.39, 1.16, 1.19, 1.71, 1.73 y 1.42 veces más, respecto al tratamiento control. Expresión de GSA-P (negro) respecto a su control sin pireno (gris). Eje y: Símbolo de genes, eje x: \log_2 FPKM

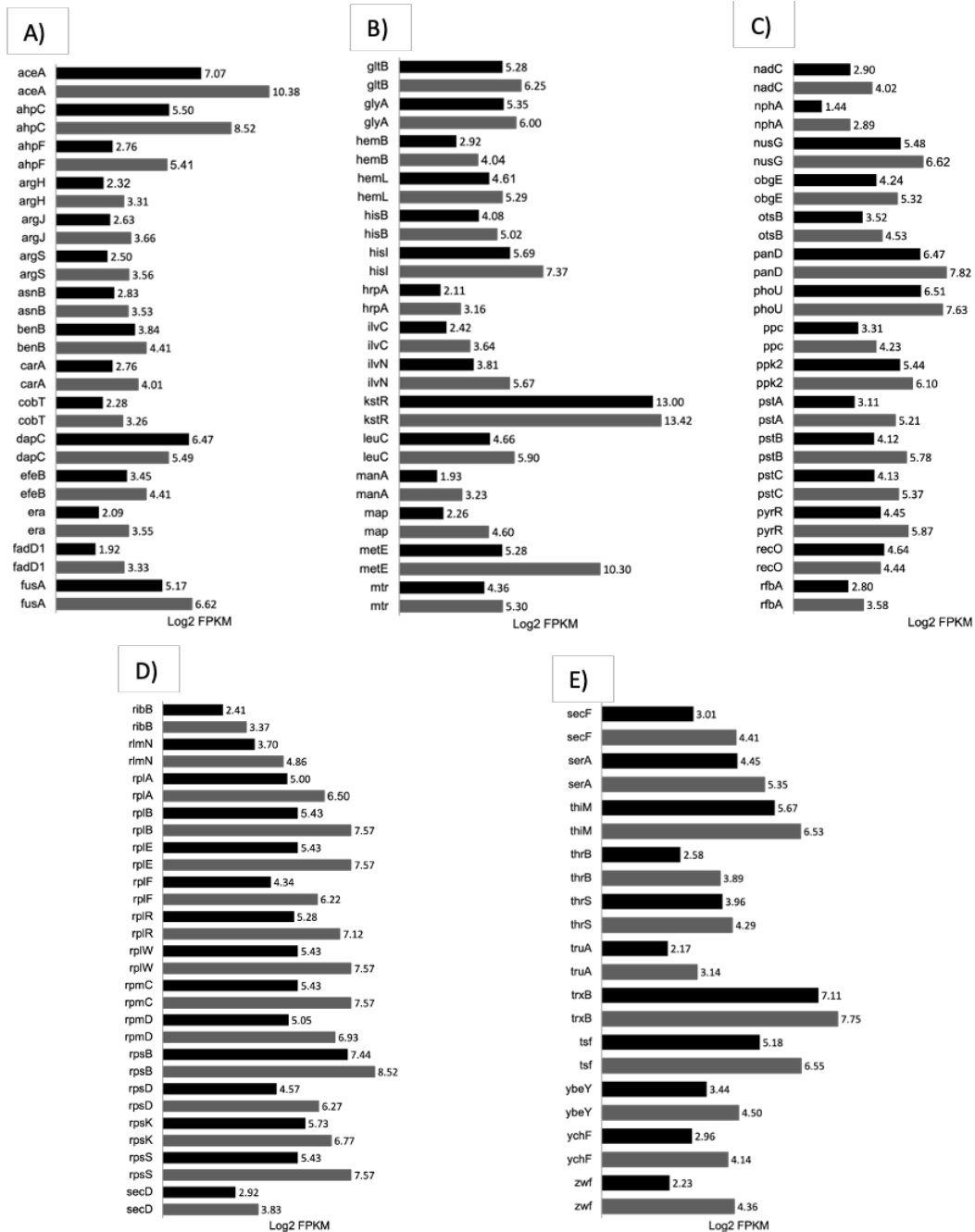


Figura 21. Validación del conjunto de GsA-P, empleando la abundancia de transcritos (FPKM) de cada gen respecto al tratamiento control. Ejemplos: *ahpF*, *nphA*, *metE*, *aceA*, *benB* y *manA*, con subexpresión aproximada de 1.96, 2, 1.95, 1.46, 1.14 y 1.67 veces más respecto al control. Expresión de GsA-P (negro) respecto a su control sin pireno (gris). Eje y: Símbolo de genes, eje x: \log_2 FPKM.

BIBLIOGRAFÍA

1. Ahmad, F., Zhu, D., & Sun, J. (2020). Bacterial chemotaxis: a way forward to aromatic compounds biodegradation. *Environmental Sciences Europe*, 32(1). doi:10.1186/s12302-020-00329-2
2. Ahn, S., Jung, J., Jang, I. A., Madsen, E. L., & Park, W. (2016). Role of glyoxylate shunt in oxidative stress response. *Journal of Biological Chemistry*, 291(22), 11928-11938.
3. Alm, E., Huang, K., & Arkin, A. (2006). The Evolution of Two-Component Systems in Bacteria Reveals Different Strategies for Niche Adaptation. *PLoS Computational Biology*, 2(11), e143. doi:10.1371/journal.pcbi.0020143
4. Alonso-Casajús N, Dauvillée D, Viale AM, Muñoz FJ, Baroja-Fernández E, Morán-Zorzano MT, Eydallin G, Ball S, Pozueta-Romero J: Glycogen phosphorylase, the product of the *glgP* gene, catalyzes glycogen breakdown by removing glucose units from the nonreducing ends in *Escherichia coli*. *J Bacteriol*. 2006, 188: 5266-5272. 10.1128/JB.01566-05.
5. Altendorf, K., Booth, I. R., Gralla, J. A. Y., Greie, J. C., Rosenthal, A. Z., & Wood, J. M. (2009). Osmotic stress. *EcoSal Plus*, 3(2), 10-1128.
6. Alvarez HM (2006). Bacterial Triacylglycerols. *Triglycerides and Cholesterol Research*. Edited by: Welson LT., New York: Nova Science Publishers Inc, 6: 159-176.
7. Alvarez, A. F., & Georgellis, D. (2023). Environmental adaptation and diversification of bacterial two-component systems. *Current Opinion in Microbiology*, 76, 102399. <https://doi.org/10.1016/j.mib.2023.102399>
8. Alvarez, H. M. (2003). Relationship between β -oxidation pathway and the hydrocarbon-degrading profile in actinomycetes bacteria. *International biodeterioration & biodegradation*, 52(1), 35-42.
9. Alvarez, H. M., & Steinbüchel, A. (2019). Biology of triacylglycerol accumulation by *Rhodococcus* (pp. 299-332). Springer International Publishing.
10. Andrews, S. C. (2010). The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1800(8), 691-705.

11. Angeles, D. M., & Scheffers, D. J. (2020). The cell wall of *Bacillus subtilis*. *Current issues in molecular biology*, 41, 539-596.
12. Arosio, P., Cairo, G., & Bou-Abdallah, F. (2024). A Brief History of Ferritin, an Ancient and Versatile Protein. *International Journal of Molecular Sciences*, 26(1), 206.
13. Arosio, P., Elia, L., & Poli, M. (2017). Ferritin, cellular iron storage and regulation. *IUBMB life*, 69(6), 414-422.
14. Ashby, M. K., & Houmard, J. (2006). Cyanobacterial two-component proteins: structure, diversity, distribution, and evolution. *Microbiology and molecular biology reviews: MMBR*, 70(2), 472–509. <https://doi.org/10.1128/MMBR.00046-05>
15. Atlas, R. M., & Hazen, T. C. (2011). Oil Biodegradation and Bioremediation: A Tale of the Two Worst Spills in U.S. History. *Environmental Science & Technology*, 45(16), 6709–6715. <https://doi.org/10.1021/es2013227>
16. Beld, J., Lee, D. J., & Burkart, M. D. (2015). Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Molecular BioSystems*, 11(1), 38-59.
17. Bennison, D. J., Nakamoto, J. A., Craggs, T. D., Milón, P., Rafferty, J. B., & Corrigan, R. M. (2021). The stringent response inhibits 70S ribosome formation in *Staphylococcus aureus* by impeding GTPase-ribosome interactions. *Mbio*, 12(6), e02679-21.
18. Borkowski, O., Ceroni, F., Stan, G. B., & Ellis, T. (2016). Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology*, 33, 123-130.
19. Borkowski, O., Ceroni, F., Stan, G.-B., & Ellis, T. (2016). Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology*, 33, 123–130. <https://doi.org/10.1016/j.mib.2016.07.009>
20. Calisto, F., Sousa, F. M., Sena, F. V., Refojo, P. N., & Pereira, M. M. (2021). Mechanisms of energy transduction by charge translocating membrane proteins. *Chemical Reviews*, 121(3), 1804-1844.

21. Capra, E. J., & Laub, M. T. (2012). Evolution of Two-Component Signal Transduction Systems. *Annual Review of Microbiology*, 66(1), 325–347. doi:10.1146/annurev-micro-092611-150039
22. Casino, P., Rubio, V., & Marina, A. (2010). The mechanism of signal transduction by two-component systems. *Current Opinion in Structural Biology*, 20(6), 763-771. <https://doi.org/10.1016/j.sbi.2010.09.010>
23. Cereijo, A. E., Ferretti, M. V., Iglesias, A. A., Álvarez, H. M., & Asencion Diez, M. D. (2024). Study of two glycosyltransferases related to polysaccharide biosynthesis in *Rhodococcus jostii* RHA1. *Biological Chemistry*, 405(5), 325-340.
24. Cheung, K. J., Badarinarayana, V., Selinger, D. W., Janse, D., & Church, G. M. (2003). A microarray-based antibiotic screen identifies a regulatory role for supercoiling in the osmotic stress response of *Escherichia coli*. *Genome research*, 13(2), 206-215.
25. Cock PJ & Whitworth DE. (2007). Evolution of prokaryotic two-component system signaling pathways: gene fusions and fissions. *Molecular Biology Evolution*, 24:2355–57
26. Corrigan RM, Bellows LE, Wood A, Gründling A. 2016. ppGpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in Gram-positive bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 113:E1710–E1719.
27. Costa, V., Angelini, C., De Feis, I., & Ciccodicola, A. (2010). Uncovering the Complexity of Transcriptomes with RNA-Seq. *Journal of Biomedicine and Biotechnology*, 2010, 1–19. doi:10.1155/2010/853916
28. Dangar, K., Ramani, K., & Changela, D. (2022). Analyzing the Capabilities of Actinobacteria for Remediation Through Metagenomic Analysis of Contaminated Habitats. In *Microbial BioTechnology for Sustainable Agriculture Volume 1* (pp. 471-487). Singapore: Springer Nature Singapore.
29. De Benedetti, S., Fisher, J. F., & Mobashery, S. (2021). Bacterial cell wall: Morphology and biochemistry. In *Practical Handbook of Microbiology* (pp. 167-204). CRC Press.

30. De Carlo, S. (2006). The structural basis for regulated assembly and function of the transcriptional activator NtrC. *Genes and Development*, 20(11), 1485–1495. doi:10.1101/gad.1418306
31. de Carvalho, C. C. (2012). Adaptation of *Rhodococcus erythropolis* cells for growth and bioremediation under extreme conditions. *Research in microbiology*, 163(2), 125-136.
32. Driessen, A. J., & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annual Review Biochemistry*, 77(1), 643-667.
33. Eberlein, C., Baumgarten, T., Starke, S., & Heipieper, H. J. (2018). Immediate response mechanisms of Gram-negative solvent-tolerant bacteria to cope with environmental stress: cis-trans isomerization of unsaturated fatty acids and outer membrane vesicle secretion. *Applied Microbiology and Biotechnology*, 102, 2583-2593. <https://doi.org/10.1007/s00253-018-8832-9>
34. Elyamine, A. M., Kan, J., Meng, S., Tao, P., Wang, H., & Hu, Z. (2021). Aerobic and anaerobic bacterial and fungal degradation of pyrene: mechanism pathway including biochemical reaction and catabolic genes. *International Journal of Molecular Sciences*, 22(15), 8202.
35. Embarcadero-Jiménez, S., Araujo-Palomares, C. L., Moreno-Perlín, T., Ramírez-Álvarez, N., Quezada-Hernández, C., Batista-García, R. A., Sanchez-Flores, A., Calcáneo-Hernández, G., & Silva-Jiménez, H. (2024). Physiology and comparative genomics of the haloalkalitolerant and hydrocarbonoclastic marine strain *Rhodococcus ruber* MSA14. *Archives of microbiology*, 206(7), 328. <https://doi.org/10.1007/s00203-024-04050-z>
36. Fedorov, A. N. (2022). Biosynthetic protein folding and molecular chaperons. *Biochemistry (Moscow)*, 87(Suppl 1), S128-S145.
37. Fuentes, S., Barra, B., Caporaso, J. G., & Seeger, M. (2015). From Rare to Dominant: a Fine-Tuned Soil Bacterial Bloom during Petroleum Hydrocarbon Bioremediation. *Applied and Environmental Microbiology*, 82(3), 888–896. doi:10.1128/aem.02625-15
38. Galperin M. Y. (2006). Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *Journal of Bacteriology*, 188(12), 4169–4182. <https://doi.org/10.1128/JB.01887-05>

39. Galperin, M. Y. (2005). A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts. *BMC Microbiology*, 5(1), 35. doi:10.1186/1471-2180-5-35
40. Gao, R., & Stock, A. M. (2009). Biological insights from structures of two-component proteins. *Annual review of Microbiology*, 63(1), 133-154
41. Gao, R., Mack, T. R., & Stock, A. M. (2007). Bacterial response regulators: versatile regulatory strategies from common domains. *Trends in Biochemical Sciences*, 32(5), 225–234. doi:10.1016/j.tibs.2007.03.002
42. Ghosh, A., Khurana, M., Chauhan, A., Takeo, M., Chakraborti, A. K., & Jain, R. K. (2010). Degradation of 4-Nitrophenol, 2-Chloro-4-nitrophenol, and 2,4-Dinitrophenol by *Rhodococcus imtechensis* Strain RKJ300. *Environmental Science and Technology*, 44(3), 1069–1077. doi:10.1021/es9034123
43. Goncalves, E. R., Hara, H., Miyazawa, D., Davies, J. E., Eltis, L. D., & Mohn, W. W. (2006). Transcriptomic Assessment of Isozymes in the Biphenyl Pathway of *Rhodococcus* sp. Strain RHA1. *Applied and Environmental Microbiology*, 72(9), 6183–6193. doi:10.1128/aem.00947-06
44. Groisman, E. A. (2016). Feedback Control of Two-Component Regulatory Systems. *Annual Review of Microbiology*, 70(1), 103–124. doi:10.1146/annurev-micro-102215-095331
45. Guermouche M'rassi, A., Bensalah, F., Gury, J., & Duran, R. (2015). Isolation and characterization of different bacterial strains for bioremediation of *n*-alkanes and polycyclic aromatic hydrocarbons. *Environmental Science and Pollution Research*, 22(20), 15332–15346. doi:10.1007/s11356-015-4343-8
46. Guevara, G., Castillo Lopez, M., Alonso, S., Perera, J., & Navarro-Llorens, J. M. (2019). New insights into the genome of *Rhodococcus ruber* strain Chol-4. *BMC Genomics*, 20, 1-17.
47. Gupta, G., Kumar, V., & Pal, A. K. (2019). Microbial degradation of high molecular weight polycyclic aromatic hydrocarbons with emphasis on pyrene. *Polycyclic Aromatic Compounds*, 39(2), 124-138.

48. Head, I., Jones, D. & Röling, W. (2006). Marine microorganisms make a meal of oil. *Nature Reviews Microbiology*, 4(3), 173–182. <https://doi.org/10.1038/nrmicro1348>
49. Hernández, M. A., Mohn, W. W., Martínez, E., Rost, E., Alvarez, A. F., & Alvarez, H. M. (2008). Biosynthesis of storage compounds by *Rhodococcus jostii* RHA1 and global identification of genes involved in their metabolism. *BMC Genomics*, 9(1). <https://doi.org/10.1186/1471-2164-9-600>
50. Horovitz, A., Tali Haviv Reingewertz, Cuéllar, J., & José María Valpuesta. (2022). Chaperonin Mechanisms: Multiple and (Mis)Understood? *Annual Review of Biophysics*, 51(1), 115–133. <https://doi.org/10.1146/annurev-biophys-082521-113418>
51. Imlay, J. A. (2003). Pathways of Oxidative Damage. *Annual Review of Microbiology*, 57(1), 395–418. <https://doi.org/10.1146/annurev.micro.57.030502.090938>
52. Imlay, J. A. (2008). Cellular defenses against superoxide and hydrogen peroxide. *Annual Review of Biochemistry*, 77(1), 755-776.
53. Inouye, M., & Dutta, R. (2002). *Histidine kinases in signal transduction*. Elsevier.
54. Juarez, A., Villa, J. A., Lanza, V. F., Lázaro, B., de la Cruz, F., Alvarez, H. M., & Moncalián, G. (2017). Nutrient starvation leading to triglyceride accumulation activates the Entner Doudoroff pathway in *Rhodococcus jostii* RHA1. *Microbial Cell Factories*, 16, 1-12.
55. Karas VO, Westerlaken I, Meyer AS. (2015). The DNA-Binding Protein from Starved Cells (Dps) Utilizes Dual Functions To Defend Cells against Multiple Stresses. *Journal of Bacteriology*. <https://doi.org/10.1128/jb.00475-15>
56. Kenney, L. J. (2010). How important is the phosphatase activity of sensor kinases?. *Current Opinion in Microbiology*, 13(2), 168-176.
57. Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M., & Ulrich Hartl, F. (2013). Molecular chaperone functions in protein folding and proteostasis. *Annual Review of Biochemistry*, 82(1), 323-355.
58. Krell, T., Lacal, J., Busch, A., Silva-Jiménez, H., Guazzaroni, M.-E., & Ramos, J. L. (2010). Bacterial Sensor Kinases: Diversity in the Recognition of Environmental Signals. *Annual Review of Microbiology*, 64(1), 539–559. doi:10.1146/annurev.micro.112408.134054

59. Lacal, J., Reyes-Darias, J. A., García-Fontana, C., Ramos, J.-L., & Krell, T. (2013). Tactic responses to pollutants and their potential to increase biodegradation efficiency. *Journal of Applied Microbiology*, 114(4), 923–933. <https://doi.org/10.1111/jam.12076>
60. Laub MT. (2011). The role of two-component signal transduction systems in bacterial stress responses. In *Bacterial Stress Responses*, ed. G Storz, R Hengge, pp. 45–58. Washington, DC: ASM Press
61. Lemuth, K., Hardiman, T., Winter, S., Pfeiffer, D., Keller, M. A., Lange, S., & Siemann-Herzberg, M. (2008). Global transcription and metabolic flux analysis of *Escherichia coli* in glucose-limited fed-batch cultivations. *Applied and environmental microbiology*, 74(22), 7002-7015.
62. Li, J., Wang, Y., Zhou, W., Chen, W., Deng, M., & Zhou, S. (2020). Characterization of a new biosurfactant produced by an effective pyrene-degrading *Achromobacter* species strain AC15. *International Biodeterioration and Biodegradation*, 152, 104959.[doi:10.1016/j.ibiod.2020.104959](https://doi.org/10.1016/j.ibiod.2020.104959)
63. Madan Babu, M., & Teichmann, S. A. (2003). Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucleic Acids Research*, 31(4), 1234–1244. <https://doi.org/10.1093/nar/gkg210>
64. Maneerat, S. & Phetrong, K (2007). Isolation of biosurfactant-producing marine bacteria and characteristics of selected biosurfactant. *Songklanakarin J. Sci. Technol.*, 2007, 29(3):781-791.
65. Martínková, L., Uhnáková, B., Pátek, M., Nešvera, J., & Křen, V. (2009). Biodegradation potential of the genus *Rhodococcus*. *Environment international*, 35(1), 162-177.
66. Mascher, T., Helmann, J. D., & Unden, G. (2006). Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases. *Microbiology and Molecular Biology Reviews*, 70(4), 910–938.[doi:10.1128/mnbr.00020-06](https://doi.org/10.1128/mnbr.00020-06)
67. Matilla, M. A., Gavira, J. A., Monteagudo-Cascales, E., Zhulin, I. B., & Krell, T. (2025). Structural and functional diversity of sensor domains in bacterial transmembrane receptors. *Trends in Microbiology*. <https://doi.org/10.1016/j.tim.2025.02.019>

68. McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJM, Holt R, Brinkman FSL, Miyauchi K, Fukuda F, Davies JE, Mohn WW, Eltis LD. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proceedings of the National Academy of Sciences of the United States of America*, 103: 15582-15587. [10.1073/pnas.0607048103](https://doi.org/10.1073/pnas.0607048103).
69. Mohammadi, T., Dam, V. van, Sijbrandi, R., Vernet, T., André Zapun, Bouhss, A., Bruin, M. D., Nguyen-Distèche, M., Kruijff, B. de, & Eefjan Breukink. (2011). Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *The EMBO Journal*, 30(8), 1425–1432. <https://doi.org/10.1038/emboj.2011.61>
70. Palma, M., Worgall, S. & Quadri, L. E. N. (2003). Transcriptome Analysis of the *Pseudomonas aeruginosa* Response to Iron. *Archives of Microbiology*, 180, 374– 379 DOI: [10.1007/s00203-003-0602-z](https://doi.org/10.1007/s00203-003-0602-z)
71. Papagianni, M. (2012). Recent advances in engineering the central carbon metabolism of industrially important bacteria. *Microbial Cell Factories*, 11, 1-13.
72. Parales, R. E., Parales, J. V., Pelletier, D. A., & Ditty, J. L. (2008). Chapter 1 Diversity of Microbial Toluene Degradation Pathways. *Advances in Applied Microbiology*, 1–73. doi:10.1016/s0065-2164(08)00401-2
73. Patel, A. B., Shaikh, S., Jain, K. R., Desai, C., & Madamwar, D. (2020). Polycyclic Aromatic Hydrocarbons: Sources, Toxicity, and Remediation Approaches. *Frontiers in Microbiology*, 11. doi:10.3389/fmicb.2020.562813
74. Peña-Sandoval, G. R., & Georgellis, D. (2010). The ArcB sensor kinase of *Escherichia coli* autophosphorylates by an intramolecular reaction. *Journal of Bacteriology*, 192(6), 1735–1739. <https://doi.org/10.1128/JB.01401-09>
75. Perteza, M., Kim, D., Perteza, G. M., Leek, J. T., & Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, 11(9), 1650–1667. <https://doi.org/10.1038/nprot.2016.095>

76. Prince, R. C. (2010). Bioremediation of Marine Oil Spills. Handbook of Hydrocarbon and Lipid Microbiology, 2617–2630. doi:10.1007/978-3-540-77587-4_194
77. Reizer, J., Bachem, S., Reizer, A., Arnaud, M., Saier Jr, M. H., & Stülke, J. (1999). Novel phosphotransferase system genes revealed by genome analysis—the complete complement of PTS proteins encoded within the genome of *Bacillus subtilis*. Microbiology, 145(12), 3419-3429.
78. Rivera, M. (2014) Bacterioferritin: Structure Function and Protein-Protein Interactions. In Handbook of Porphyrin Science; Kadish, K. K.; Smith, K. M.; Guillard, R., Eds.; Vol. 30; pp 135– 178, DOI: DOI: 10.1142/9789814407755_0041
79. Rivera, M. (2017). Bacterioferritin: structure, dynamics, and protein–protein interactions at play in iron storage and mobilization. Accounts of Chemical Research, 50(2), 331-340.
80. Santos-Beneit, F. (2015). The Pho regulon: a huge regulatory network in bacteria. Frontiers in Microbiology, 6, 402.
81. Seibold G. & Eikmanns BJ (2007) The *glgX* gene product of *Corynebacterium glutamicum* is required for glycogen degradation and for fast adaptation to hyperosmotic stress. Microbiology. 153: 2212-2220. 10.1099/mic.0.2006/005181-0.
82. Seo, J.-S., Keum, Y.-S., & Li, Q. X. (2009). Bacterial Degradation of Aromatic Compounds. International Journal of Environmental Research and Public Health, 6(1), 278-309. <https://doi.org/10.3390/ijerph6010278>
83. Sevvana, M., Vijayan, V., Zweckstetter, M., Reinelt, S., Madden, D. R., Herbst-Irmer, R., & Becker, S. (2008). A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA. Journal of Molecular Biology, 377(2), 512-523.
84. Shajani, Z., Sykes, M. T., & Williamson, J. R. (2011). Assembly of bacterial ribosomes. Annual Review of Biochemistry, 80(1), 501-526.
85. Shimodaira, J., Furusawa, Y., Miyazawa, Y., Kasai, D., Miyauchi, K., Masai, E., & Fukuda, M. (2012). The 24-bp consensus sequence responsible for regulation of the BphS1T1 two-component system in a hybrid promoter. Journal of Bioscience and Bioengineering, 113(3), 279–285. <https://doi.org/10.1016/j.jbiosc.2011.10.021>

86. Silva-Jiménez, H., Araujo-Palomares, C.L., Macías-Zamora, J.V., Ramírez-Álvarez, N., García-Lara, B. & Corrales-Escobosa, A.R. (2018). Identification by MALDI-TOF MS of Environmental Bacteria with High Potential to Degrade Pyrene. *Journal of the Mexican Chemical Society*, 62(2), 214- 225. <https://doi.org/10.29356/jmcs.v62i2.411>
87. Silva-Jiménez, H., García-Fontana, C., Cadirci, B. H., Ramos-González, M. I., Ramos, J. L., & Krell, T. (2012). Study of the TmoS/TmoT two-component system: towards the functional characterization of the family of TodS/TodT like systems. *Microbial Biotechnology*, 5(4), 489-500.
88. Starosta, A. L., Lassak, J., Jung, K., & Wilson, D. N. (2014). The bacterial translation stress response. *FEMS microbiology reviews*, 38(6), 1172-1201.
89. Stock, Ann M.; Robinson, Victoria L.; Goudreau, Paul N. (2000). Two-component signal transduction. *Annual Review of Biochemistry*, 69:183–215. <https://doi.org/10.1146/annurev.biochem.69.1.183>
90. Sun, T., & Altenbuchner, J. (2010). Characterization of a mannose utilization system in *Bacillus subtilis*. *Journal of Bacteriology*, 192(8), 2128-2139.
91. Takeda, H., Shimodaira, J., Yukawa, K., Hara, N., Kasai, D., Miyauchi, K., Masai, E., & Fukuda, M. (2010). Dual two-component regulatory systems are involved in aromatic compound degradation in a polychlorinated-biphenyl degrader, *Rhodococcus jostii* RHA1. *Journal of Bacteriology*, 192(18), 4741–4751. <https://doi.org/10.1128/JB.00429-10>
92. Terán-Melo, J. L., Rodríguez-Rangel, C., Dimitris Georgellis, & Álvarez, A. F. (2019). Mecanismos de autofosforilación y transfosforilación en sistemas de dos componentes bacterianos. *TIP Revista Especializada En Ciencias Químico-Biológicas*. <https://doi.org/10.22201/fesz.23958723e.2019.0.162>
93. Toro-Roman, A., Mack, T. R., & Stock, A. M. (2005). Structural analysis and solution studies of the activated regulatory domain of the response regulator ArcA: a symmetric dimer mediated by the $\alpha 4$ - $\beta 5$ - $\alpha 5$ face. *Journal of Molecular Biology*, 349(1), 11-26.
94. Ulrich, L. E., & Zhulin, I. B. (2007). MiST: a microbial signal transduction database. *Nucleic Acids Research*, 35(Database issue), D386–D390. <https://doi.org/10.1093/nar/gkl932>
95. Ulrich, L. E., & Zhulin, I. B. (2009). The MiST2 database: a comprehensive genomics resource on microbial signal transduction. *Nucleic Acids Research*, 38(suppl_1), D401–D407. [doi:10.1093/nar/gkp940](https://doi.org/10.1093/nar/gkp940)

96. Urbano, S. B., Di Capua, C., Cortez, N., Farías, M. E., & Alvarez, H. M. (2014). Triacylglycerol accumulation and oxidative stress in *Rhodococcus* species: differential effects of pro-oxidants on lipid metabolism. *Extremophiles*, 18, 375-384.
97. Volkman, B. F. (2001). Two-State Allosteric Behavior in a Single-Domain Signaling Protein. *Science*, 291(5512), 2429–2433. doi:10.1126/science.291.5512.2429
98. Wang, Z., Hou, X., Shang, G., Deng, G., Luo, K., & Peng, M. (2024). Exploring Fatty Acid β -Oxidation Pathways in Bacteria: From General Mechanisms to DSF Signaling and Pathogenicity in *Xanthomonas*. *Current Microbiology*, 81(10), 336.
99. West, A. H., & Stock, A. M. (2001). Histidine kinases and response regulator proteins in two-component signaling systems. *Trends in Biochemical Sciences*, 26(6), 369–376. [https://doi.org/10.1016/s0968-0004\(01\)01852-7](https://doi.org/10.1016/s0968-0004(01)01852-7)
100. Wuichet, K., & Zhulin, I. B. (2010). Origins and Diversification of a Complex Signal Transduction System in Prokaryotes. *Science Signaling*, 3(128), ra50–ra50. doi:10.1126/scisignal.2000724
101. Yakimov, M. M., Timmis, K. N., & Golyshin, P. N. (2007). Obligate oil-degrading marine bacteria. *Current opinion in biotechnology*, 18(3), 257-266. <https://doi.org/10.1016/j.copbio.2007.04.006>
102. Yang, P., Liu, W., Chen, Y., & Gong, A. D. (2022). Engineering the glyoxylate cycle for chemical bioproduction. *Frontiers in Bioengineering and Biotechnology*, 10, 1066651.
103. Yu, K., & Zhang, T. (2012). Metagenomic and Metatranscriptomic Analysis of Microbial Community Structure and Gene Expression of Activated Sludge. *PLoS ONE*, 7(5), e38183–e38183. <https://doi.org/10.1371/journal.pone.0038183>
104. Zampolli, J., Di Canito, A., Manconi, A., Milanese, L., Di Gennaro, P., & Orro, A. (2020). Transcriptomic analysis of *Rhodococcus opacus* R7 grown on o-xylene by RNA-Seq. *Frontiers in Microbiology*, 11, 1808.
105. Zampolli, J., Zeaiter, Z., Di Canito, A., & Di Gennaro, P. (2018). Genome analysis and -omics approaches provide new insights into the biodegradation potential of *Rhodococcus*. *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-018-9539-7

106. Zhao, J., Baba, T., Mori, H., & Shimizu, K. (2004). Effect of *zwf* gene knockout on the metabolism of *Escherichia coli* grown on glucose or acetate. *Metabolic Engineering*, 6(2), 164-174.
107. Zobell, C. E., & Upham, H. C. (1944). A list of marine bacteria including descriptions of sixty new species. University of California Press. *Bull Scripps Inst Oceanog* 5:239–292
108. Zschiedrich, C. P., Keidel, V., & Szurmant, H. (2016). Molecular Mechanisms of Two-Component Signal Transduction. *Journal of Molecular Biology*, 428(19), 3752–3775.[doi:10.1016/j.jmb.2016.08.003](https://doi.org/10.1016/j.jmb.2016.08.003)