

UCH-FC
Doc-BMCN
L358
C1

FACULTAD DE CIENCIAS
UNIVERSIDAD DE CHILE



**“Identificación de redes de regulación transcripcional activadas en respuesta a cobre en
Enterococcus faecalis”**

Tesis
Entregada A La
Universidad De Chile
En Cumplimiento Parcial De Los Requisitos
Para Optar al Grado De
Doctor en Ciencias con Mención en Biología Molecular,
Celular Y Neurociencias
Facultad De Ciencias
Por

Mauricio Alejandro Latorre Mora

Diciembre, 2012
Santiago – Chile

Director de Memoria de Título: Dr. Mauricio González Canales.
Laboratorio de Bioinformática y Expresión Génica.
INTA – Universidad de Chile.

FACULTAD DE CIENCIAS

UNIVERSIDAD DE CHILE
INFORME DE APROBACION
TESIS DE DOCTORADO



Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por el candidato.

MAURICIO ALEJANDRO LATORRE MORA

Ha sido aprobada por la comisión de evaluación de la tesis como requisito para optar al grado de Doctor en Ciencias con mención en Biología Molecular Celular y Neurociencias, en el examen de Defensa Privada de Tesis rendido el día

Director de Tesis

Dr. Mauricio González Canales

A handwritten signature in blue ink, reading 'Mauricio González Canales', written over a horizontal line.

Comisión de Evaluación de la Tesis

Dr. Marco Tulio Núñez (presidente)

Dra. Ana María Preller

Dr. Miguel Allende

Dr. Bernardo González

Dr. Juan Carlos Letelier

Four handwritten signatures in blue ink, each written over a horizontal line. From top to bottom, they correspond to: Marco Tulio Núñez, Ana María Preller, Miguel Allende, and Juan Carlos Letelier.

*A todos aquellos que hicieron posible
dar este importante paso de mi vida*



En la célebre ciudad de Santiago de Chile, un 17 de septiembre de 1981 nace el Señor Mauricio Alejandro Latorre Mora. A sus cortos cinco años de edad, parte su aventura en las áreas del conocimiento al ingresa al Glorioso Liceo Experimental Manuel de Salas, lugar donde concluye con éxito toda su enseñanza escolar. En su andar por las diversas materias y caminos de la vida, el año 2001 entra a la Ilustre Universidad de Chile como estudiante de Ingeniería en Biotecnología Molecular. Es aquí donde, desarrolla una pasión incondicional por las ciencias biológicas que posteriormente, en el año 2007, lo hacen tomar la decisión retornar a su querida Universidad para continuar y así profundizar sus conocimientos adquiridos, tomando el desafío de convertirse en Doctor en Ciencias con mención en Biología Molecular, Celular y Neurociencias, siendo este el último paso en su formación como estudiante y así pasar finalmente a ser un flamante líder de investigación científica, cuya historia está aún por escribirse....

AGRADECIMIENTOS



En esta oportunidad quiero agradecer a todas las personas que me apoyaron y creyeron en mí nuevamente, brindándome su confianza, ayuda y cariño.

Partir agradeciendo a mi Mamá *Georgina*, por su constante preocupación y amor en cada uno de los instantes de mi vida, a mi tía *Nina* por estar siempre conmigo, cuidándome y protegiéndome. A mi polola *Carola*, por vivir todo este camino junto a mí, dándome su amor y compañía. A mis amigos *Inos*, *Gori*, *Heri*, *Possy* y a mis segundas familias *Rusiñol* y *Medi*, gracias por su apoyo y cariño.

A mi estimado tutor, profe *Mauricio* gracias por la confianza depositada, consejos y reflexiones, y por sobre todo la libertad para desarrollarme a plenitud como científico. A la Dra. *Cambiazo*, por su ejemplo como profesional y ayuda en los momentos indicados.

A mis compañeros y amigos del Laboratorio de Bioinformática y Expresión Génica (destacando como es de costumbre a la gerencia de LBEG). Al Laboratorio de Bioinformática y Matemática del Genoma, en particular al Dr. *Maass* por brindarme un espacio en su Laboratorio y a *Marko*, por su constante e incansable disposición. Al grupo de "*Division of Infectious Diseases*", muchas gracias *Barbara* por la posibilidad de formar parte activa del Laboratorio, *Jung* por tus consejos, *Jessica* por tu ayuda incondicional, *KV*, *Karen* y al grupo del Dr. *Cesar Arias* gracias por todo.

A mis amigas de pregrado, *Lore*, *Cinthia*, *Sole* y *Cata*, por esas incansables conversaciones de nuestro futuro y buenos deseos. A mis amigos y compañeros del Doctorado, por esas largas horas de estudio, trabajos y presentaciones que posteriormente se transformaron en carretes, bailes y comilonas, muchas gracias.

A los profesores de la Universidad, por su formación y disposición para entregar sus conocimientos. Agradecer a la comisión, por sus críticas asertivas y constructivas, por esos consejos que permitieron finalizar esta tesis con éxito, aportando siempre en mi formación como doctor. Al programa de becas CONICYT, Vicerrectoría de Asuntos Académicos de la Universidad de Chile, programa MECESUP UCH0713, proyectos FONDECYT 1110427 y FONDAP 15090007, gracias por el financiamiento otorgado.

Y a todos aquellos quienes con un simple consejo o una palabra de ánimo impulsaron este trabajo y mi desarrollo como doctor,

gracias.

INDICE DE CONTENIDOS

DEDICATORIA.....	ii
AGRADECIMIENTOS.....	iii
INDICE DE CONTENIDOS.....	v
INDICE DE TABLAS.....	ix
INDICE DE FIGURAS.....	x
INDICE DE ANEXOS.....	xi
LISTA DE ABREVIATURAS.....	xii
RESUMEN.....	xiv
ABSTRACT.....	xv
INTRODUCCION.....	1
1. El cobre.....	1
2. Respuestas transcripcionales globales frente a la disponibilidad de cobre.....	2
3. Redes de regulación transcripcional mediadas por factores de transcripción.....	3
4. Enterococcus faecalis como modelo biológico elegido.....	9
5. Hipótesis.....	13
6. Objetivo general.....	14
7. Objetivos específicos.....	15

MATERIALES Y METODOS.....	16
1. MATERIALES.....	16
1.1 Cepas y condiciones de cultivo.....	16
1.2 Sales, soluciones y reactivos.....	19
1.3 Partidores.....	19
1.4 Algoritmos bioinformáticos.....	19
2. METODOS.....	21
2.1. Análisis <i>in silico</i>	21
2.1.1. Identificación de elementos que van a constituir la red de regulación transcripcional <i>in silico</i>	21
2.1.2. Construcción <i>in silico</i> de la red de regulación transcripcional mediada por factores de transcripción (escala global).....	22
2.1.3 Identificación de una red de regulación transcripcional activada por Cu y selección de regulones para verificaciones conectividades <i>in vivo</i>	23
2.1.4 Análisis topológicos y generación de logos.....	24
2.2. Biología Molecular.....	25
2.2.1. Reacción en cadena de la polimerasa (PCR).....	25
2.2.2. Electroforesis.....	26
2.2.3. Construcción y verificación de cepas mutantes para factores de transcripción en <i>E. faecalis</i> OG1RF.....	27

2.3. Ensayos de expresión génica.....	29
2.3.1. Extracción de RNA total.....	29
2.3.2. Cuantificación y calidad del RNA.....	31
2.3.3. Síntesis de cDNA.....	31
2.3.4. PCR cuantitativo en tiempo real (qPCR).....	32
2.4. Protocolos de caracterización microbiológica.....	33
2.4.1. Curva de crecimiento.....	33
2.4.2. Contenido intracelular de cobre.....	34
2.5. Análisis estadístico.....	35
 RESULTADOS	 36
 1. Construcción <i>in silico</i> de una posible red de regulación transcripcional activada en respuesta a la exposición a Cu.....	 36
1.1 Identificación y clasificación de los componentes de la red.....	36
1.2 Construcción <i>in silico</i> de la red de regulación transcripcional mediada por factores de transcripción (escala global).....	43
1.3 Análisis topológicos y asignación de propiedades de la red transcripcional global de <i>E. faecalis</i>	50
1.4 Identificación de una posible red de regulación transcripcional activada por Cu y aplicación de filtros de selección.....	57

2. Determinación de la participación del(los) factor(es) de transcripción en la regulación de la expresión de los componentes de la red inducidos por Cu.....	61
2.1 Verificación de los resultados de microarreglos a través de ensayos de PCR en tiempo real (qPCR).....	61
2.2 Construcción y verificación de cepas mutantes de <i>E. faecalis</i> OG1RF para los factores de transcripción CopY, Fur y LexA.....	63
2.3 Análisis de cambios transcripcionales mediante qPCR en las cepas mutantes.....	67
3. Evaluación de la dependencia de la regulación transcripcional mediada por los factores de transcripción en la resistencia de <i>E. faecalis</i> a Cu.....	71
3.1 Curvas de crecimiento.....	71
3.2 Contenido intracelular de Cu.....	74
DISCUSION.....	76
CONCLUSIONES.....	115
REFERENCIAS BIBLIOGRAFICAS.....	117
ANEXOS.....	125

INDICE DE TABLAS

Tabla 1. Cepas bacterianas y plásmido.....	18
Tabla 2. Distribución del total de los genes dentro de los operones identificados en el genoma de <i>E. faecalis</i>	39
Tabla 3. Posibles factores de transcripción agrupados por familia presentes en el genoma de <i>E. faecalis</i>	42
Tabla 4. Porcentaje de categorías COG de los genes contenidos en la red global en <i>E. faecalis</i>	47
Tabla 5. Número de sitios de unión para cada familia de factores de transcripción y clasificación según tipo de factor de acuerdo con datos reportados en otras bacterias.....	55
Tabla 6. Patrones presentes en la red global de regulación transcripcional de <i>E. faecalis</i> agrupados por el tipo de conectividad.....	56
Tabla 7. Verificación por qPCR de la activación por Cu de los regulones CopY, Fur y LexA.....	57
Tabla 8. Verificación por qPCR de las conectividades predichas in silico mediante el uso de mutantes para los regulones CopY, Fur y LexA.....	69
Tabla 9. Tiempo generacional para las cepas silvestre y mutantes para los genes <i>copY</i> , <i>fur</i> y <i>lexA</i> de <i>E. faecalis</i> expuestas a diferentes concentraciones de Cu.....	73

INDICE DE FIGURAS

Figura 1. Distribución del tipo de operones en genomas bacterianos.....	40
Figura 2. Proporción del total de factores de transcripción en relación al total de genes presentes en diversos genomas.....	41
Figura 3. Red de regulación transcripcional a escala global de <i>E. faecalis</i>	46
Figura 4. Distribución porcentual de la posición del total de sitios de unión encontrados dentro de su región promotora dentro de la red global.....	48
Figura 5. Logos consenso para los sitios de unión contenidos en la red global.....	49
Figura 6. Distribución del grado de entrada para los genes presentes en la red de regulación transcripcional de <i>E. faecalis</i>	54
Figura 7. Sub-red de regulación transcripcional activada por Cu en <i>E. faecalis</i>	60
Figura 8. Verificación de la cepa mutante <i>lexA</i> de <i>E. faecalis</i>	65
Figura 9. Verificación de las cepas mutante <i>copY</i> y <i>fur</i> de <i>E. faecalis</i>	66
Figura 10. Sub-red de regulación transcripcional activada por Cu curada <i>in vivo</i> para los regulones CopY, Fur y LexA de <i>E. faecalis</i>	70
Figura 11. Curva de crecimiento para las cepas silvestre y mutantes para los genes <i>copY</i> , <i>fur</i> y <i>lexA</i> expuestas a concentraciones de Cu.....	72
Figura 12. Contenido intracelular de Cu cepas silvestre y mutantes para los genes <i>copY</i> , <i>fur</i> y <i>lexA</i> de <i>E. faecalis</i>	75

INDICE DE ANEXOS

ANEXO 1. Tabla de partidores en la construcción de mutantes y ensayos de qPCR.....	126
ANEXO 2. Publicación: Hödar C, Moreno P, di Genova A, Latorre M, Reyes-Jara A, Maass A, González M, Cambiazo V.” Genome wide identification of <i>Acidithiobacillus ferrooxidans</i> (ATCC 23270) transcription factors and comparative analysis of ArsR and MerR metal regulators”. <i>Biometals</i> . 2012 Feb;25(1):75-93.....	128
ANEXO 3. Sub-red de regulación transcripcional activada por 0,05 mM de CuSO ₄ (Red de Abrantes).....	129
ANEXO 4. Publicación: López G, Latorre M, Reyes-Jara A, Maass A, Cambiazo V, González M.” Transcriptomic response of <i>Enterococcus faecalis</i> to iron excess”. <i>Biometals</i> . 2012 Aug;25(4):737-47.....	130
ANEXO 5. Contenido intracelular de Fe en cepas silvestre y mutantes <i>fur</i> de <i>E. faecalis</i> expuestas a 0,5 mM de CuSO ₄	131
ANEXO 6. Resumen congreso internacional “Tema 14-Trace Elements in Man and Animals”.....	132
ANEXO 7. Resumen congreso internacional “Fourth International Workshop on Iron and Copper Homeostasis”.....	133
ANEXO 8. Resumen congreso internacional “Plant and Animal Genome XX”.....	134

LISTA DE ABREVIATURAS

AAS:	Espectrometría de absorción atómica
Ace1p:	Factor de transcripción de cobre
ArgR:	Familia de reguladores transcripcionales de respuesta a estrés
ATP7a:	ATPasa transportadora de eflujo de cobre Menkes
ATP7b:	ATPasa transportadora de eflujo cobre Wilson
BHI:	Infusión de cerebro y corazón
Blast:	<i>Basic Local Alignment Search Tool</i>
BlastP:	Blast de proteínas
CDS:	Secuencias codificantes
CMI:	Concentración mínima inhibitoria
COG:	Grupos de genes ortólogos
CopA:	ATPasa transportadora de eflujo cobre
<i>copA</i> :	Gen que codifica para la proteína CopA
CopB:	ATPasa transportadora de captación de cobre
<i>copB</i> :	Gen que codifica para la proteína CopB
CopY:	Regulador transcripcional represor del operon <i>cop</i>
<i>copY</i> :	Gen que codifica para la proteína CopY
CopZ:	Chaperona de cobre
<i>copZ</i> :	Gen que codifica para la proteína CopZ
Cp:	Punto de cruce
CRP-FNR:	Familia de reguladores transcripcionales de metabolismo basal
CTR1:	Transportador de cobre 1
CTR3:	Transportador de cobre 3
Cu:	Cobre
CueR:	Factor de transcripción de respuesta a cobre
CusRS:	Sistema de dos componentes de respuesta a cobre
CutC:	Proteína de homeostasis de cobre
DBTBS:	<i>Data Base for Transcription Binding Sites</i>
DnaA:	Factor de transcripción de replicación
DO:	Densidad óptica
EBI:	<i>European Bioinformatic Institute</i>
EDTA:	Ácido etilendiamino tetra acético
Ef:	Eficiencia de la reacción de qPCR
EMSA:	Ensayo de retardo electroforético
Fe:	Hierro
Fur:	Regulador transcripcional de homeostasis de hierro (familia Fur)
<i>fur</i> :	Gen que codifica para la proteína Fur

GalR-S:	Familia de reguladores transcripcionales de metabolismo galactosa
LexA:	Regulador transcripcional de respuesta a estrés
<i>lexA</i> :	Gen que codifica para la proteína LexA
LysR:	Familia de reguladores transcripcionales de respuesta a estrés
Mac1p:	Factor de transcripción de cobre
MalT:	Factor de transcripción de metabolismo de maltosa
MOPS:	Acido 3-(N-morfolino) propanosulfónico
MPP	Matriz de peso probabilística
MTF1:	Factor de transcripción de metales
NCBI:	<i>National Center of Biotechnology Information</i>
ODB:	<i>Operon data base</i>
PBM:	Microarreglos de unión de proteína
PBS:	Tampón fosfato salino
PcoRS:	Sistema de dos componentes de respuesta a cobre
PCR:	Reacción en cadena de la polimerasa
PDB:	<i>Protein Data Bank</i>
Per:	Regulador transcripcional de respuesta a estrés (familia Fur)
Pfam:	<i>Protein family data base</i>
PurR:	Factor de transcripción de metabolismo de purinas
qPCR:	PCR en tiempo real
ROS:	Especies reactivas de oxígeno
SHOPS:	<i>Show Operon Structures</i>
SOD:	Super óxido dismutase
sRNA: :	RNA de tamaño pequeño no codificante
TAE:	Tampón tris-acetato EDTA
THB:	Caldo <i>Todd-Hewitt</i>
Tm:	Tiempo de fusión
Tris:	Tris-(hidroximetil)-aminometano
U:	Unidades enzimáticas
Uber:	Algoritmo de búsqueda de operones
UV:	Ultra violeta
VIS:	Espectro de luz visible
Zn:	Zinc
Zur:	Regulador transcripcional de homeostasis de zinc (familia Fur)

RESUMEN

El cobre (Cu) es un micronutriente requerido para el óptimo desarrollo y funcionamiento celular, sin embargo en exceso, resulta tóxico para todos los sistemas biológicos. Frente a cambios en la disponibilidad extracelular del Cu, estudios de expresión génica global, indican la existencia de cambios transcripcionales de genes que codifican para proteínas involucradas en la homeostasis de Cu, mecanismos antioxidantes y otras rutas metabólicas, que refuerzan la respuesta celular a cambios de exposición del metal. Sin embargo, desconocemos los elementos regulatorios de la transcripción capaces de establecer conectividades en la respuesta a nivel global.

En relación a esto, utilizando como modelo de estudio la bacteria *Enterococcus faecalis*, se construyó *in silico* una red de regulación transcripcional global de factores de transcripción y sus correspondientes genes blanco, capaz integrar diferentes datos de expresión génica. La red resultante de este cruce de información y su posterior verificación en la bacteria, permitió proponer que la respuesta transcripcional frente al Cu se establece mediante la activación de módulos coordinados transcripcionalmente por un conjunto común de elementos reguladores capaces de traducir diferentes estímulos generados por la exposición sostenida al metal, cuyo grado de conectividad depende esencialmente, de la participación en diferentes procesos celulares de las proteínas codificadas en cada uno de ellos, destacándose sistemas relacionados con homeostasis de Cu, captación de hierro, reparación de DNA, metabolismo basal y producción de energía.

Los resultados obtenidos en esta tesis, además permitieron brindar a la comunidad científica, un nuevo modelo global de regulación transcripcional, el cual puede ser utilizado para el estudio de la respuesta celular frente a otros estímulos y contrastar las propiedades descritas en otros modelos bacterianos.

ABSTRACT

Copper (Cu) is a micronutrient required for optimal development and functioning of cells. However, in excess, it is toxic for all biological systems. Global gene expression studies indicate the existence of transcriptional changes of genes that encode for proteins involved in Cu homeostasis, antioxidant mechanisms and other metabolic pathways, which enhance the cellular response to metal exposure changes in response to changes in the Cu extracellular availability. However, the transcriptional regulatory elements that help to establish connections in a global level response are largely unknown.

In this regard, a global transcriptional regulatory network of transcription factors and their target genes, was constructed *in silico* using the bacteria *Enterococcus faecalis* as a study model, allowing to integrate different gene expression data. The network resulting from this information intersection, and its subsequent verification in the bacteria, allowed to suggest that the transcriptional response against Cu is established by the activation of modules transcriptionally coordinated by a common set of regulatory elements capable of translating different stimulus generated by a sustained metal exposure, whose connectivity degree depends essentially of the participation in different cellular processes of the proteins encoded in each of them, among them, Cu homeostasis, iron uptake, DNA repair, basal metabolism and energy production.

The results obtained in this thesis allows to provide to the scientific community a new global model of transcriptional regulation, which can be used for the study of the cellular response to other stimulus and contrast the described properties in other bacterial models.

INTRODUCCION

1. El cobre

El Cobre (Cu) es un elemento traza requerido para el correcto funcionamiento de diversas enzimas tanto de organismos procariontes como eucariontes. En los sistemas biológicos el Cu puede existir en sus estados reducido (Cu^{+1}) u oxidado (Cu^{+2}), posibilitando su participación en la oxidación de diferentes sustratos¹. Su intervención en procesos de respiración celular, transporte de hierro (Fe) y protección frente a radicales libres, hacen del Cu un elemento de transición fundamental para la realización de una serie de procesos bioquímicos en forma normal². Sin embargo, el exceso de éste metal puede inducir estrés oxidativo generando especies reactivas del oxígeno (ROS) tales como los radicales hidroxilos a través de las reacciones de Haber-Weiss y Fenton³. El desequilibrio entre la generación de ROS y los sistemas antioxidantes pueden causar peroxidación lipídica, modificación proteica y daño al DNA^{1,4}. Sumado a esto, el Cu puede unirse directamente a grupos sulfhidrilos de proteínas e incluso desplazar a otros metales que actúan como co-factores, conduciendo a la inactivación enzimática y/o a la pérdida de estructura proteica⁵.

Esta dualidad, requiere de la existencia de mecanismos homeostáticos, capaces de controlar la concentración intracelular de Cu. En este sentido, se reconocen tres etapas involucradas principalmente en el manejo celular de Cu, a saber⁶: i) Captación: ejecutada por proteínas de membrana encargadas de la incorporación del metal hacia el interior de la célula, ii) Manejo intracelular (almacenamiento y distribución): realizada por proteínas que unen transitoriamente el metal, por ejemplo, metalotioneinas y chaperonas de Cu), iii) Eflujo:

realizada vía proteínas de membrana capaces de transferir el metal hacia el medio extracelular, siendo las ATPasas tipo P, las más comunes y mejor descritas a la fecha.

La respuesta celular a Cu depende entonces de la acción coordinada de un conjunto de proteínas cuya actividad controla la tasa de incorporación, acumulación y salida del metal, procesos que en organismos procariontes y eucariontes unicelulares son regulados fuertemente a nivel transcripcional por proteínas (factores de transcripción) cuya función es modulada directamente por la unión al metal^{7, 8}.

Disfunciones de los componentes homeostáticos de Cu así como de los sistemas de protección antioxidantes generan, en humanos, diferentes patologías. Ejemplo de ello son las patologías de Menkes y Wilson^{9, 10}, generadas por mutaciones en genes que codifican para las ATPasa de eflujo ATP7a y ATP7b respectivamente. El Síndrome de Menkes es una enfermedad ligada al cromosoma X, que produce deficiencia de Cu generando un deterioro mental progresivo, hipotermia, anormalidades del tejido conectivo, que finalmente llevan a la muerte aproximadamente a los tres años de edad. La enfermedad de Wilson es autosómica y se caracteriza por excreción biliar de Cu defectuosa lo que lleva a una acumulación del metal en hígado y ojos, produciendo los característicos anillos oculares o Kayser-Fleischer, seguida por trastornos neurológicos, psiquiátricos, renales, hematológicos, endocrinos y cirrosis hepática¹¹. Por otro lado, la pérdida del control de los mecanismos de protección frente a la toxicidad del Cu se ha vinculado con enfermedades multifactoriales como el Alzheimer e incluso a problemas cardíacos presentes en la diabetes^{12, 13}, abriendo la interrogante de que además de los mecanismos homeostáticos, otros procesos metabólicos vitales para un correcto funcionamiento celular estarían siendo afectados por el desbalance de este micronutriente.

2. Respuestas transcripcionales globales frente a la disponibilidad de cobre

Diversos han sido los esfuerzos para entender cual es la respuesta de la célula a nivel global frente a la exposición a Cu y como es su capacidad para coordinar diversos procesos metabólicos en post de mantener la homeostasis celular. En esta línea de investigación, experimentos de microarreglos realizados en diferentes modelos biológicos, enfocados a determinar los efectos del Cu (déficit o exceso) a nivel transcripcional¹⁴, cuentan que, además de los mecanismos homeostáticos de Cu, existen otros transcritos que cambian su abundancia, los cuales codifican para proteínas con funciones asociadas a adhesión celular, organización del citoesqueleto, modificación y síntesis de proteínas, señalización, metabolismo energético basal (glicólisis, vía de las pentosas) y protección contra estrés oxidativo, los que abarcan cerca de un 80 % del total de los cambios observados. Estos resultados dan cuenta que el Cu es capaz de inducir cambios transcripcionales globales a nivel celular de genes que codifican para componentes involucrados en otros procesos celulares, sugiriendo mecanismos de regulación génica capaces de responder al estímulo coordinando la expresión de estos elementos frente a la exposición al metal.

A la fecha, se han descrito diversos mecanismos para entender la sensibilidad de los elementos de regulación transcripcional que responden a Cu, capaces de influir directamente sobre la expresión de genes que participan en la homeostasis intracelular del metal¹⁵. En el modelo eucarionte de *Saccharomyces cerevisiae* se postula que cambios en la disponibilidad del metal, activan la acción coordinada de dos factores de transcripción^{16, 17}. Frente a un déficit del Cu, la proteína Mac1p activa la expresión de los transportadores de captación (CTR1/CTR3), por

el contrario el factor Ace1p, en altas condiciones del metal activa la expresión de Cup1 (chaperona de acumulación) y genes de respuesta a estrés (SOD).

En bacterias, el modelo más simple y mejor estudiado de regulación transcripcional es el operon *cop* comandado por el represor CopY, presente en diversos organismos procariontes. Este factor, impide la transcripción de genes encargados del manejo intracelular del metal a través de la unión sobre sus regiones reguladoras. El modelo describe que en un estado de homeostasis, CopY reprime genes involucrados con la homeostasis del metal. Ahora bien, en un escenario de exceso, la chaperona CopZ le entrega un átomo de Cu a CopY, el complejo CopY-Cu posee una menor afinidad hacia el DNA, liberándose de las regiones reguladoras, permitiendo que ocurra la transcripción génica de componentes involucrados principalmente en resistencia, como la ATPasa de Cu CopA¹⁸.

En humanos, la regulación transcripcional sensible a Cu es de forma indirecta; proteínas chaperonas denominadas metalotioneínas¹⁹, como respuesta a un exceso del metal intracelular, intercambian átomos de Cu por átomos de Zinc (Zn) presentes en su estructura, los átomos de Zn liberados se unen al factor de transcripción MTF1, el complejo MTF1-Zn activa la expresión de genes que codifican para las metalotioneínas, las cuales actúan como un quelante de Cu evitando la toxicidad éste en su estado libre.

Las evidencias derivadas de estudios en procariontes y eucariontes indican que el Cu modula de forma directa o indirecta, factores de transcripción encargados de regular la expresión de genes involucrados en la homeostasis del metal y mecanismos antioxidantes (sistemas de protección), genes que abarcan aproximadamente solo el 10 % del total de cambios transcripcionales observados cuando células son expuestas a diferentes tratamientos con el metal. En cuanto a los sistemas regulatorios, poco conocemos de otros factores de transcripción capaces

también de responder al estímulo Cu, o cuál es la coordinación que se establece entre estos elementos para generar una respuesta a nivel global que de cuenta de cambios en las abundancias transcripcionales de genes que codifican para los componentes involucrados en otros procesos celulares.

3. Redes de regulación transcripcional mediadas por factores de transcripción

De los componentes involucrados en regulación transcripcional, los factores de transcripción tienen una participación fundamental en censar un estímulo para luego traducirlo en una serie de cascadas de interacción que involucran reacciones bioquímicas que terminan con la coordinación y regulación de eventos de regulación de la transcripción²⁰. En general, el mecanismo de acción de estos factores es promover (activador) o imposibilitar (represor) el reclutamiento de la maquinaria transcripcional en lugares particulares dentro del genoma²¹. Para esto, estas proteínas son capaces de reconocer secuencias nucleotídicas específicas denominados sitios de unión o *binding sites*, las cuales están presentes principalmente en las regiones promotoras de uno o varios operones bacterianos (blancos de la regulación)²².

Dado este mecanismo, las relaciones que se establecen presentan direccionalidad y secuencialidad en cada uno de sus pasos, iniciándose con la capacidad de un determinado factor de transcripción de reconocer un estímulo, seguido con el reconocimiento del sitio de unión por parte de la proteína en las regiones promotoras de sus genes blanco lo que finalmente conduce a la activación o represión de la transcripción.

Recientemente, múltiples esfuerzos se han efectuado para entender como es el funcionamiento a nivel global de los factores de transcripción sobre la regulación de un genoma.

Con el propósito de abordar el problema desde una visión sistémica, en este trabajo de investigación se ha definido operacionalmente que los factores de transcripción y sus genes blancos forman parte de un circuito de interacciones complejo, cuya conectividad entre ellos va a depender principalmente de identificación o no de un determinado sitio de unión²³. Esta visión permite integrar toda la información, pudiendo la suma total de las conectividades establecidas entre los elementos del sistema ser representada como una red de regulación transcripcional a escala global, en la cual los genes y los factores son los nodos dentro de la red y los sitios de unión las interacciones entre ellos²⁴.

En ésta línea de investigación, mediante el uso de diversas herramientas bioinformáticas y experimentos a escala global (genómica, transcriptómica, proteómica, etc.) se ha logrado generar diversas hipótesis de redes de regulación transcripcional. La primera red de regulación construida capaz de integrar y describir a nivel sistémico las interacciones que se establecen entre los factores de transcripción y sus correspondientes genes blancos se propuso en la levadura *S. cerevisiae*²⁵. En esta red se describió por primera vez la presencia de unidades de regulación que se repiten a lo largo del genoma, se caracterizó la estructura de una red de regulación transcripcional donde se distinguen módulos particulares para dar forma a una red compleja de interacciones, describiendo niveles de conectividades donde destacan diferentes factores de transcripción capaces de regular otros factores. En procariontes, los modelos actualmente más depurados de redes de regulación transcripcional se encuentran en *Escherichia coli*²⁶. El uso de este organismo modelo, no sólo permitió estudiar las conectividades transcripcionales a gran escala, sino también logró dar cuenta de la dinámica de la regulación en función al tipo de estímulo, estableciendo en general, que la topología que presenta la red, limita en gran medida la capacidad de la bacteria para adecuarse a los diferentes medios²⁶.

Con la integración del conocimiento generado a la fecha sobre de redes de regulación transcripcional, no solo se ha logrado entender la composición e interacción de los elementos que las constituyen, sino también se ha conseguido describir diversas propiedades comunes en estos sistemas. En este contexto, diversos estudios tanto topológicos como funcionales, concuerdan que las principales ventajas que le otorga a la célula contar con este tipo de redes de regulación transcripcional radican en²⁷: i) Velocidad de propagación del estímulo: por ejemplo, en el caso de que un sitio de unión esté en más de una región promotora, el factor de transcripción puede eventualmente activar o reprimir más de un gen, confiriéndole al sistema una alta y rápida propagación del estímulo, así como una coordinación en la expresión cuya funcionalidad es una pronta y diversa respuesta transcripcional por parte de la célula; ii) Variabilidad de la respuesta: la capacidad limitada de censar un estímulo por parte de los factores de transcripción conduce a un alto grado de especificidad. No obstante, existen estímulos capaces de afectar la función de forma directa o indirectamente a un alto número de factores de transcripción, tal es el caso por ejemplo, del estrés inducido por privación de alimentos, cambios de temperatura o presencia de ROS.

Estas características en definitiva, permiten diferenciar dentro de una red de regulación transcripcional a nivel global, nodos capaces de responder frente a ciertos estímulos, cuyo grado de especificidad en la respuesta va a depender del tipo de perturbación al cual se enfrenta la célula y de la capacidad de mediación de los factores de transcripción por traducir esta señal en regulación y coordinación de la expresión de uno o varios genes. A finales del año 2009, se construyó una posible red global de regulación transcripcional en *Mycobacterium tuberculosis* capaz de integrar datos de microarreglos temporales obtenidos para dos estados de desarrollo (replicativa y no-replicativa)²⁸. La red permitió sugerir cuales eran las conectividades entre los

genes que se activaban o reprimían según las condiciones de replicación de la bacteria y los factores de transcripción involucrados, destacando sub-conjuntos de nodos particulares dentro de la red global (sub-redes) involucradas en cada uno de los procesos. Estas sub-redes describían la activación de nuevos genes con sus respectivos factores de transcripción capaces de detectar las diferentes condiciones metabólicas de crecimiento, cuya finalidad supone establecer un sistema regulado para componentes involucrados directamente en los procesos de replicación así como de otras proteínas encargadas de generar la energía necesaria para llevar a cabo estos procesos. Además, con la clasificación de estas sub-redes se logró determinar cuales eran los nodos de respuesta temprana y tardía, cuya temporalidad y dinámica de expresión va a depender de la capacidad del factor de responder de forma directa o indirecta a las condiciones generadas en cada uno de los estados de desarrollo.

Este tipo de estudios además de integrar la respuesta transcripcional global de un organismo, permiten entender a nivel sistémico, como los cambios de expresión génica contribuyen al proceso de adaptación de una célula frente a un estímulo del medio, identificando los mecanismos específicos involucrados en estos procesos. En términos del Cu, experimentos sobre la respuesta transcripcional global realizados a la fecha, a excepción de los elementos involucrados directamente en la homeostasis del metal, no describen mecanismos regulatorios capaces de controlar o coordinar la expresión de genes activados o reprimidos frente a la exposición al estímulo, o cuales podrían ser las sub-redes involucradas de respuesta al Cu. En este contexto, la presente tesis de doctorado propuso identificar y caracterizar posibles sub-redes transcripcionales involucradas en la adaptación celular al Cu, siendo el propósito fundamental del trabajo determinar dentro de una red compleja de interacciones, cuales de ellas son capaces de responder frente a la exposición al metal, sugiriendo conectividades entre los diferentes

elementos, describiendo estructuras y proponiendo dinámicas de coordinación en la regulación transcripcional sensibles al estímulo.

4. *Enterococcus faecalis* como modelo biológico elegido

En primera instancia, para abordar la construcción de una red de regulación a escala global, no sólo se debe tener un alto grado de conocimiento de los elementos que van a constituir parte de la red y las consiguientes sub-redes (factores de transcripción, sitios de unión y genes blanco), además es necesario manejar en la mayor medida posible, los estímulos involucrados en la respuesta, en este caso la exposición a Cu^{29} . Estos dos requerimientos ponen de manifiesto que la elección de un modelo biológico implica básicamente elegir un organismo capaz de integrar tanto las factibilidades para la construcción de la red así como controlar el estímulo ambiental (en este caso la concentración extracelular de Cu), lo que permitirá finalmente estudiar e identificar la coordinación que se establece entre estos elementos para entender una respuesta transcripcional a nivel global.

Si bien, contamos con el modelo de redes transcripcionales de *S. cerevisiae*, el trabajar con organismos eucariontes actualmente, presenta una alta complejidad a la hora de abordar una pregunta en esta línea de investigación, producto de: mecanismos complementarios a la acción de factores de transcripción involucrados en la expresión génica (modificación de histonas, metilaciones y fosforilaciones del DNA, co-activación, etc.), diversos factores de transcripción capaces de regular la expresión de un gen (combinatoria), múltiples modificaciones tanto post-transcripcionales como traduccionales y la compartimentalización intracelular, dificultan en gran medida tanto la identificación de sub-redes activadas por el estímulo Cu como el adjudicar la

acción de la red mediada por los factores de transcripción sobre la expresión de los genes. Todo esto, sumado al tiempo requerido para trabajar a nivel sistémico con grandes volúmenes de información, restringen la elección de un organismo eucarionte como modelo de estudio dado los márgenes de tiempo establecidos en una tesis de doctorado.

Por el contrario, la utilización de modelos procariontes, presenta grandes ventajas las que radican principalmente en el número total de genes, dado el tamaño de los genomas bacterianos significativamente menor (órdenes de magnitud) en comparación a organismos eucariontes; además cerca del 70 % de estos comparten al menos una vez la misma región de regulación transcripcional (genes co-regulados dado su organización en operones)³⁰, la existencia actualmente de una gran cantidad y diversidad de genomas bacterianos secuenciados lo cual facilita en gran medida la búsqueda de potenciales elementos reguladores. Todas estas características, presentan a un modelo procarionte como una gran alternativa de elección.

Tal como se mencionó anteriormente, la red de regulación transcripcional de *E. coli* cuenta con una alta información y caracterización de las conectividades entre los factores de transcripción y sus respectivos genes blanco. Sin embargo, en términos de elegir un modelo que presente un fácil manejo y estudio del estímulo Cu, *E. coli* al pertenecer al grupo de las Gram negativas, cuenta con mecanismos homeostáticos independientes los cuales van a controlar las concentraciones y flujos de Cu en el periplasma y el citoplasma de la bacteria³¹. A nivel transcripcional, el sistema de dos componentes PcoRS controla la expresión de proteínas involucradas en la detoxificación de niveles extremos de Cu en el periplasma, la proteína CueR activa la expresión de la ATPasa encargada del eflujo del metal desde el citoplasma hacia el periplasma, además un tercer mecanismo de detoxificación de características híbridas regulado por el complejo transcripcional CusRS, que podría estar involucrado en la detoxificación tanto

del citoplasma como del periplasma en la bacteria. Pensando en un efecto directo y posterior interpretación del efecto del metal sobre la expresión génica, la compartimentalización que se establece en *E. coli* a nivel de citoplasma y periplasma, así como la opción de precisar la implicancia de esta regulación transcripcional en la protección contra el metal, dificultan la elección de este organismo como modelo de estudio.

Por el contrario, el operon *cop* da cuenta del modelo más simple de homeostasis y regulación transcripcional descrito a la fecha³². Si bien, el mayor grado de conocimiento sobre el funcionamiento de este sistema se encuentra en *Enterococcus hirae*, esta bacteria al no contar con su genoma secuenciado limita significativamente las posibilidades de construir una red de regulación a nivel global. Con el propósito de identificar nuevos componentes involucrados con la homeostasis del metal y considerando la información disponible sobre la homeostasis de Cu en *E. hirae* presente en la literatura, nuestro laboratorio en el año 2010 presentó el modelo del operon *cop* en *Enterococcus faecalis* (bacteria filogenéticamente más cercana a *E. hirae* con genoma secuenciado)³³. Este operon está compuesto por los genes *copY*, *copA* y *copZ*; frente a diversos tratamientos con Cu, transcritos para estos genes exhiben un aumento en su abundancia cercana a 60 veces en comparación al medio basal, evidenciando la respuesta frente al metal a nivel de la expresión de estos genes, posiblemente regulados a través de CopY. En este mismo trabajo, además mostramos la respuesta transcripcional global (microarreglo) de la bacteria frente a la exposición a Cu. Los resultados indicaron que un total de 234 genes cambiaron su abundancia transcripcional frente al tratamiento, los cuales codifican para diferentes procesos metabólicos (incluyendo homeostasis de Cu y respuesta frente a ROS). Sumado a este trabajo, recientemente en este mismo modelo, publicamos un nuevo artículo donde describimos un nuevo

componente de respuesta al metal involucrado en la homeostasis del Cu (CutC), el cual posiblemente podría estar complementando la acción del operon *cop* en *E. faecalis*³⁴.

En resumen, considerando la disponibilidad de un genoma secuenciado el cual permite identificar, seleccionar y clasificar los elementos que van a constituir parte de la red, el conocimiento adquirido del operon *cop*, el efecto del Cu tanto a nivel celular como a nivel transcripcional global (microarreglo) y la experiencia obtenida en nuestro laboratorio para manejar la bacteria, sumado al desafío de construir una red de regulación transcripcional y brindar a la comunidad científica un nuevo modelo bacteriano dotado de una propuesta de redes, propusimos a *E. faecalis* como el modelo biológico adecuado para abordar la hipótesis de la presente tesis.

En atención a la evidencia presentada se planteó la siguiente hipótesis de trabajo:

5. Hipótesis de trabajo

Genes que se activan frente a la exposición a cobre y que codifican para proteínas que participan en diversos procesos celulares, son regulados por un conjunto común de factores de transcripción.



6. Objetivo general

Identificar redes de regulación transcripcional activadas en respuesta a Cu, compuestas por factores de transcripción capaces de coordinar la expresión de genes que codifican para componentes asociados a la protección directa contra el metal (genes del operon *cop*), respuesta a estrés, metabolismo basal, producción de energía, reparación celular, transporte, etc. Así como, verificar si esta regulación es determinante en la resistencia contra el metal en la bacteria *Enterococcus faecalis*.

7. Objetivos específicos

Objetivo 1: Construir *in silico* una posible red de regulación transcripcional activada en respuesta a la exposición a Cu.

Mediante el uso de diferentes algoritmos bioinformáticos, se construyó y caracterizó topológicamente una red de regulación transcripcional en el genoma de *E. faecalis*, la cual contiene factores de transcripción conectados con los genes codificados en la bacteria a través de la presencia de un sitio de unión en sus regiones promotoras. Posteriormente, se identificó dentro de la red el sub-conjunto de genes activados en respuesta a Cu (sub-red).

Objetivo 2: Determinar la participación del(los) factor(es) de transcripción en la regulación de la expresión de los componentes de la red inducidos por Cu.

Se generaron criterios de selección que permitieron escoger dentro de la sub-red activada por Cu, factores de transcripción a los que se verificó conectividades con sus respectivos operones blancos mediante la generación y caracterización de cepas mutantes.

Objetivo 3: Evaluar la dependencia de la regulación transcripcional mediada por los factores de transcripción en la resistencia a Cu.

Se realizaron diferentes protocolos microbiológicos de exposición a Cu que lograron determinar si, la ausencia de la regulación transcripcional impactó sobre la resistencia al metal, comparando principalmente la viabilidad celular y el contenido intracelular de Cu entre las cepas silvestre y mutantes.

MATERIALES Y METODOS

1. MATERIALES

1.1. Cepas y condiciones de cultivo

Las características de las cepas utilizadas en este trabajo se detallan en la tabla 1. Cepas de *E. faecalis* se cultivaron en medio mínimo N líquido (1 % (p/v) glucosa (Merck, New Jersey, EUA), 1 % (p/v) peptona (Oxoid, Basingstoke, Inglaterra), 1 % (p/v) fosfato monohidrógeno de sodio (Merck), 0,5 % (p/v) extracto de levadura (Oxoid)), Caldo BHI (BD, New Jersey, EUA) y Caldo THB (BD). Para el crecimiento de *E. coli* se utilizó medio Luria-Bertoni líquido (LB) (1 % de triptona (Oxoid), 1 % de cloruro de sodio (Merk), 0,5 % de extracto de levadura (Oxoid)). Para cultivos sólidos en placa, se usó la misma composición del medio líquido correspondiente adicionando agar bacteriológico al 1,5 % (p/v) (Oxoid). La selección de colonias de *E. coli* transformada con el vector pGEM-T (Promega, Wisconsin, EUA) se efectuó agregando al medio sólido el antibiótico ampicilina (Sigma-Aldrich, Illinois, EUA) y X-Gal (Gold Biotechnology, St. Louis, EUA) a una concentración de 100 µg/mL y 200 µg/mL respectivamente. En cultivos de células transformadas con el vector pCJK47-Gen se agregó el antibiótico gentamicina (Sigma-Aldrich) a una concentración de 5 µg/mL para *Enterococcus* y 200 µg/mL para *E. coli*. Para cultivos de la cepa mutante para el gen *lexA* de *E. faecalis*, fue necesario adicionar al cultivo el antibiótico eritromicina (50 µg/mL).

Los cultivos sólidos se efectuaron a una temperatura de 37 °C en una estufa Precision Incubator 300 (Mettler, Schwabach, Alemania). Los cultivos líquidos se realizaron en una

estufa L5I-7016R (Latech, Salt Lake, EUA) con agitación de 140 r.p.m. a 37 °C. Las cepas se almacenaron a -80 °C en el medio líquido correspondiente a 80 % (v/v) y glicerol 20 % (v/v). Cuantificaciones de absorbancia de cultivos líquidos y ácidos nucleicos se realizaron en un espectrofotómetro MBA2000-UV/VIS (Perkin Elmer, New York, EUA) a una densidad óptica (DO) de 600 nm y 240/260/280 nm respectivamente. Las verificaciones finales de las mutantes se realizaron vía secuenciación Sanger por la empresa GeneWiz (New Jersey, EUA).

Tabla 1. Cepas bacterianas y plásmidos.

Cepas y plásmidos	Genotipo relevante	Referencia
Cepas		
<i>E. faecalis</i>		
OG1RF	Cepa silvestre <i>E. faecalis</i> , receptora de conjugación	35
CK111	Cepa dadora conjugativa	Laboratorio Dra. Barbara E. Murray*
$\Delta copY$	Mutante cepa <i>E. faecalis</i> OG1RF, deleción por conjugación del gen <i>copY</i>	³⁶
Δfur	Mutante cepa <i>E. faecalis</i> OG1RF, deleción por conjugación del gen <i>fur</i>	Este trabajo
$\nabla lexA$	Mutante cepa <i>E. faecalis</i> OG1RF, inserción del vector Tn917 en gen <i>lexA</i>	37
<i>E. coli</i>		
DH5 α	Clonamiento del vector pGEM-T, cepa comercial Invitrogen	38
EC1000	Clonamiento del vector pCJK- Δfur	39
Plásmidos		
pGEM-T	Vector comercial, Promega	40
pGEM-T- Δfur	Vector pGEM-T contenedor del segmento homólogo para construcción de mutante para el gen <i>fur</i>	Este trabajo
pCJKgen	Contiene sitio oriTpCF10, gen <i>lacZ</i> , gen <i>pheS</i> (selección), gen <i>repA</i> (replicación), gen <i>aph2</i> (resistencia a gentamicina).	41
pCJKgen- Δfur	Vector pCJKgen contenedor del segmento homólogo para construcción de mutante para el gen <i>fur</i>	Este trabajo

* Division of Infectious Diseases, University of Texas Medical School. Houston, Texas, EUA.

1.2. Sales, soluciones y reactivos

Sales utilizadas en los ensayos de exposición al metal (CuSO_4) y posteriores soluciones para el lavado celular (PBS 1 X, solución 1: NaCl 0,15 M, solución 2: glicina 50 mM a pH 3,5 y solución 3: EDTA 1,0 mM) fueron adquiridas en la empresa Merck. Las soluciones y medios se prepararon en agua destilada y desionizada (TCL, Santiago, Chile) y posteriormente esterilizada mediante autoclave LS-1 (Rexall, Mississauga, Canadá) durante 30 min. Todos los kits y enzimas utilizados en los ensayos de biología molecular se adquirieron en las empresas Invitrogen (California, EUA), Promega, Qiagen (Duesseldorf, Alemania) y Roche (Basilea, Suiza).

1.3. Partidores

Los partidores utilizados tanto para las reacciones de PCR convencional como en tiempo real (ANEXO 1), se sintetizaron por Alpha DNA (Quebec, Canadá) y Sigma-Aldrich. El diseño de éstos se llevó a cabo mediante el Programa Primer3plus⁴², utilizando como templatado el genoma secuenciado de *E. faecalis* V583⁴³.

1.4. Algoritmos bioinformáticos

La predicción de la organización de los genes dentro de *E. faecalis* (asignación de operones) se realizó a través del uso de los programas SHOPS⁴⁴ y Uber⁴⁵. La identificación y clasificación de los posibles factores de transcripción codificados en la bacteria, se llevó a cabo

automáticamente mediante Blastp bi-direccionales⁴⁶ y asignación de dominios funcionales a través del programa InterProScan⁴⁷, éste último permitió rastrear información de motivos descritos en diferentes familias de proteínas recogidas desde las bases de datos de Pfam⁴⁸, Interpro⁴⁹ y del European Bioinformatic Institute (EBI)⁵⁰. La construcción de la red de regulación transcripcional se llevó a cabo utilizando los algoritmos MOTIF SAMPLER-SCANNER⁵¹ y CONSENSUS-PATSER⁵². Como elementos de búsqueda, se rastrearon un total de 215 matrices de peso probabilísticas (MPPs) para sitios de unión bacterianos⁵³, información obtenida por recopilación de secuencias desde plataformas tales como: RegulonDB⁵⁴, TractorDB⁵⁵ y DBTBS⁵⁶ y posteriores curaciones manuales basados en referencias experimentales. Esta información esta presente en una base de datos previamente generada por el Laboratorio de Bioinformática y Matemática del Genoma (LBMG). Como templado de la búsqueda de las MPPs, se utilizó el genoma secuenciado de *E. faecalis* V583⁴³. Para la visualización y análisis topológicos posteriores de la red construida, se utilizó las plataformas ARTEMIS⁵⁷, CYTOSCAPE⁵⁸, Microsoft Excel 2010⁵⁹ y el paquete Igraph contenido en el software R⁶⁰. Para la selección de genes diferencialmente activados por Cu, se utilizaron datos de expresión génica global publicados previamente por nuestro laboratorio³⁶. El cruce de ésta información con la lista de genes incluidos en la red de regulación generada *in silico* se realizó mediante el programa Microsoft Access 2010⁵⁹. Todos los alineamientos de las secuencias nucleotídicas se realizaron con el algoritmo ClustalW v1.4 disponible a través del editor de secuencias BIOEDIT⁶¹. La generación de posteriores Logos se efectuó mediante el programa WEBLOGO disponible en línea (<http://weblogo.berkeley.edu/>)⁶².

Todas las herramientas bioinformáticas se utilizaron en forma estándar, basadas en las referencias citadas y siguiendo las recomendaciones correspondientes indicadas por los autores.

2. MÉTODOS

2.1. Análisis *in silico*

2.1.1. Identificación de elementos que van a constituir la red de regulación transcripcional *in silico*

La construcción de la red de regulación transcripcional mediada por factores de transcripción requiere en primera instancia, identificar cuales van a ser los componentes constituyentes, los que corresponden básicamente a los elementos reguladores (factores de transcripción) y el total de genes presentes dentro del genoma de la bacteria (blancos de regulación). Como se mencionó anteriormente, el modelo de estudio utilizado fue la bacteria *E. faecalis*, la anotación completa del genoma de la cepa V583 cuenta con un total de 3113 posibles secuencias codificantes (CDS, posibles genes)⁶³.

En primer lugar se determinó la organización genómica que presenta *E. faecalis*, a fin de identificar grupos de genes que pueden transcribirse como una sola unidad transcripcional y que por tanto podrían compartir la misma región promotora (definición operacional de operones⁶⁴). De esta forma se constituyó el conjunto de secuencias que se utilizaron en la búsqueda de los sitios de unión (ver detalles en sección 2.1.2.) que a su vez permitieron establecer la hipótesis de conectividades entre factor de transcripción y el conjunto de operones de nuestra red.

2.1.2. Construcción *in silico* de la red de regulación transcripcional mediada por factores de transcripción (escala global)

Determinados los elementos reguladores (agrupados en familias de factores de transcripción) y la organización de los genes en operones, se procedió a construir la red de regulación transcripcional. Esta tarea se realizó en colaboración con el Laboratorio de Bioinformática y Matemática del Genoma (LBMG) dirigido por el Dr. Alejandro Maass en el Centro de Modelamiento Matemático (CMM) de la Universidad de Chile. Se determinó en el genoma de *E. faecalis* V583, el número de secuencias que coinciden con aquellas descritas para diferentes sitios de unión de bacterias. Para ello, las MPPs correspondientes a cada sitio de unión, fueron generadas mediante alineamientos de archivos FASTA los cuales contienen en promedio 10 secuencias nucleotídicas asignadas como posibles sitios de reconocimiento para factores de transcripción bacterianos. Las secuencias contenidas en cada archivo, fueron previamente agrupadas de acuerdo a su procedencia (obtenidas desde una sola especie o de todo un orden bacteriano) y reconocimiento (por un factor en particular o una familia de factores de transcripción). La base de datos contiene un total de 362 MPPs (91 familias de factores de transcripción representadas), en este trabajo se consideraron, solo aquellas matrices asignadas a un factor de transcripción el cual tuviese un posible homólogo por identidad dentro del genoma de *E. faecalis* (149 en total distribuidas en 91 familias de factores). Para esto, se cruzó la clasificación por factores presente en la base de datos con la predicción de factores de transcripción obtenida en el punto 2.1.1., obteniendo un conjunto final de 149 MPPs de búsqueda. Posteriormente, cada una de ellas fueron rastreadas 300 pb río arriba y 50 pb río abajo

del inicio de cada operon de *E. faecalis* V583 mediante dos protocolos automatizados de búsqueda (MOTIF SAMPLER-SCANNER y CONSENSUS-PATSER).

El archivo final resultante, integra los elementos reguladores y los genes (agrupados en operones) conectados entre se en base a la presencia de un posible sitio de unión en su región promotora. Esta información fue registrada en un archivo de texto en formato genebank⁶³ y formato .xls (Microsoft Excel), para luego ser analizado en plataformas de visualización y acceso a la información de genomas y redes biológicas (ARTEMIS y Cytoscape).

2.1.3. Identificación de una red de regulación transcripcional activada por Cu y selección de regulones para verificaciones conectividades *in vivo*

Durante el desarrollo de este trabajo de tesis, nuestro laboratorio ha publicado una serie de experimentos utilizando microarreglos sembrados con casi la totalidad de los genes anotados de *E. faecalis* V583 (Nimblegen, N° de catálogo A4359- 00-01, GSE20453)³⁶. Brevemente, los vidrios sembrados fueron hibridados con cDNA sintetizado desde RNA genómico, aislado de la bacteria crecida en medio N estándar de cultivo (referencia) y un medio N suplementado con 0,5 mM de CuSO₄ (tratamiento) hasta una fase exponencial intermedia (DO_{600nm} = 1). Del total de genes diferencialmente expresados en forma significativa producto de la exposición a Cu, se consideraron lo 109 aumentaron su abundancia de transcrito. Los operones que contienen estos genes fueron rastreados manualmente en la red global construida en el punto 2.1.2. Se incluyeron en la red, operones que presentaban al menos uno de los 109 genes detectados, descartándose aquellos operones con un segundo gen que cambio su expresión significativa e inversa. El total

de genes inducidos por el metal dentro de la red transcripcional global, constituyó finalmente la red de regulación transcripcional inducida por Cu.

Considerando el alto volumen de información contenida dentro de la red, se aplicaron 2 filtros que pretenden seleccionar regulones para su verificación de conectividades en la bacteria.

i) Filtro por conectividades: Se seleccionaron regulones que presentaron un alto número de conectividades. Nuestro argumento considera que este tipo de redes arrastran un alto número de conectividades falsas positivas, por lo tanto elegir un número alto aumenta la probabilidad de encontrar factores de transcripción conectados con algunos de sus blancos predichos.

ii) Redundancia en el genoma: Se pretende validar la conectividad dentro de un regulon directamente en la bacteria, a través de la cuantificación transcripcional de sus genes en un escenario ausente de regulación (cepa mutante para su correspondiente factor de transcripción versus la cepa silvestre). Dado este diseño experimental, la redundancia se presenta como un problema a la hora de verificar los cambios transcripcionales, puesto que un segundo regulador con características similares podría anular el fenotipo esperado en la cepa mutante. Es por esto, que se seleccionaron regulones en los cuales el factor solo posea una copia en el genoma de *E. faecalis*.

2.1.4 Análisis topológicos y generación de logos

Se utilizaron protocolos estandarizados, previamente descritos en la literatura para examinar las diferentes propiedades topológicas y estructurales de la red a escala global e inducida pro Cu construidas en *E. faecalis*, estas fueron:

i) Distribución del grado de salida P(K): Este valor da la fracción de genes conectados con los factores de transcripción, obtenido al contar el número de genes en la red N (K) que tienen K = 1, 2, 3 ... n de conectividades dividido por el total conectividades.

ii) Tipo de conectividades: La identificación del número de orígenes, loops de alimentación, módulos de entrada y autorregulación dentro de la red se determinó automáticamente a través de la aplicación del paquete Igraph contenido en el programa R, ingresando como elemento de entrada la red transcripcional en formato .xls.

La construcción de Logos se realizó mediante el programa WEBLOGO, utilizando las secuencias nucleotídicas de los posibles sitios de unión para cada regulon identificado.

2.2. Biología Molecular

2.2.1. Reacción en cadena de la polimerasa (PCR)

La determinación de la calidad del RNA en todas las muestras utilizadas y la verificación de los constructos utilizados en la generación de cepas mutantes de la cepa OG1RF de *E. faecalis* se efectuaron a través de experimentos de PCR convencional. En ambos casos, la amplificación se realizó en una mezcla final de 20 μ L, compuesta por 8 ng/ μ L de cDNA templado, tampón PCR 1 X, MgCl₂ 1,5 mM, dNTPs 0,2 mM, partidores 0,2 pM c/u, TAQ Polimerasa 5 U (Invitrogen). La reacción se llevó a cabo en un termociclador PTC-100 (MJ Research, Massachusetts, EUA), siendo el programa de reacción de PCR utilizado el que se describe a continuación: a) Desnaturalización inicial a 94 °C por 10 mins, b) 30 ciclos de desnaturalización a 94 °C por 30 s, asociación de los partidores por 30 s con T_m específica para cada pareja de

partidores utilizados (ANEXOS 1), extensión a 72 °C por 60 s, y c) extensión final, 72 °C durante 10 mins. En los ensayos de PCR convencional en colonia, el cDNA se reemplazó por 1 µL de una resuspensión de una colonia bacteriana en 20 µL de agua libre de nucleasas.

2.2.2. Electroforesis

El fraccionamiento de muestras de DNA y cDNA se realizó en geles de agarosa al 1 % p/v en tampón TAE 1 X (ácido acético 20 mM, Tris 40 mM, EDTA 2 mM, pH 8,5), utilizando la misma solución como tampón electroforético. Se adicionó bromuro de etidio al 0,1 % v/v con el propósito de teñir los geles para su posterior visualización.

Muestras de RNA (1 µg aprox.), previo a la electroforesis se desnaturalizaron en una solución de formamida 36 % (v/v), formaldehído 4 % (v/v), MOPS 0,8 X y bromuro de etidio 0,4 % (v/v), a 70 °C por 15 min, luego fueron enfriadas en hielo por 5 min. Finalmente, las muestras se fraccionaron en geles de agarosa al 1,2 % p/v en condiciones desnaturalizantes, adicionando a la muestra formaldehído al 7 % (v/v) y MOPS 1 X. Como tampón electroforético se utilizó una solución MOPS 1 X.

La inspección visual de las muestras de DNA y RNA se realizó en un transiluminador de luz UV modelo TFX 20.M (Vilber Lourmat). Las fotografías de los geles se capturaron a través de una cámara digital Dc290 (Kodak).

2.2.3. Construcción y verificación de cepas mutantes para factores de transcripción en *E. faecalis* OG1RF

Gran parte de este objetivo se realizó en el laboratorio "Division of Infectious Diseases" dirigido por la Dra. Barbara Murray en la Universidad de Texas (Houston, EUA) durante un período de 4 meses. La estadía tuvo como propósito principal construir y verificar las mutantes para los genes que codifica para los factores de transcripción seleccionados en el punto 2.1.3. Los partidores utilizados en cada uno de los pasos de construcción y verificación de las mutantes se describen en la tabla 2.

i) Mutante *fur*: La generación de la cepa mutante para *fur* (EF1575) se llevó a cabo mediante la técnica de delección de genes por doble recombinación usando el sistema PheS⁴¹, estrategia que permite generar delecciones específicas en el genoma. El protocolo, estandarizado en laboratorio de la Dra. Murray, consistió primero en amplificar por PCR convencional dos segmentos correspondientes a la región río arriba (up, partidores *fur_up*) y río abajo (down, partidores *fur_down*) del gen *fur*, al partidor antisentido del segmento up se adicionó la secuencia reversa complementaria (10 pb) de la región 5' del segmento down. Del mismo modo, se adicionó al partidor sentido de la región down, la secuencia reversa complementaria de la región 3' del segmento up. La adición de estos segmentos complementarios, permitió mediante una nueva reacción de PCR, en la cual se utilizaron ambos segmentos (up y down) como templado y los partidores sentido de la región up y antisentido de la región down (*fur_up_s* y *fur_down_a*), obtener un amplicón que corresponde a la fusión de ambos segmentos (up/down).

El amplicón obtenido fue ligado al vector comercial pGEM-T (Promega) y posteriormente clonado en la bacteria *E. coli* DH5 α tomando en cuenta las recomendaciones del fabricante. El

vector resultante pGEM-T- Δfur fue purificado mediante el kit Wizard Plus Minipreps (Promega) y luego secuenciado para verificar su correcta clonación. Luego, se digirió el vector con las enzimas de restricción NotI y Sall (Promega), el fragmento *fur* digerido fue ligado en el vector pCJKgen previamente tratado con ambas enzimas. El vector pCJKgen- Δfur fue clonado en cepas de *E. coli* EC1000, seleccionando los transformantes con el antibiótico gentamicina (10 mg/mL), paso que permitió aumentar la eficiencia en la recuperación de masa del plásmido. El constructo luego fue purificado (Wizard Plus Minipreps, Promega) y electro-transformado en la cepa OG1RF de *E. faecalis* (200 mg/mL de gentamicina como medio de selección) utilizando el equipo MicroPulser Electroporator #165-2100 (BioRad, California, EUA). El vector pCJKgen no puede replicar en esta cepa, de tal forma que las bacterias resistentes al antibiótico son aquellas en que el vector se ha integrado al cromosoma. El evento de segunda recombinación y pérdida del plásmido es favorecido por traspaso de las células OG1RF, que contienen el constructo clonado en pCJK47gen, a un medio con p-Cl-fenilalanina. El vector pCJK47-Gen lleva un alelo dominante para una mutante del gen que codifica para la fenilalanina transferasa. Esta mutación permite aceptar compuestos con alta densidad electrónica en la posición *para* del anillo aromático de la fenilalanina. La incorporación de este tipo de compuestos (en este caso p-Cl-fenilalanina), a la cadena aminoacídica en síntesis, es altamente tóxica para la célula. Esta propiedad conferida por el vector permite seleccionar a mutantes de OG1RF que han perdido el plásmido y el gen blanco ya que estas bacterias son favorecidas para crecer en un medio con p-Cl-fenilalanina. Por lo tanto, las potenciales cepas mutantes, fueron aquellas que no contienen el plásmido y que perdieron el gen de interés. Finalmente el genotipo de las cepas mutantes para *fur* se verificó con ensayos de PCR y secuenciación (partidores *dfur*).

ii) Mutante *lexA*: La cepa mutante de *E. faecalis* OG1RF para este gen fue donada directamente por el laboratorio de la Dra. Danielle Garsin durante la pasantía realizada en EUA. Esta mutante fue generada por la inserción del transposón Tn917 en la región codificante para *lexA*³⁷. La mutación en este caso, se verificó mediante PCR convencional y posterior secuenciación (partidores *dlexA*), lo que además permitió precisar la posición en la cual se interrumpió el gen.

iii) Mutante *copY*: Esta mutante fue previamente construida y publicada por nuestro laboratorio en colaboración con el de la Dra. Murray³⁶, utilizando de igual forma método de doble recombinación homóloga empleado en la mutante para el gen *fur*. Del mismo modo que para las demás cepas mutantes utilizadas en esta tesis, se verificó la mutación por PCR convencional y secuenciación (partidores *dcopY*).

2.3. Ensayos de expresión génica

2.3.1. Extracción de RNA total

Para la verificación de datos obtenidos previamente en el microarreglo de *E. faecalis* expuesta a Cu, la población se ajustó desde un cultivo con la cepa OG1RF crecido durante toda la noche en medio N, una concentración de $DO_{600nm} = 0,05$ final en dos matraces (A control, B tratamiento con Cu) con medio N fresco. Cuando los cultivos alcanzaron una $DO_{600nm} = 0,5$, se suplementó el matraz B con $CuSO_4$ hasta alcanzar una concentración de 0,5 mM. Luego, de 2 horas de exposición al metal, se tomó una alícuota independiente (6 mL de cultivo con una $DO_{600nm} = 2$) de cada matraz (A y B), el experimento se realizó 3 veces en días diferentes, lo que

constituyó finalmente un total de 6 muestras, correspondientes a 3 réplicas biológicas para cada uno de los matraces control y tratamiento.

Del mismo modo, para los ensayos de verificación de cambios transcripcionales entre la cepa silvestre y las mutantes generadas para los factores de transcripción seleccionados en el punto 2.1.3., se ajustó desde un cultivo de *E. faecalis* OG1RF y mutante crecidas durante toda la noche en medio N una concentración de $DO_{600nm} = 0,05$ final en dos matraces (A silvestre, B mutante) con medio N fresco. Luego de alcanzar ambos cultivos una $DO_{600nm} = 0,5$ (réplica temporal de la extracción de RNA del experimento con Cu), se tomó una alícuota independiente (6 mL de cultivo con una $DO_{600nm} = 2$) de cada matraz (A y B). Al igual que en el experimento anterior, el ensayo se realizó 3 veces en días diferentes, constituyendo finalmente un total de 18 muestras, correspondientes a 3 réplicas biológicas para cada mutante con su respectiva muestra silvestre.

Posteriormente, todas las alícuotas obtenidas se centrifugaron a 14.000 r.p.m durante 2 min. El pellet celular resultante se lavó 3 veces en forma sucesiva con una solución tampón de fosfato salino pH 7,4 (PBS). Luego se procedió con la extracción de RNA, para ello se utilizó el kit de extracción de RNA de Qiagen, previa adición de lisozima a una concentración final de 10 mg/mL por 20 min a 37 °C al pellet celular obtenido luego de los lavados.

Remanentes de DNA en las muestras, fueron tratados con 10 U de DNAasa (Qiagen), siguiendo las recomendaciones del fabricante. Finalmente, el RNA se resuspendió en 20 µL de agua libre de nucleasas para sus posteriormente cuantificaciones y verificaciones de calidad e integridad.

2.3.2. Cuantificación y calidad del RNA

La integridad de la muestra se verificó a través de electroforesis en geles desnaturalizantes (punto 2.2.2), observando como resultado de la migración, dos bandas características correspondientes a las unidades ribosomales 16S y 23S. La posterior cuantificación se realizó mediante espectrofotometría (punto 1.1), calculando las razones de absorbancia entre 260/280 nm y entre 260/240 nm, criterio que permitió determinar una eventual presencia de proteínas y fenoles en la muestra respectivamente. Contaminaciones con DNA se controlaron a través de PCR convencional (punto 2.2.1), utilizando arbitrariamente la pareja de partidores *gdh_s/gdh_a* (tabla 2), esperando no obtener ninguna amplificación en las muestras de RNA. Como control positivo de la amplificación se realizó un PCR en colonia utilizando la misma mezcla de partidores empleada para las muestras de RNA.

2.3.3. Síntesis de cDNA

La síntesis de cDNA de doble hebra se realizó *in vitro*. La reacción de transcripción reversa se realizó a partir de 2 µg de RNA total y 1 µL partidores al azar o *random primers* (3 µg/µL) en un volumen de 15 µL. La reacción se incubó por 5 min a 70 °C. Terminado este tiempo, se agregó 1,25 µL de dNTP (10 nM), 0,5 µL de RNAsin (40 U/µL), 5 µL de M-MLV buffer (5 X) y 1 µL de la enzima Transcriptasa Reversa M-MLV (Promega 200 U/µL). Finalmente, se adicionó agua libre de nucleasas hasta completar un volumen de 25 µL. La mezcla resultante se incubó durante 1 hora a 37 °C para luego ser almacenadas a -20 °C hasta su posterior uso en ensayos de qPCR.

La cuantificación del cDNA se efectuó a través de espectrofotometría (punto 1.1), verificando su calidad mediante la amplificación arbitraria del gen *gdh* por PCR convencional (punto 2.2.1).

2.3.4. PCR cuantitativo en tiempo real (qPCR)

Se determinó mediante qPCR, la abundancia relativa de los transcritos para los genes diferencialmente expresados seleccionados desde los microarreglos (activados por Cu dentro de la red) y los genes seleccionados como posibles blancos de regulación dentro de la red inducida por Cu (punto 2.1.3). Cada reacción se llevó a cabo utilizando el kit qPCR Supermix UDG (Invitrogen) en capilares de vidrio a partir de 50 ng de cDNA en un volumen final de 10 μ L. La mezcla final contenía además BSA 1 X, 0,1 pM de cada partidor y 5 μ L de enzima Platinum SYBR Green (Invitrogen). El protocolo de amplificación se ingresó al termociclador a través del programa Light Cycler v3.5 (Roche) y se desglosa en: a) activación de la enzima, 95 °C por 10 mins, b) amplificación, 40 ciclos de: 1) desnaturalización, 95 °C por 3 min, 2) asociación de los partidores, por 10 segundos a la temperatura correspondiente para cada pareja de partidores incluida en la tabla 2) extensión, por 15 segundos a 72 °C y 4) curva de desnaturalización, incremento de la temperatura a una tasa de 0,1 °C/segundo desde la temperatura de asociación de los partidores hasta 99 °C, etapa final que permitió verificar la ausencia de productos inespecíficos en la reacción. Todos los transcritos se cuantificaron por triplicado para cada una de las muestras de cDNA. La obtención de datos se realizó mediante el programa LightCycler Software v3.5 (Roche).

Los resultados generados se expresaron como la razón de cambio entre la abundancia normalizada del transcrito para el gen pregunta en la muestra control y la muestra tratada, utilizando como referencia la abundancia de un transcrito que no varía su abundancia en las diferentes condiciones analizadas (*housekeeping*), expresados según la siguiente ecuación⁶⁵:

$$\text{Ec 1: Razón de cambio normalizado} = \frac{\text{Ef gen en estudio}^{(\Delta C_p)}}{\text{Ef normalizador}^{(\Delta C_p)}}$$

Ef = Eficiencia promedio de la pareja de partidores durante todas las reacciones

ΔC_p = Diferencia del ciclo en que se inicia la detección de fluorescencia (punto de cruce o *crossing point*, C_p) entre la muestra control y muestra experimental

2.4. Protocolos de caracterización microbiológica

2.4.1. Curva de crecimiento

A partir de un cultivo líquido, crecido toda la noche en medio N a 37 °C con agitación, se ajustó un inóculo con una $DO_{600nm} = 0,05$ en un matraz control conteniendo 50 mL de medio N y en otros tres matraces experimentales con 50 mL de medio N suplementados con 0,5 y 1,0 mM de $CuSO_4$. Los cultivos bacterianos se incubaron a 37°C con agitación a 140 rpm. El crecimiento bacteriano se monitoreó por registro de absorbancia cada una hora durante las primeras 9 hs y luego una medición final a las 24 h. Este ensayo se realizó en triplicado para cada cepa mutante comparando en paralelo su crecimiento con la cepa silvestre OG1RF de *E. faecalis*. Los valores

de densidad óptica, se utilizaron en el cálculo del tiempo generacional requerido para duplicar la masa de bacterias en un lapso determinado, a través de la siguiente ecuación:

$$\text{Ec 2: } \text{Tiempo generacional} = \frac{\text{Log (B2)} - \text{Log (B1)}}{t2 - t1} \times 95,6$$

B1 = Densidad óptica cuantificada a tiempo t1.

B2 = Densidad óptica cuantificada a tiempo t2.

t1 = Tiempo inicial, tiempo arbitrario en el cual el cultivo se encuentra en fase exponencial

t2 = Tiempo posterior al tiempo t1 (final), en el cual el cultivo aún encuentra en fase exponencial

95,6 = Factor de conversión Ln / Log

2.4.2. Contenido intracelular de cobre

Se determinó el contenido intracelular de Cu para las cepas silvestre y mutantes, utilizando las condiciones de crecimiento, concentración y tiempo de exposición al metal que se definieron para los experimentos de extracción de RNA y curvas de crecimiento. Siguiendo el mismo protocolo publicado previamente por nuestro laboratorio³⁶, cuando los cultivos alcanzaron una $DO_{600\text{nm}} = 1$, se extrajo 6 mL los cuales se centrifugaron a 14.000 r.p.m. durante 1 min. Eliminado el sobrenadante, el exceso del metal fue removido mediante 3 lavados sucesivos con las soluciones 1, 2 y 3 respectivamente, resuspendiendo finalmente las células en 1 mL de solución PBS 1X. Las muestras fueron lisadas a través de 5 pulsos de sonicación de 15 segundos, con 5 segundos de reposo entre cada pulso utilizando un sonicador modelo OMNI-

ruptor 250 (OMNI, Kennesaw, EUA). Inmediatamente lisadas las células, se centrifugaron a 14.000 r.p.m. durante 30 min a 4 °C, posteriormente se extrajo independientemente 100 µL del sobrenadante de cada tubo para cuantificación de proteínas, y otros 100 µL para digerirlos con 250 µL HNO₃ 0,1 N ultra puro (Merck) a 75 °C por 24 hs. Finalmente, luego de la digestión ácida se determinó el contenido de Cu mediante espectrometría de absorción atómica (AAS) en un horno de grafito SIMMA 6100 (Perkin Elmer), cuyo límite de detección circunda las partes por billón (ppb). Los resultados de cada medición fueron normalizados con los valores obtenidos de su cuantificación de proteínas correspondientes.

2.5. Análisis estadístico

Las réplicas técnicas y biológicas generadas en cada uno de los experimentos se detallan en el pie de las figuras o en la descripción de los protocolos cuando corresponda, donde se declara el promedio de las mediciones obtenidas con sus respectivas desviaciones estándar y análisis estadísticos aplicados.

La determinación de significancias en los experimentos de qPCR se llevó a cabo a través del algoritmo REST⁶⁶. Para los ensayos microbiológicos, se utilizó el software GRAPHPAD PRISMA trial v5.0. En ambos programas se manejaron los parámetros establecidos de base, considerando diferencias estadísticamente significativas si $p < 0,05$.

RESULTADOS

1. Construcción *in silico* de una posible red de regulación transcripcional activada en respuesta a la exposición a Cu.

1.1 Identificación y clasificación de los componentes de la red

Las redes en general se componen de nodos que interactúan entre si a través de bordes o enlaces⁶⁷. En el caso de redes de regulación transcripcional mediadas por factores de transcripción, es posible identificar dos tipos de nodos, los factores de transcripción o elementos reguladores y los genes blanco de estos factores o elementos regulados (blancos de regulación). Las conectividades establecidos entre ambos nodos corresponden a secuencias nucleotídicas regulatorias capaces de ser reconocidas por los factores de transcripción (sitios de unión), presentes principalmente en las regiones promotoras de los genes blanco los cuales van a generar la conectividad entre ambos tipos de nodos²⁵.

En bacterias, los genes que van a constituir los posibles blancos de regulación se encuentran organizados en operones, lo cual implica que dos o más genes puedan compartir la misma región promotora (operon policistrónico), posibilitando que un factor de transcripción regule a través de un solo sitio de unión la expresión de más de un gen⁶⁴.

Dado esta características de los genomas bacterianos, con la intención de identificar y precisar los genes blanco que van a formar parte de la red, en primer lugar se procedió a determinar los posibles operones dentro del genoma de *E. faecalis* V583. El resultado luego de la asignación automática entregó que del total de genes codificados en la bacteria (3113)⁴³, 1663

fueron predichos como organizados en operones, 991 son del tipo monocistrón (gen único) y 672 del tipo policistrón (2122 genes contenidos) (tabla 2). Esta distribución fue comparada con resultados declarados en otras bacterias. La figura 1 da cuenta que la predicción obtenida de operones en *E. faecalis* concuerda con valores porcentuales promedios calculados para un conjunto de 550 genomas bacterianos (datos extraídos de la base de datos públicas ODB³⁰).

Obtenida la organización de los blancos de regulación, se procedió a identificar los posibles factores de transcripción codificados en el genoma de *E. faecalis* que van a constituir los elementos reguladores de la red. Mediante el uso de bases de datos actualizadas y herramientas de anotación de genomas desarrolladas en el Laboratorio de Bioinformática y Matemática del Genoma (LBMG) dirigido por el Dr. Maass, se generó una nueva anotación y clasificación de los posibles factores de transcripción presentes en la bacteria. Un total de 183 genes fueron identificados como codificantes para posibles factores de transcripción, grupo que corresponde al 5,5% del total de los genes codificados en la bacteria, porcentaje que se encuentra dentro del rango obtenido en otros genomas procariontes de tamaño similar (figura 2)⁶⁸.

Posteriormente, mediante la comparación de secuencias aminoacídicas e información de motivos funcionales de proteínas disponibles en diversas bases de datos, se logró agrupar los factores de transcripción por familias (tabla 3). El análisis de esta clasificación, da cuenta que familias de factores de transcripción con un alto número de integrantes (mayor a 8) se encuentran en menor proporción en comparación a las familias que presentan un menor número de componentes, dato que concuerda con lo descrito en bacterias y arqueas⁶⁹.

Tanto los resultados de la organización de operones como la identificación y clasificación de factores de transcripción en *E. faecalis*, nos permitió suponer para estos elementos, que no existen características particulares a nivel global dentro del cromosoma de la bacteria en relación

a lo descrito en otros organismos, punto relevante puesto que la estrategia posterior a utilizar para la construcción de la red, pretende rastrear información sobre sitios de unión para factores de transcripción descrita en otras bacterias dentro del genoma de *E. faecalis*.

Tabla 2. Distribución del total de los genes dentro de los operones identificados en el genoma de *E. faecalis*.

Número de genes por operon	Número de operones	Número de genes totales	Porcentaje de operones totales (%)	Porcentaje de genes totales (%)
1	991	991	59,59	31,83
2	342	684	20,57	21,97
3	163	489	9,80	15,71
4	66	264	3,97	8,48
5	32	160	1,92	5,14
6	29	174	1,74	5,59
7	14	98	0,84	3,15
8	7	56	0,42	1,80
9	8	72	0,48	2,31
10	3	30	0,18	0,96
11	4	44	0,24	1,41
12	2	24	0,12	0,77
13	1	13	0,06	0,42
14	1	14	0,06	0,45
total	1663	3113		

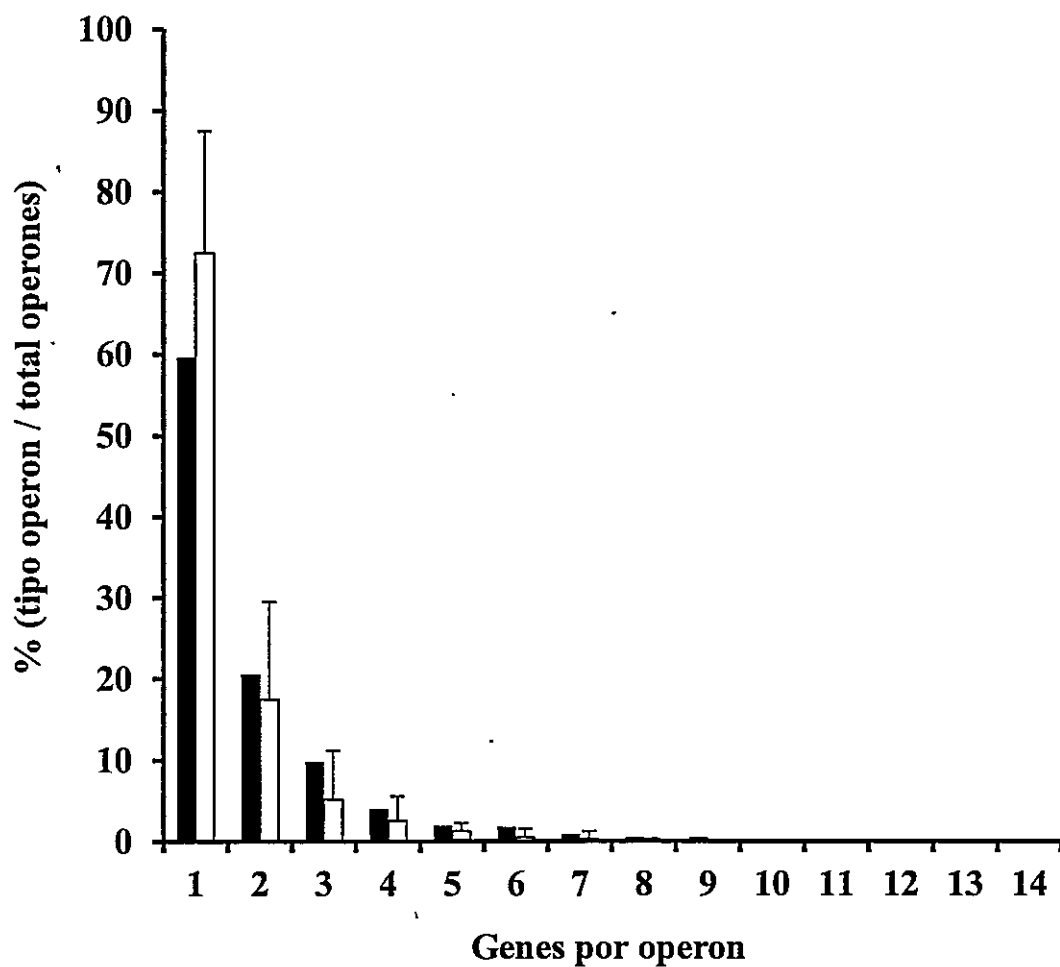


Figura 1. Distribución del tipo de operones en genomas bacterianos. El gráfico presenta el porcentajes de cada tipo de operones (columna 1 en tabla 2) con respecto al total del genoma para: *E. faecalis* (barras negras) y el promedio de 550 especies bacterianas (barras blancas, base de datos de operones ODB³⁰).

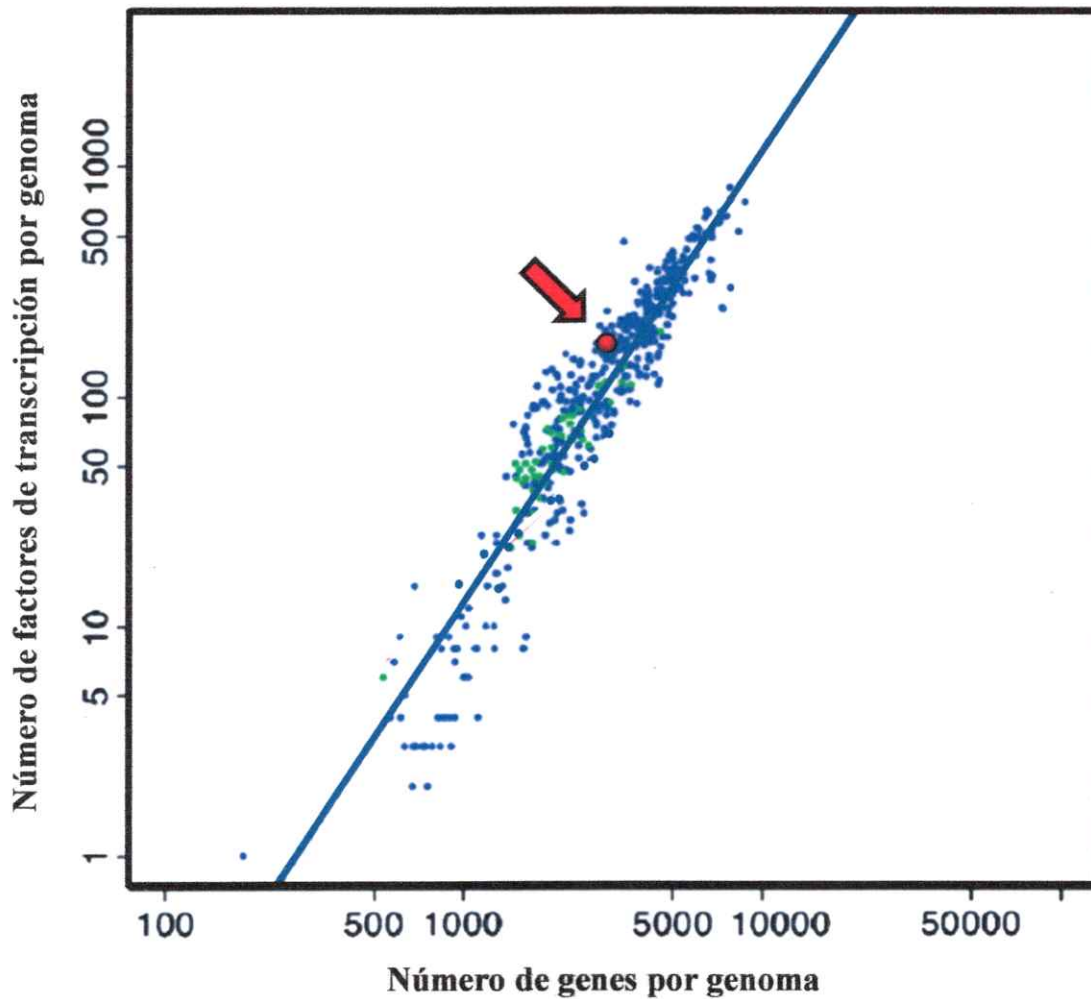


Figura 2. Proporción del total de factores de transcripción en relación al total de genes presentes en diversos genomas. Puntos azules y verdes corresponden a genomas bacterianos y arqueas respectivamente. Flecha roja señala la posición en el gráfico del valor calculado para *E. faecalis* (punto rojo). Ambos ejes graduados a escala logarítmica. Figura adaptada⁶⁸.

Tabla 3. Posibles factores de transcripción agrupados por familia presentes en el genoma de *E. faecalis*.

Familia	Número de factores de transcripción
AraC	10
ArgR	4
ArsR	2
Che-Y	14
Cro-CI	14
CRP-FNR	3
DeoR	5
Fur	3
GalR-S	12
GntR	14
LacI	11
LysR	9
MarR	6
MerR	3
Pho	10
PurR	13
TetR	5
Otros	45
Total	183

1.2 Construcción *in silico* de la red de regulación transcripcional mediada por factores de transcripción (escala global).

Luego de la selección de los nodos reguladores y blancos de regulación, el paso siguiente fue identificar los enlaces (sitios de unión) en las regiones promotoras de los operones de *E. faecalis*, capaces de establecer conectividades entre estos elementos para construir la red. En este contexto, actualmente *E. faecalis* no cuenta con una base de datos propia de los sitios de unión para el genoma. Por lo tanto fue necesario emplear una estrategia bioinformática capaz de buscar dentro del total de las posibles regiones promotoras, secuencias nucleotídicas consensos descritas en otras especies (MPPs), generadas por alineamientos de sitios de unión, potencialmente reconocidas por factores de transcripción codificados en nuestro modelo de estudio.

El número de MPPs utilizadas ascendió a un total de 149 (28 familias de factores de transcripción representadas). La búsqueda de estas matrices se realizó en las regiones promotoras comprendidas entre la posición +50 y -300 del inicio de la traducción del primer gen codificado en cada uno de los operones predichos en la bacteria. Esta región fue seleccionada según consensos reportados que describen entre estas posiciones al tamaño máximo donde se han encontrado sitios de unión en bacterias^{70, 71}.

Definidos estos parámetros, se determinaron los puntos de corte para la aplicación de los programas de búsqueda. Cada uno de los sitios de unión encontrados presenta un puntaje calculado en relación al nivel de homología con su MPP correspondiente utilizada en la búsqueda. El puntaje mínimo aceptable se consideró en función de la razón entre el total de sitios de unión encontrados y el total de operones contenidos en la red. A modo de ejemplo, se ha descrito en diferentes modelos de redes como en *E. coli* y *B. subtilis*, que el total de sitios de

unión no supera al doble de los blancos de regulación (operones) presentes en el genoma⁷². Utilizando estos criterios, se logró construir una posible red de regulación transcripcional mediadas por factores de transcripción en *E. faecalis* que presenta un total 716 enlaces o posibles sitios de unión y 597 nodos divididos en: i) blancos de regulación: 542 operones (1007 genes totales, cercano al 33% del genoma) y ii) elementos reguladores: 13 familias de factores de transcripción (55 factores independientes, 30% del total).

La figura 3 describe un grafo de conectividades, en la cual se muestran los elementos reguladores (familias de factores de transcripción) conectados con los blancos de regulación (operones) a través de la presencia de potenciales sitios de unión (enlaces). En particular, la base de datos de MPPs utilizada en la búsqueda contiene un alto grado de precisión para algunos componentes particulares de ciertas familias de factores de transcripción, lo que permitió asignar específicamente sitios de unión para los factores Fur, Per y Zur de la familia Fur, PhoP de la familia Pho, FruR de la familia DeoR y MalI de la familia LacI.

Dentro de la red, tanto la distribución de posibles funciones asignadas como la proporción de operones mono y policistrónicos se mantiene con respecto al total del genoma de *E. faecalis* (tabla 4), destacándose blancos de regulación con funciones vinculadas con homeostasis de metales y estrés oxidativo regulados por los factores locales CopY, Fur, Zur, Per.

En relación a la posición de los sitios de unión dentro de la región promotora y la secuencia consenso generada a través de alineamiento de cada uno de las secuencias agrupadas por familia de factor de transcripción (Logo), ambos poseen un alto grado de correspondencia con datos reportados (figura 4 y 5)^{33, 73-80}, resultado que en conjunto con la distribución de funciones ratifican el hecho de que en la operación de los programas no hubo sesgos biológicos

para la construcción de la red y que los elementos encontrados guardan relación con lo descrito en otras bacterias.

En resumen, la ausencia de limitaciones en cuanto al total de operones del genoma, así como la ausencia de criterios que pudiesen haber limitado o sesgado los posibles blancos de regulación, permitieron describir a la red de regulación transcripcional de *E. faecalis* con características globales, donde la única limitante a nivel de la búsqueda se encuentra en la información disponible sobre el número y tipo de matrices utilizadas como elementos de entrada.

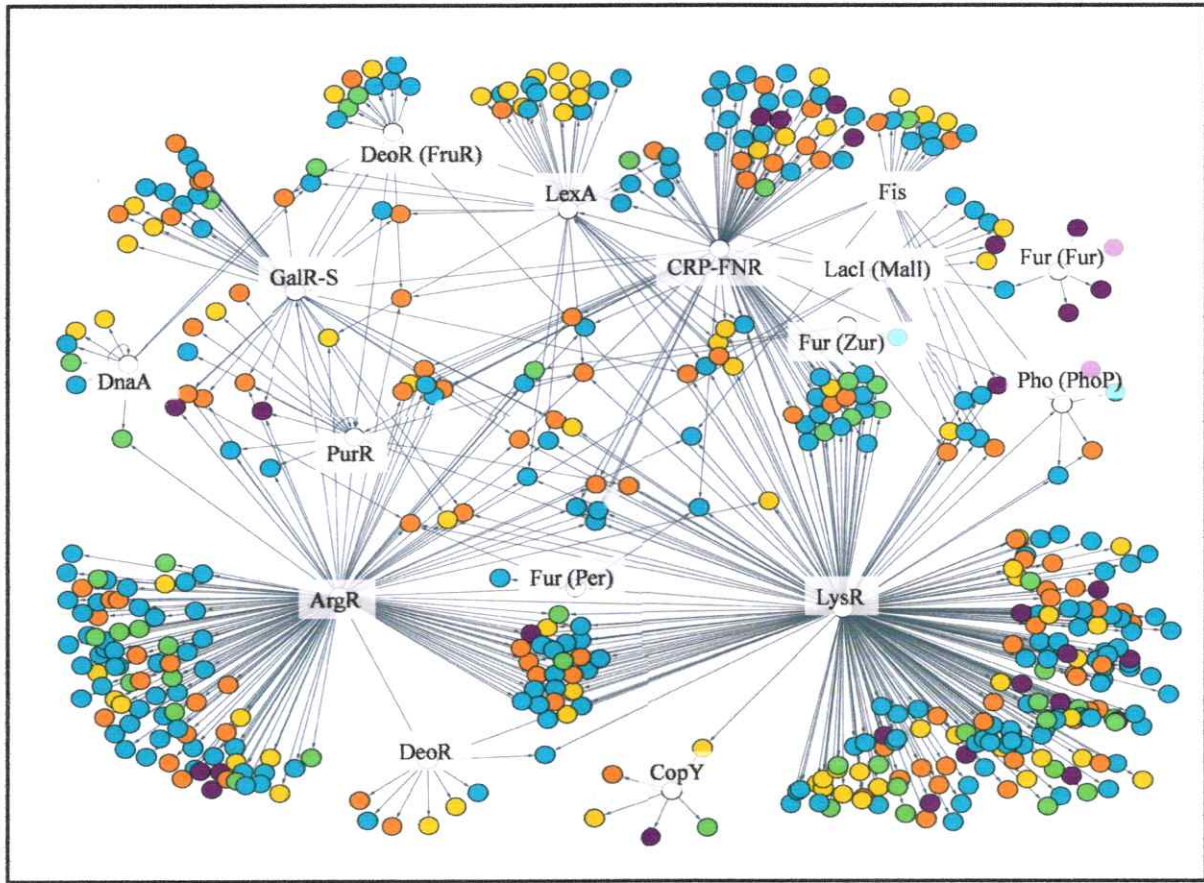


Figura 3. Red de regulación transcripcional a escala global de *E. faecalis*. El grafo muestra la visualización del archivo “E.faecalisV583.cys” a través del programa Cytoscape. Simbología de colores (agrupados por posibles funciones) para nodos (esferas): i) blancos de regulación (542), amarillo transcripción, traducción y replicación; morado transporte y señalización; naranja metabolismo basal y generación de energía; verde otras funciones; celeste sin función asignada o proteínas hipotéticas; ii) elementos reguladores en blanco (13 familias de factores de transcripción). Los enlaces (716 sitios de unión) se indican con flechas plumas.

Tabla 4. Porcentaje de categorías COG de los genes contenidos en la red global construida en *E. faecalis*. El porcentaje para cada categoría se calculó en relación al total de genes presentes en el genoma de la bacteria para cada una de las categorías.

Categorías COG	Genoma (%)		Red Global (%)	
	monocistrónicos	policistrónicos	monocistrónicos	policistrónicos
Transcripción, traducción y replicación	9,92	5,83	10,08	5,95
Transporte y señalización	4,25	2,52	3,13	3,13
Metabolismo basal y generación de energía	10,08	11,18	10,08	9,98
Otras funciones	5,51	4,41	4,84	5,14
Sin función asignada o proteínas hipotéticas	33,70	12,60	34,27	13,41

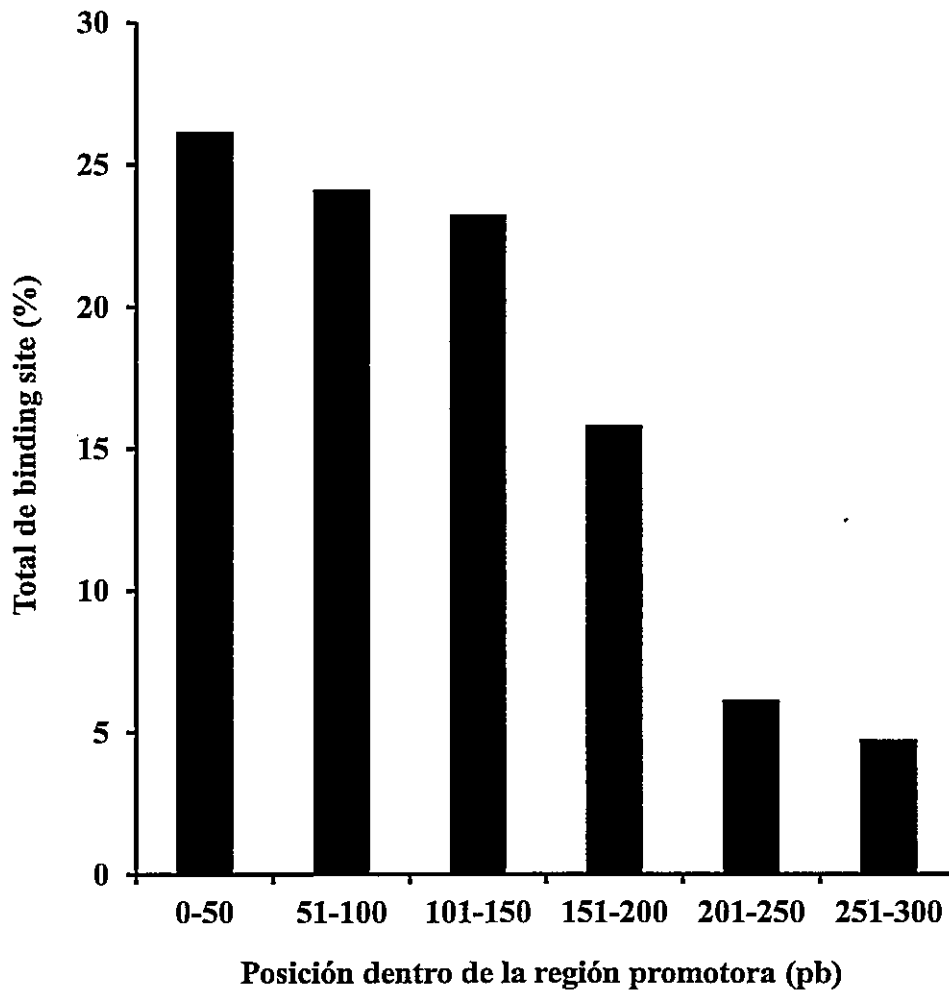


Figura 4. Distribución porcentual de la posición del total de sitios de unión encontrados dentro de su región promotora dentro de la red global. Los rangos de posiciones se consideran desde el inicio de la traducción. El porcentaje se calculó con la suma total de sitios de unión presente en cada rango de posición dividido por total de sitios de unión contenidos en la red (n = 716).

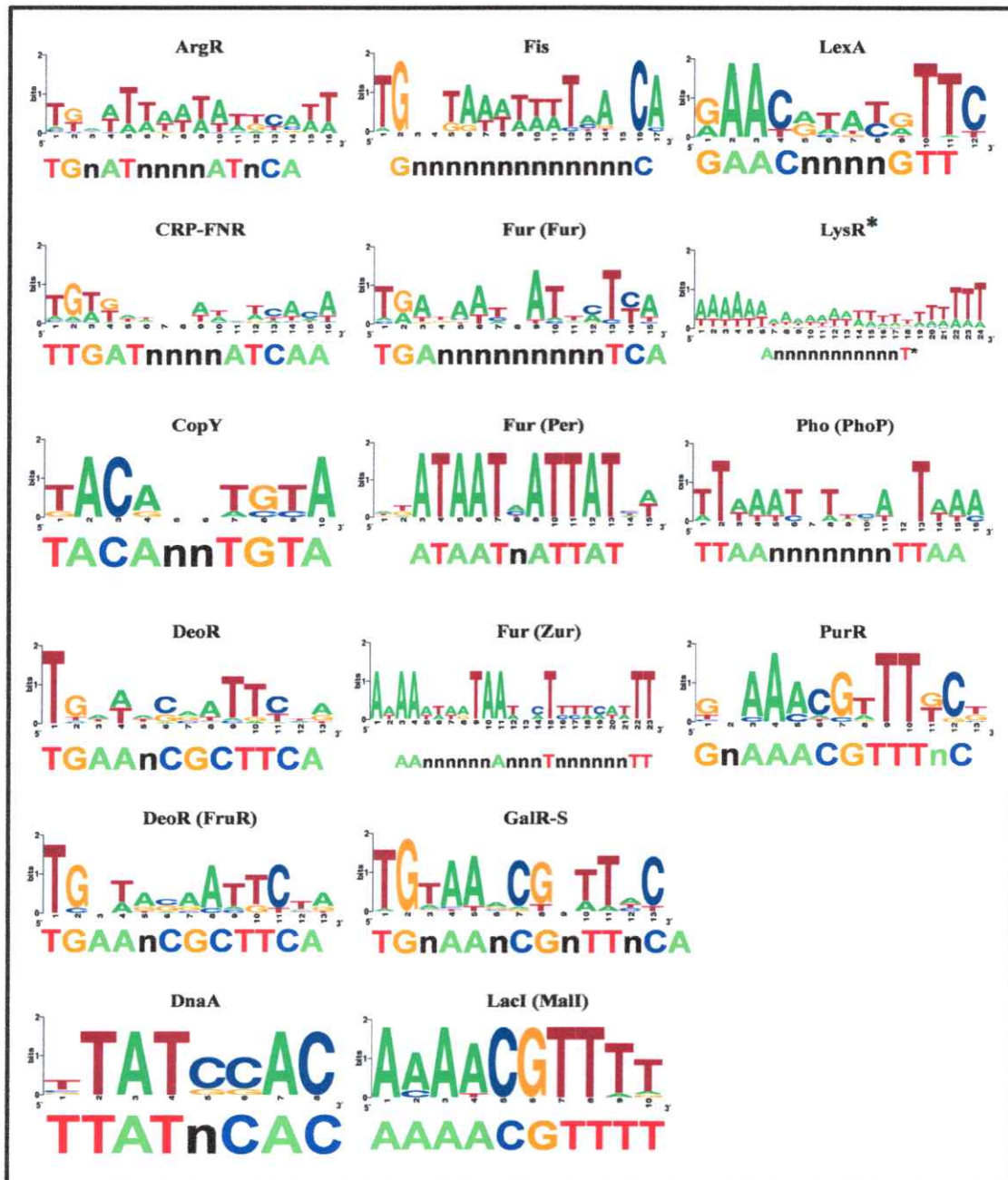


Figura 5. Logos consenso para los sitios de unión contenidos en la red global. Representación gráfica de MPPs construidas con los sitios de unión asignados a cada familia de factores de transcripción presentes en la red. La secuencia consenso reportada (indicada en la base de cada Logo) fue obtenido directamente de información disponible en la literatura. *Consenso de LysR puede ajustarse en más de una posición dentro del Logo.

1.3 Análisis topológicos y asignación de propiedades de la red transcripcional global de *E. faecalis*.

A través de la identificación y análisis de propiedades estructurales en cualquier tipo de red, se logra precisar importantes características sistémicas, las cuales se conservan en diversos modelos de interacciones biológicas descritos a la fecha⁸¹. Esta caracterización se alcanza principalmente a través de la determinación de parámetros topológicos, los cuales finalmente entregan información sobre propiedades cuantitativas y cualitativas que presenta una red en términos de su estructura y composición de elementos⁸².

Considerando los protocolos actuales utilizados en la determinación de parámetros topológicos de redes de regulación transcripcional, actualmente es posible diferenciar dos niveles de análisis y clasificación.

A escala global

A este nivel encontramos la distribución del grado o $P(K)$, la cual describe el número de conectividades establecidas entre los nodos que conforman la red⁸³. En una red de regulación transcripcional, las conectividades generadas son direccionadas, es decir, la relación se establece por la capacidad del factor de transcripción de reconocer su correspondiente sitio de unión en la región promotora de un gen. Esta restricción de direccionalidad tiene por consecuencia que la distribución del grado sea de entrada (hacia el gen) o de salida (desde el factor de transcripción).

Estudios topológicos realizados en diferentes redes de regulación transcripcional, señalan que esta distribución conserva una función del tipo $P(K) = AK^{-\gamma}$ (*power law*) o de "Escala Libre", lo que biológicamente se manifiesta dentro de una red en una alta probabilidad de encontrar un

nodo con un grado bajo de conectividades, probabilidad que decae en forma exponencial en la medida que aumenta el valor del grado, siendo el exponente γ en la ecuación el que describe el decaimiento exponencial que va a presentar la distribución⁸³.

En el contexto de nuestra red, esto se traduce como una baja probabilidad de encontrar un operon que presente un alto número de sitios de unión en su región promotora. La figura 6 muestra la distribución del grado de entrada obtenida en la red de *E. faecalis*, como se puede observar, nuestra red cumple con las propiedades de escala libre, cuya ecuación presenta un valor $\gamma = 3,57$, valor que se encuentra dentro del rango esperado⁸³.

Por otro lado, se observó el “grado de salida” de nuestra red, parámetro que describe la distribución del número de conectividades que surgen desde un factor de transcripción hacia sus operones blanco, esta propiedad puede ser utilizada para clasificar los factores en⁸⁴: i) locales (regulan pocos genes, < 5), genes relacionados en su mayoría funcionalmente, ii) globales, regulan un alto número de genes (> 20); dentro de este grupo, si los genes regulados codifican para proteínas vinculadas al mismo proceso o función metabólica dentro de la célula son designados como maestros.

En la red de *E. faecalis* (tabla 5), en general el número de sitios de unión encontrados concuerda con la clasificación descrita en la literatura para cada uno de los factores de transcripción, destacándose LysR, ArgR y CRP-FNR como los que contienen el mayor número de conectividades.

A escala local

El tipo de conectividades describe de forma particular, elementos conservados que deberían estar presentes al momento de analizar cualquier red de regulación transcripcional. Básicamente, son patrones conservados descritos en todas las redes transcripcionales construidas a la fecha que le adjudican propiedades locales al sistema⁸⁵. La tabla 6 muestra los resultados obtenidos para la red de *E. faecalis* en estudio.

En relación a los elementos identificados en *E. faecalis*, trece patrones fueron clasificados como regulones, los cuales definen en nuestra red transcripcional sub-redes integradas por un factor de transcripción con sus respectivos genes blancos. Además, se identificaron cuatro orígenes, que constituyen variantes de regulones, definidos por su capacidad de detectar directamente el estímulo, y por lo tanto no dependen de otro regulador para traducir la señal. Regulones y orígenes describen componentes que en definitiva permiten entender cómo la bacteria logra controlar etapas de adaptación frente a diferentes estímulos del medio⁸¹.

Por otra parte, el patrón más simple identificado en nuestra red es la autorregulación (positiva o negativa, seis casos), definido como un factor de transcripción capaz de regular su propia expresión, lo que permite aumentar o disminuir la velocidad en el control de la transcripción sobre él y sus genes blanco. Un número similar de casos fue observado para patrones de regulación en cadena, en los cuales un regulador se une al promotor de un segundo regulador y este a su vez, reconoce el promotor de un tercero, y así sucesivamente. Esta cadena representa el circuito lógico más simple para ordenar de forma temporal eventos transcripcionales. Los patrones más representados son los denominados "loops" de alimentación, descritos como operones regulados transcripcionalmente por dos o más factores, los cuales a su vez se regulan entre ellos. Esta configuración le entrega al sistema la propiedad de potenciar o

disminuir el efecto del regulador sobre la transcripción en función de los estímulos a lo que se enfrenta la bacteria (control fino sobre la activación o represión del gen). Finalmente se identificaron módulos de entradas singulares, definidos como un factor de transcripción capaz de regular un gran número de genes, lo que permite ampliar y diversificar la traducción de la señal en el corto plazo.

La identificación de estos patrones dentro de la red de *E. faecalis*, permite por un lado entender desde un punto transcripcional, cómo la bacteria logra controlar temporalmente diferentes etapas de adaptación frente a diferentes estímulos del medio. Además posibilita la clasificación de elementos de la red en sub-redes que presentan propiedades particulares basadas en su configuración de conectividades.

Los resultados obtenidos en la caracterización topológica de la red de regulación transcripcional de *E. faecalis*, sugieren un modelo útil para trabajar. No obstante, dado la ausencia de un estímulo que va a activar o reprimir parte de la red construida, se imposibilita a este nivel dar una explicación del funcionamiento de la red. Por lo tanto, la identificación y caracterización de sub-redes se hará en la sección siguiente, cuando se incluya la información derivada de los experimentos de expresión génica obtenidos en el modelo bacteriano expuesto a un estímulo de Cu.

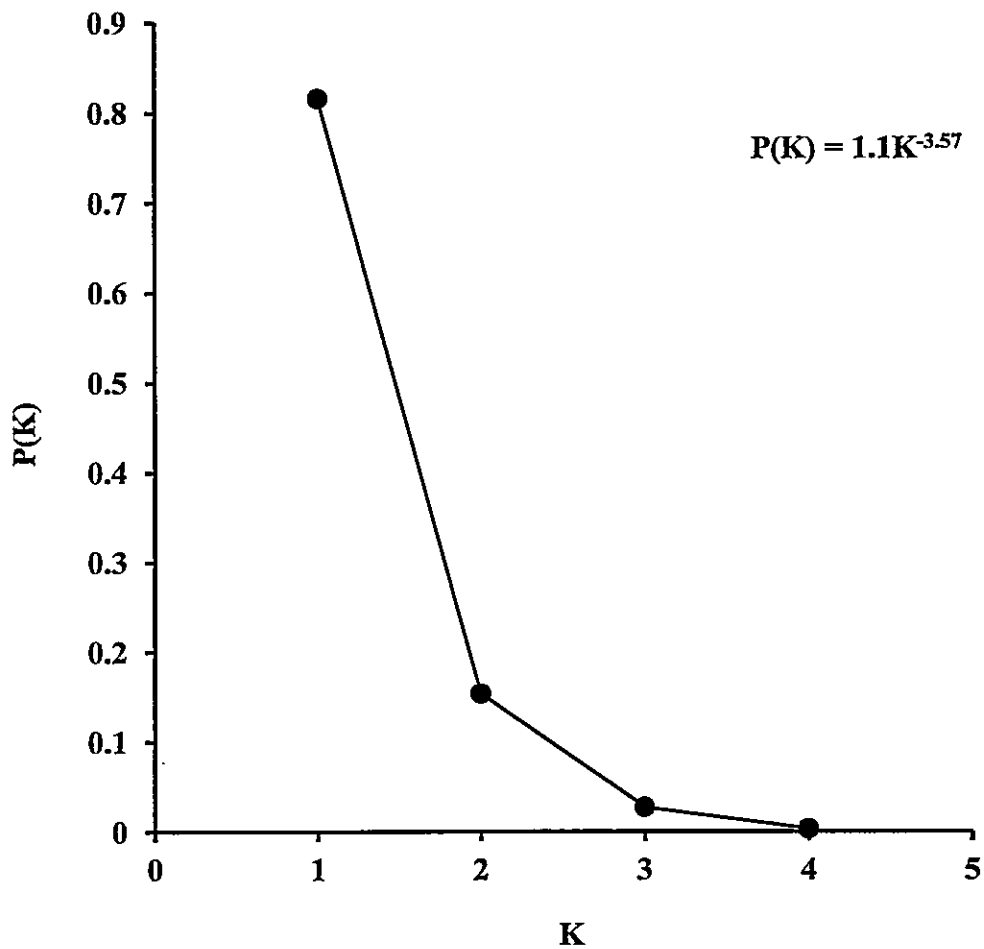


Figura 6. Distribución del grado de entrada para los genes presentes en la red de regulación transcripcional de *E. faecalis*. El gráfico muestra la probabilidad $P(K)$ para un gen en particular de tener $K(1,2,\dots,n)$ interacciones dentro de la red global construida.

Tabla 5. Número de sitios de unión para cada familia de factores de transcripción y clasificación según tipo de factor de acuerdo con datos reportados en otras bacterias. Según literatura⁸⁶, factores clasificados como globales/maestros reconocen dentro del genoma sobre 20 sitios de unión, para el caso de los locales, el número de conectividades es inferior a 5.

Familia	Sitios de unión	Tipo
ArgR	142	Global*
CopY	5	Local*
CRP-FNR	74	Global*
DeoR	7	Local
DeoR (FruR)	12	Maestro
DnaA	9	Local
Fis	17	Local
Fur (Fur)	5	Local
Fur (Per)	3	Local*
Fur (Zur)	6	Local*
GalR-S	30	Global*
LacI (Mall)	12	Local
LexA	33	Maestro*
LysR	237	Global*
Pho (PhoP)	5	Local
PurR	10	Local

* Familia de factores de transcripción descritos en otros organismos bajo la misma clasificación de Global, Maestro o Local

Tabla 6. Patrones presentes en la red global de regulación transcripcional de *E. faecalis* agrupados por el tipo de conectividad.

Tipo de conectividad	Cantidad de patrones
Regulones (familia)	13
Orígenes	4
Autorregulación	6
Regulación en cadena	8
Loops de alimentación	56
Módulo de entrada singular	10

1.4 Identificación de componentes de la red de regulación transcripcional activada por Cu.

La construcción de la red de *E. faecalis* activada por el metal, implicó identificar dentro de la red global construida un conjunto particular de genes capaces de aumentar su abundancia transcripcional frente a la exposición a Cu, lo que constituyó finalmente una sub-red dentro del universo de conectividades en estudio. En el año 2010, publicamos la respuesta transcripcional global de *E. faecalis* OG1RF frente a una exposición elevada no letal de Cu (0,5 mM)³⁶. Utilizando esta información como referente, se rastreó dentro de la red de regulación transcripcional de *E. faecalis*, un total de 84 operones (109 genes) que aumentaron significativamente su abundancia transcripcional en el microarreglo.

Del cruce de los datos (figura 7), se obtuvo una sub-red activa compuesta por 33 operones (87 genes) conectados entre si por 56 sitios de unión para 9 familias de factores de transcripción (34 reguladores individuales), conjunto de elementos que constituyen la sub-red de regulación transcripcional activada en *E. faecalis* por el metal.

El análisis de la estructura de esta sub-red mantiene, al igual que la red global, características topológicas de red a escala libre ($\kappa = 2,05$). En cuanto a nivel de las características particulares de los operones identificados, los datos indican que la traducción directa de la señal activada por el Cu sobre la regulación transcripcional estaría afectando solo a genes vinculados con la homeostasis de Cu (CopY). No obstante, además de este origen encontramos al factor de transcripción Fur, que podría responder a cambios en el estado redox producto de la exposición sostenida a Cu, regulando a su vez la expresión de genes vinculados a la incorporación de Fe y al gen que codifica para el regulador DnaA, que a su vez podría

controlar la expresión del gen EF3027 que codifica posiblemente para una serina proteasa vinculada a la protección frente a estrés oxidativo⁸⁷. En general, el resto de los factores eventualmente podrían participar en procesos que detectan cambios metabólicos producto de la acumulación del metal, vinculados principalmente al posible estrés oxidativo generado. Tal es el caso de LysR y LexA, cuya función es regular genes de estrés en general y reparación de DNA, respectivamente^{80, 88}. En relación con las conectividades, los regulones CRP-FNR, LysR y ArgR en su calidad de factores globales, podrían ampliar la señal transcripcional a un gran número de genes, destacándose el operon EF0955-0957 (síntesis de trehalosa) el cual presenta el mayor grado de regulación (mayor número de sitios de unión encontrados en su promotor). Interesantemente la acumulación de este azúcar se ha vinculado directamente con la protección frente a ROS⁸⁹.

Nuestro siguiente paso fue verificar experimentalmente si las conectividades predichas en la red entre los factores de transcripción y sus respectivos operones blanco se cumplen en nuestro modelo de estudio. La estrategia experimental incluyó la generación de cepas mutantes de la bacteria para los factores de transcripción de interés, con el propósito de evaluar si la expresión de los genes activados por Cu en la red depende de la presencia de los elementos reguladores propuestos. En este contexto, dado el alto número de conectividades a estudiar y el tiempo establecido para la finalización de la tesis de doctorado, se hizo necesario seleccionar dentro de la sub-red activada por el metal, un número abordable de factores de transcripción para ser mutados en la bacteria. Para esto, se aplicaron dos filtros de selección diseñados específicamente en consideración a las limitaciones experimentales y no a criterios biológicos, decisión que nos permitió evitar sesgos en la interpretación y posterior discusión de los resultados. A continuación se describen los filtros de selección utilizados:

a) Conectividades: Como podemos observar en la figura 7, los regulones LysR, ArgR, CRP-FNR, Fur y LexA presentan el mayor número de conectividades (5 o más blancos de regulación). Tal como se mencionó en la sección de Materiales y Métodos, considerando la posibilidad que los sitios de unión predichos en el análisis *in silico* fuesen falsos positivos, sólo se eligieron estos regulones para optimizar la probabilidad de encontrar conexiones reales durante la curación *in vitro/in vivo* de la red.

b) Redundancia en el genoma: La estrategia experimental posterior incluyó la generación de mutantes para los factores de transcripción en cada uno de los regulones. En el caso de CRP-FNR, ArgR y LysR existen 3, 4 y 9 posibles genes respectivamente, que codifican para las proteínas de estas familias. La redundancia a este nivel es un problema dado que se pretende verificar cambios de expresión entre la cepa mutante y la silvestre para el factor de transcripción que comanda el regulon, resultado que podría ser apantallado por la presencia de un homólogo con funciones similares al regulador seleccionado. Tomando esto en consideración, se excluyó estos tres regulones de los posteriores análisis.

Estudios previos realizados en nuestro laboratorio, caracterizaron al factor CopY como un elemento regulador de operon *cop* de *E. faecalis*³⁶, el cual codifica para proteínas involucradas en la homeostasis del Cu. Dentro de la red activada por Cu, encontramos además del operon *cop* como parte del regulon CopY, otros 2 operones que presentan un sitio de unión para este represor. Si bien, CopY no satisface el primer criterio propuesto, dado su vínculo directo con el metal, fue considerado en los experimentos posteriores. Con esto, el conjunto final de regulones a estudiar directamente en la bacteria está compuesto por los factores LexA (5 operones), Fur (5 operones) y CopY (3 operones).

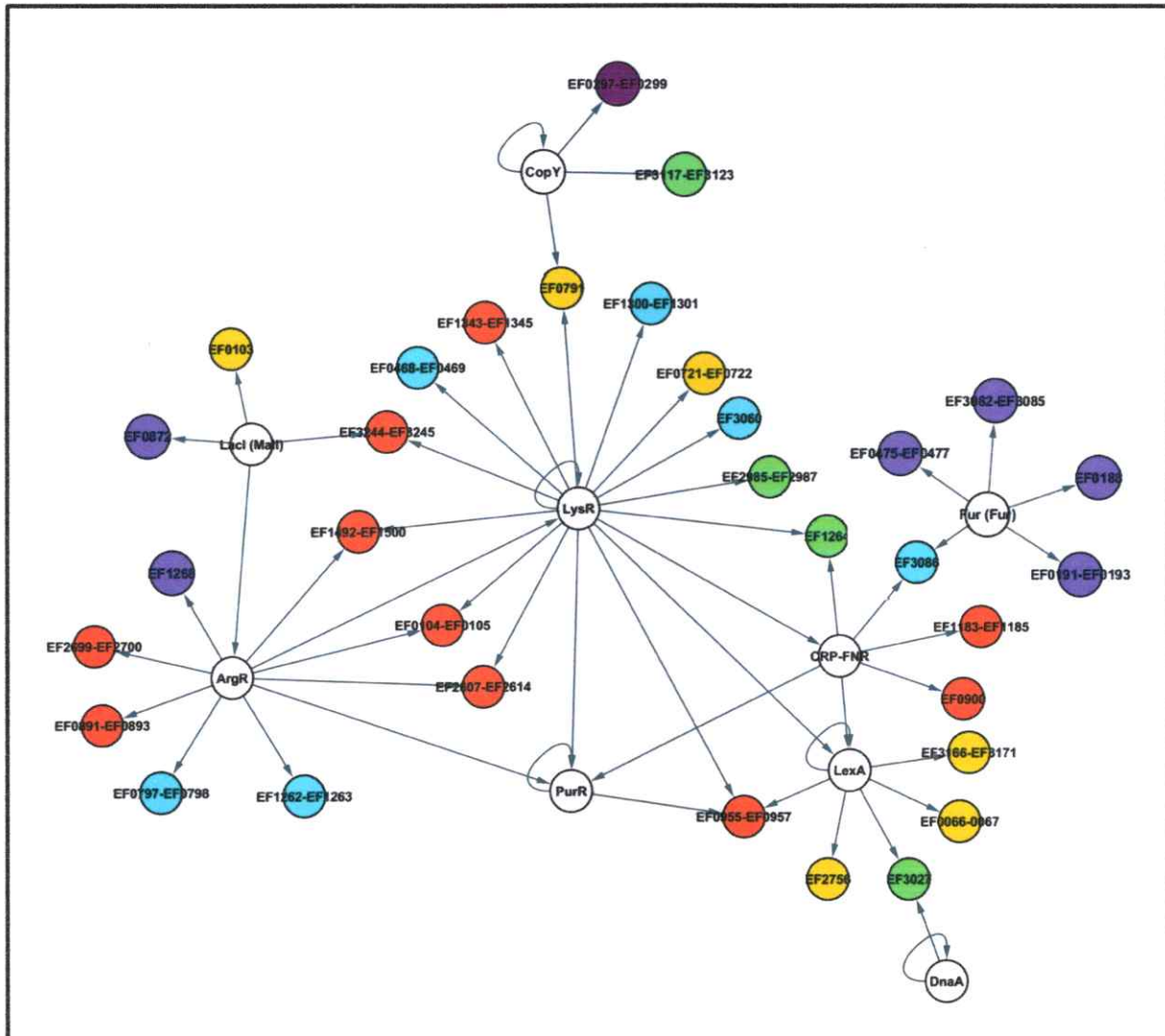


Figura 7. Sub-red de regulación transcripcional activada por Cu en *E. faecalis*. El grafo contiene 42 nodos (9 reguladores o familias de factores de transcripción y 33 blancos de regulación u operones) conectados a través de 56 enlaces (sitios de unión). Simbología de colores (agrupados por posibles funciones) para nodos (esferas): i) blancos de regulación (542), amarillo transcripción, traducción y replicación; morado transporte y señalización; naranja metabolismo basal y generación de energía; verde otras funciones; celeste sin función asignada o proteínas hipotéticas; ii) elementos reguladores en blanco (13 familias de factores de transcripción). Los enlaces se indican con flechas plomas.

2. Determinación de la participación del(los) factor(es) de transcripción en la regulación de la expresión de los componentes de la red inducidos por Cu.

2.1 Verificación de los resultados de microarreglos a través de ensayos de PCR en tiempo real (qPCR).

Previo a la verificación en la bacteria de las conectividades predichas *in silico* para los 3 regulones seleccionados, es necesario dado el nivel de precisión de los microarreglos, mediante ensayos de qPCR se ratificó el aumento en la abundancia transcripcional observada en cada uno de los blancos de regulación activados (13 operones). Para ello, el primer paso fue validar el gen normalizar EF1004 en las muestras de RNA obtenidas luego de los tratamientos con Cu. El promedio de 3 mediciones realizadas independiente en cada una de las muestras, no mostró diferencias significativas en la cantidad de transcrito para el gen EF1004 entre el ciclo en el cual se empieza a detectar el aumento de fluorescencia (punto de cruce o *crossing point*, Cp, valores: control $15,19 \pm 0,25$, tratamiento $15,28 \pm 0,22$), ratificando que el gen EF1004 no fue afectado por la exposición a Cu y por lo tanto sirve como normalizador.

Confirmado el normalizador, se procedió a cuantificar los transcritos para los 13 operones seleccionados en el punto anterior. En el caso de unidades transcripcionales que posiblemente contienen información de 2 o más genes (policistrones), se diseñaron dos parejas de partidores capaces de amplificar el primer y último gen del operon. La tabla 7 presenta el cociente entre las muestras tratadas con Cu y control relativas al normalizador. Como se puede apreciar los valores obtenidos en todos los casos ratificaron que los regulones seleccionados se activan en presencia de Cu bajo las condiciones de exposición utilizadas.

Tabla 7. Verificación por qPCR de la activación por Cu de los regulones CopY, Fur y LexA. Razón de cambio normalizada (Ef^{ACp}) calculada a través de la ecuación 1 (ver sección 2.3.4. de MyM), utilizando abundancia calculadas en la muestra control y tratada con 0,5 mM de $CuSO_4$. Para policistrones el valor corresponde al promedio del primer y último gen del operon medidos en tres réplicas biológicas con tres réplicas técnicas (todos los valores mostraron diferencias significativas, REST $p < 0,05$).

Regulon	Anotación del operon	Ef^{Cp}
CopY		
EF0297-EF0299	Operon <i>cop</i>	$56,3 \pm 11,5$
EF0791	Factor de transcripción de la familia TetR	$8,5 \pm 2,8$
EF3117-EF3123	Pirofosfoquinasa	$4,8 \pm 1,1$
Fur		
EF0188	Transportador ABC de Fe	$8,4 \pm 1,4$
EF0191-EF0193	Transportador ABC de ferrocromos	$9,5 \pm 1,8$
EF0475-EF0477	Transportador de Fe	$4,1 \pm 1,6$
EF3082-EF3085	Transportador ABC de Fe	$3,6 \pm 0,8$
EF3086	Proteína hipotética	$6,8 \pm 1,4$
LexA		
EF0066-EF0067	DNA helicasa	$10,4 \pm 1,1$
EF0955-EF0957	Glucomutasa hidrolasa	$5,8 \pm 1,3$
EF2756	Proteína inducible por daño al DNA	$4,5 \pm 1,5$
EF3027	Serina proteasa	$4,1 \pm 0,7$
EF3166-EF3172	Proteína de reparación de DNA	$4,4 \pm 1,2$

2.2 Construcción y verificación de cepas mutantes de *E. faecalis* OG1RF para los factores de transcripción CopY, Fur y LexA.

Los resultados obtenidos en el objetivo 1 permitieron establecer el set final de elementos a estudiar que constituyen una parte de la red transcripcional inducida en respuesta a Cu, sub-red cuya validación fue abordada a través de experimentos *in vivo*. La estrategia utilizada para cumplir este objetivo involucró encontrar cambios transcripcionales en los blancos de regulación de cada regulon, en ausencia de su correspondiente regulador transcripcional, como podría ocurrir en cepas mutantes de *E. faecalis* para cada uno de los factores de transcripción en estudio.

Nuestro laboratorio cuenta con la cepa mutante de *E. faecalis* OG1RF para el gen *copY*, construida previamente en colaboración con el laboratorio de la Dra. Barbara Murray en la Universidad de Texas³⁶. La mutante para el gen *lexA* generada por la inserción del transposón Tn917³⁷, fue donada directamente por el laboratorio de la Dra. Danielle Garsin durante la estadía realizadas en Estados Unidos y verificada por PCR utilizando partidores que alinean en las regiones flanqueantes al gen (figura 8A). Un aspecto crucial dentro de la verificación de esta mutante, fue identificar exactamente el lugar de la inserción del transposón para suponer una interrupción efectiva de traducción de la proteína en su estado nativo funcional. En el modelo de *B. subtilis* y *E. coli*, la proteína LexA es su estado activo reprime la expresión de sus genes blanco, frente a daños en el DNA, la proteína RecA (EF3171) activa la autoproteólisis de LexA lo que permite su liberación desreprimiendo la expresión de genes involucrados en reparación⁸⁸. La figura 8B presenta un alineamiento global de la secuencia aminoacídica de LexA de *E. faecalis* y sus homólogos por identidad en *B. subtilis* y *E. coli*. La figura detalla específicamente el lugar de inserción del transposón (dato obtenido por secuenciación del fragmento amplificado

en la cepa mutante) y los residuos clivados en la autoproteólisis. Como se puede observar, la inserción interrumpe gran parte del gen (> 45 %) incluyendo dos puntos clave de autoclivaje, lo que supone la expresión de una proteína trunca o inactivada. Basados en esta información se decidió utilizar esta cepa mutante de *E. faecalis* en los posteriores ensayos de expresión.

En el caso particular del factor de transcripción Fur, la no disponibilidad de la mutante en *E. faecalis* para este gen, condujo a la utilización de una estrategia de delección por doble recombinación homóloga, estandarizada en el laboratorio de la Dra. Bárbara Murray. El trabajo realizado permitió escindir específicamente el gen *fur* del cromosoma de la bacteria sin dejar el remanente del uso de antibióticos o generación de mutaciones polares las cuales pueden afectar la expresión de otros genes. La figura 9 expone los resultados obtenidos de la amplificación por PCR convencional de la cepa silvestre, mutante *copY* (construida mediante la misma estrategia de mutación) utilizando partidores capaces de amplificar el gen mutado con sus respectivas regiones flanqueantes. En la migración de los productos generados, se logró distinguir claramente una disminución del tamaño del producto amplificado en la cepa mutante correspondiente al tamaño del gen delecionado. Ambos amplicones se purificaron y secuenciaron lo que permitió finalmente verificar la correcta delección por doble recombinación.



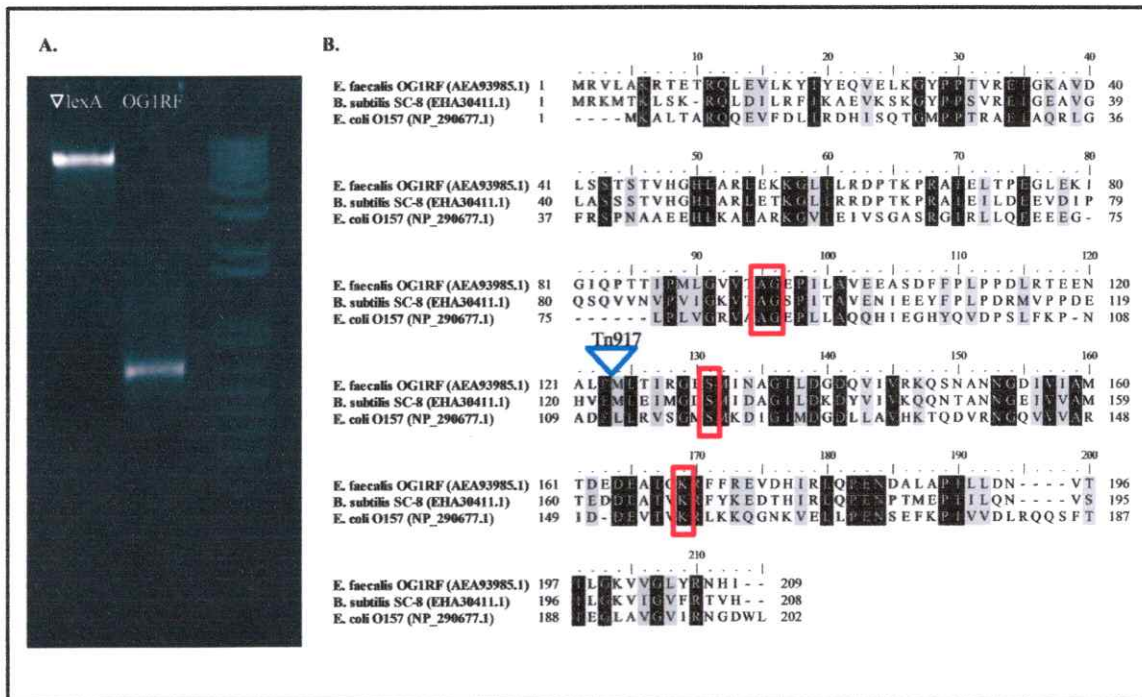


Figura 8. Verificación de la cepa mutante *lexA* de *E. faecalis*. A. Migración electroforética en geles de agarosa de los productos *lexA* amplificados por PCR convencional en la cepa mutante ∇ *lexA* y silvestre (OG1RF). Estándar de tamaño 1 kb DNA plus *ladder* (Invitrogen, primer carril de derecha a izquierda). B. Alineamiento global de las secuencias aminoacídicas de los homólogos por identidad de LexA de *E. faecalis*, *B. subtilis* y *E. coli*. Triángulo azul indica el lugar específico de la inserción del trasposón Tn917, cuadros rojos indican los residuos clivados en la autólisis.

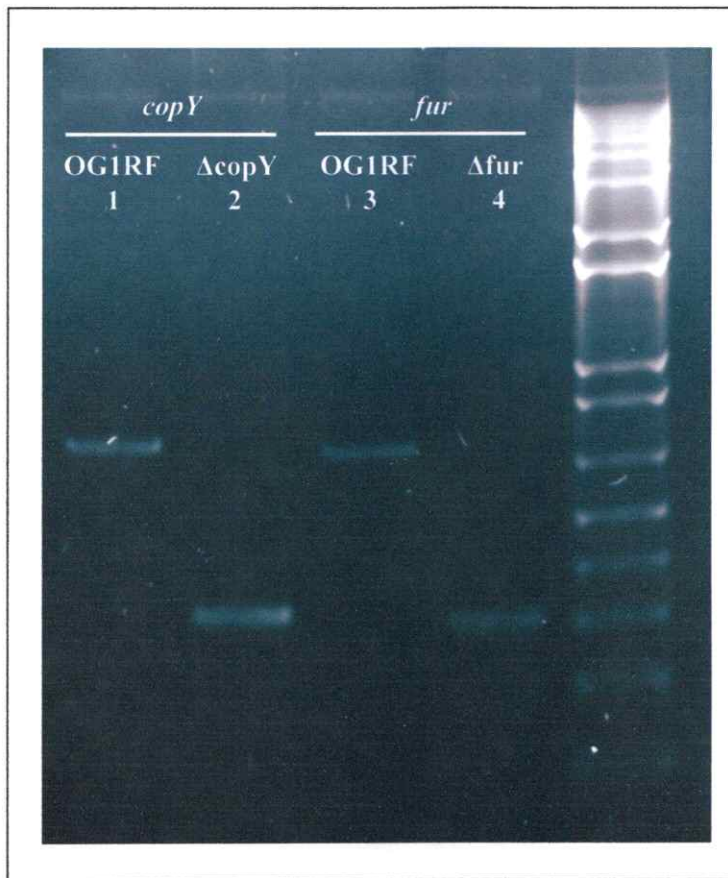


Figura 9. Verificación de las cepas mutante *copY* y *fur* de *E. faecalis*. El gel muestra la migración realizada para productos de PCR amplificados de los genes: carril 1 *copY* cepa silvestre OG1RF, carril 2 *copY* cepa mutante *copY* (Δ *copY*), carril 3 *fur* cepa silvestre OG1RF, carril 4 *fur* cepa mutante *fur* (Δ *fur*). En el primer carril de derecha a izquierda, se cargó el estándar de tamaño 1 kb DNA plus *ladder* (Invitrogen).

2.3 Análisis de cambios transcripcionales mediante qPCR en las cepas mutantes

Construidas y verificadas las mutantes para los tres regulones en estudio, tal como se mencionó anteriormente, este objetivo buscó determinar la existencia de cambios transcripcionales en los blancos para los regulones CopY, Fur y LexA utilizando mutantes para sus factores de transcripción. Para ello, cultivos independientes de las cepas silvestre y mutantes fueron crecidos en paralelo ($DO_{600nm} = 1$, aproximadamente 3 horas de cultivo en medio basal N, réplica temporal del control utilizado en los ensayos de microarreglo y qPCR de exposición a Cu) se les extrajo RNA y sintetizó el cDNA correspondiente. Al igual que lo observado en el punto 2.1 para el normalizador EF1004 no se obtuvo diferencias significativas en sus valores de abundancia transcripcional para todas las muestras en estudio (Cp valores: $15,54 \pm 0,32$ silvestre, $15,34 \pm 0,29$ mutante *copY*, $15,14 \pm 0,23$ mutante *fur*, $15,44 \pm 0,30$ mutante *lexA*).

Tomando en consideración que la adición de Cu activa los operones que constituyen los 3 regulones, así como que en otros organismos los factores de transcripción LexA, Fur y CopY actúan como represores transcripcionales, el esperado en la cuantificación en la mutante (sin exposición a Cu) es un aumento significativo en la abundancia transcripcional para cada uno de los transcritos que codifican para sus correspondientes blancos de regulación (como sucede cuando la cepa silvestre es expuesta a Cu). En el caso particular de la mutante para el gen *lexA*, dado su construcción producto de la inserción del transposón Tn917, aunque esta proteína está codificada dentro de un operon monocistrónico, previo a la cuantificación de sus blancos de regulación, a modo de control se midió la abundancia transcripcional de los dos genes próximos codificados río arriba y abajo de *lexA*, en ambos casos no hubo diferencias significativas de expresión entre la cepa silvestre y mutante, indicando que no se generaron efectos en la

expresión de estos genes. La tabla 8 da cuenta de los valores de qPCR obtenidos para los blancos de regulación de los 3 regulones propuesto en la red activados por el Cu con sus correspondientes sitios de unión, lo que supone que estos regulones solo controlarían la expresión de genes vinculados a componentes de homeostasis de metales y reparación frente a estrés oxidativo se activaron en la respectiva mutante.

Estos resultados finalmente permitieron curar parte de la red transcripcional generada *in silico*, cuyo alcance de los resultados obtenidos para la red activada, solo nos permite interpretar con mayor profundidad y seguridad las conectividades dentro de los regulones LexA, Fur y CopY. Tomando esto en consideración, la curación manual además permitió generar una nueva sub-red activa por Cu en *E. faecalis* (figura 10), en el cual podemos notar la aparición de dos nuevos módulos: i) uno de tipo central o general de regulación coordinado principalmente por los factores globales LysR, ArgR y CRP-FNR, cuyas función posiblemente son las de regular genes que codifican para proteínas de metabolismo basal y generación de energía, ii) otro de tipo independiente, compuesto por los regulones esencialmente por CopY y Fur, los cuales presentan un alto grado de especificidad en el reconocimiento del sitio de unión (alto grado de secuencias palindrómicas), vinculados directamente a regular genes que codifican para componentes con homeostasis de Cu y Fe.

El objetivo final de la tesis pretende estudiar si la ausencia en el control transcripcional de genes activados en presencia de Cu impacta sobre la resistencia celular al metal. Para esto, a través de ensayos microbiológicos de viabilidad y contenido del metal se comparó en todos los casos, el comportamiento entre las cepas mutantes *lexA*, *copY* y *fur* con la cepa silvestre OG1RF frente a la toxicidad mediada por Cu, estrategia ya reportada por nuestro laboratorio utilizada en la caracterización del modelo de homeostasis de Cu de *E. faecalis*.

Tabla 8. Verificación por qPCR de las conectividades predichas *in silico* mediante el uso de mutantes para los regulones CopY, Fur y LexA. La razón de cambio normalizado (Ef^{ACp}) se calculó a través de la ecuación 1 (ver sección 2.3.4. de MyM), utilizando los valores de abundancia para la muestra silvestre y mutante correspondiente para cada regulon. En el caso de policistrones el valor presentado es el promedio del primer y último gen del operon medidos en tres réplicas biológicas con tres réplicas técnicas. Asterisco demarca diferencias significativas (REST $p < 0,05$). Secuencias palindrómicas dentro del sitio de unión se subrayaron en negro.

Operon	Anotación del operon	Razón de cambio normalizado	Sitio de unión
CopY			
EF0297-EF0299	Operon <i>cop</i>	36,6 ± 7,5*	<u>TACATTTGTA</u>
EF0791	Factor de transcripción de la familia TetR	1,1 ± 0,4	<u>TACATTTGAA</u>
EF3117-EF3123	Pirofosfoquinasa	1,0 ± 0,2	<u>AACAGTTGTA</u>
Fur			
EF0188	Transportador ABC de Fe	8,1 ± 1,6*	<u>TGAGAACGATTCTCA</u>
EF0191-EF0193	Transportador ABC de ferrocromos	5,5 ± 1,1*	<u>TGATAACGATTCTCA</u>
EF0475-EF0477	Transportador de Fe	19,2 ± 2,8*	<u>TGAGAACGATTCTCA</u>
EF3082-EF3085	Transportador ABC de Fe	125,1 ± 14,8*	<u>TGAGAATGATTCTCA</u>
EF3086	Proteína hipotética	1,9 ± 0,5	<u>TGAGAATGATTCTCA</u>
LexA			
EF3166-EF3172	Proteína de reparación de DNA	2,9 ± 0,7*	<u>GAACAAACATTC</u>
EF0066-EF0067	DNA helicasa	2,6 ± 0,3*	<u>GAACATATGTAC</u>
EF2756	Proteína inducible por daño al DNA	2,4 ± 0,4*	<u>GAACGTCCGTTC</u>
EF0955-EF0957	Glucomutasa hidrolasa	0,9 ± 0,5	<u>GAATAAACGTTT</u>
EF3027	Serina proteasa	0,8 ± 0,3	<u>AAACACATATTC</u>

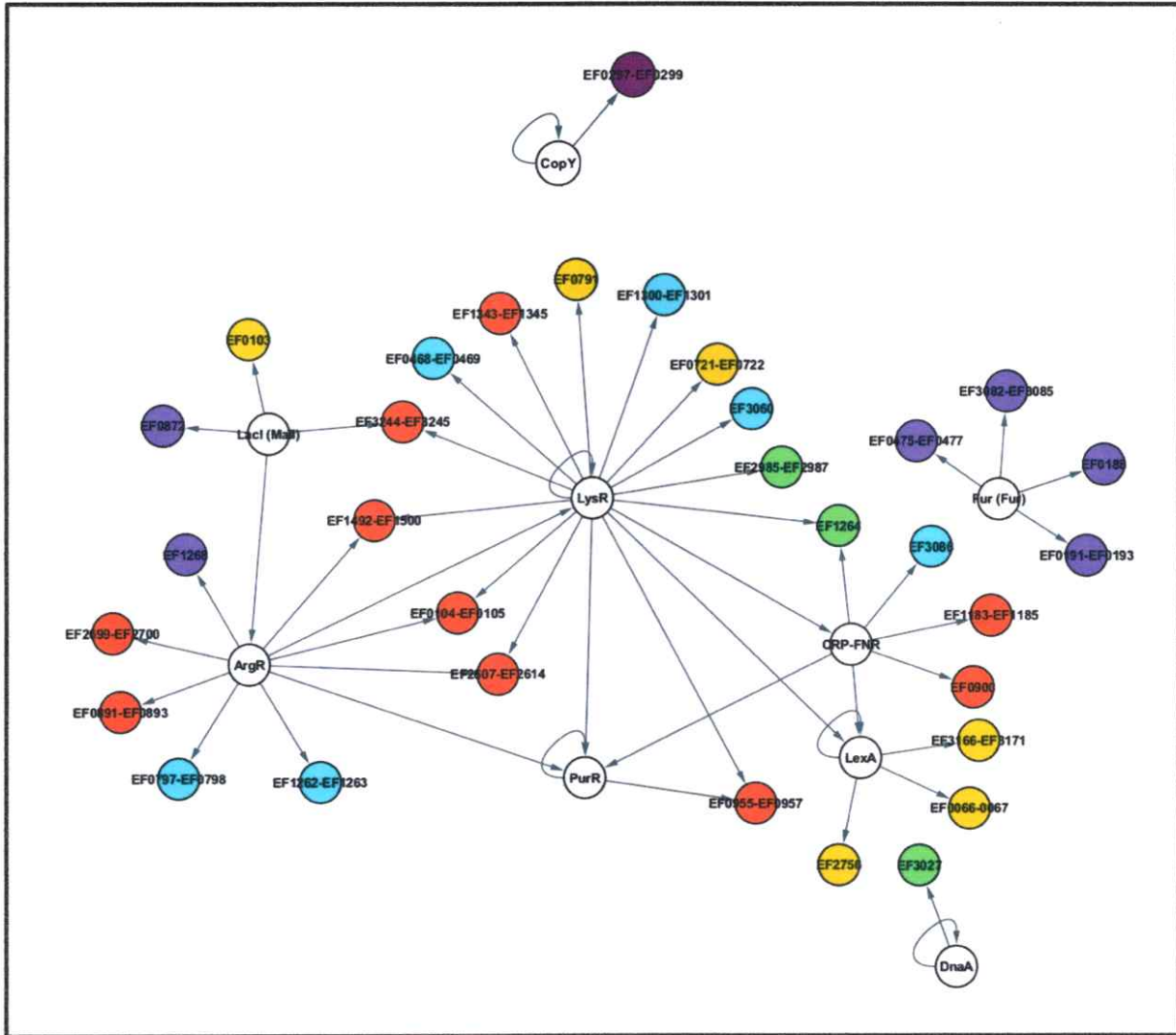


Figura 10. Sub-red de regulación transcripcional activada por Cu curada *in vivo* para los regulones CopY, Fur y LexA de *E. faecalis*. El grafo contiene 41 nodos (9 reguladores o familias de factores de transcripción y 32 blancos de regulación u operones) conectados a través de 51 enlaces (sitios de unión). Simbología de colores (agrupados por posibles funciones) para nodos (esferas): i) blancos de regulación (542), amarillo transcripción, traducción y replicación; morado transporte y señalización; naranja metabolismo basal y generación de energía; verde otras funciones; celeste sin función asignada o proteínas hipotéticas; ii) elementos reguladores en blanco (13 familias de factores de transcripción). Los enlaces se indican con flechas plumas.

3. Evaluación de la dependencia de la regulación transcripcional mediada por los factores de transcripción en la resistencia de *E. faecalis* a Cu.

3.1. Curvas de crecimiento

La activación de genes frente a altas concentraciones de Cu supone que sus productos podrían participar activamente en la resistencia frente al metal¹⁴. De acuerdo a los resultados del punto anterior, los tres factores de transcripción en estudio actuarían como represores transcripcionales de genes que se inducen en presencia del metal. Estudios previos, identificaron que la cepa silvestre OG1RF de *E. faecalis* crecida en un medio suplementado con concentraciones superiores a 0,5 mM de CuSO₄, aumenta significativamente su tiempo generacional en comparación al crecimiento en medio basal, reflejando la sensibilidad en la viabilidad que presenta la bacteria frente a la exposición a Cu^{34,36}. La caracterización previa del operon *cop*, indica que la ausencia del regulador CopY genera un aumento significativo cercano a las 36 veces de cambio en el transcrito para la ATPasa de eflujo Cu CopA (tabla 8), diferencia que no se traduce en cambios en su viabilidad en comparación a la cepa silvestre³⁶.

Considerando estos resultados, en paralelo se realizaron curvas de crecimiento bajo las mismas condiciones de tratamientos con Cu (basal, 0,5 mM, 1 mM y 2 mM de CuSO₄) para la cepa silvestre y las 3 mutantes en estudio (figura 11, tabla 9). Los resultados manifiestan al igual que lo observado en la mutante *copY*, que las cepas ausentes de los reguladores Fur y LexA no presentaron diferencias significativas con la cepa silvestre de *E. faecalis* tanto en las curvas de crecimiento como en los tiempos generacionales calculados.

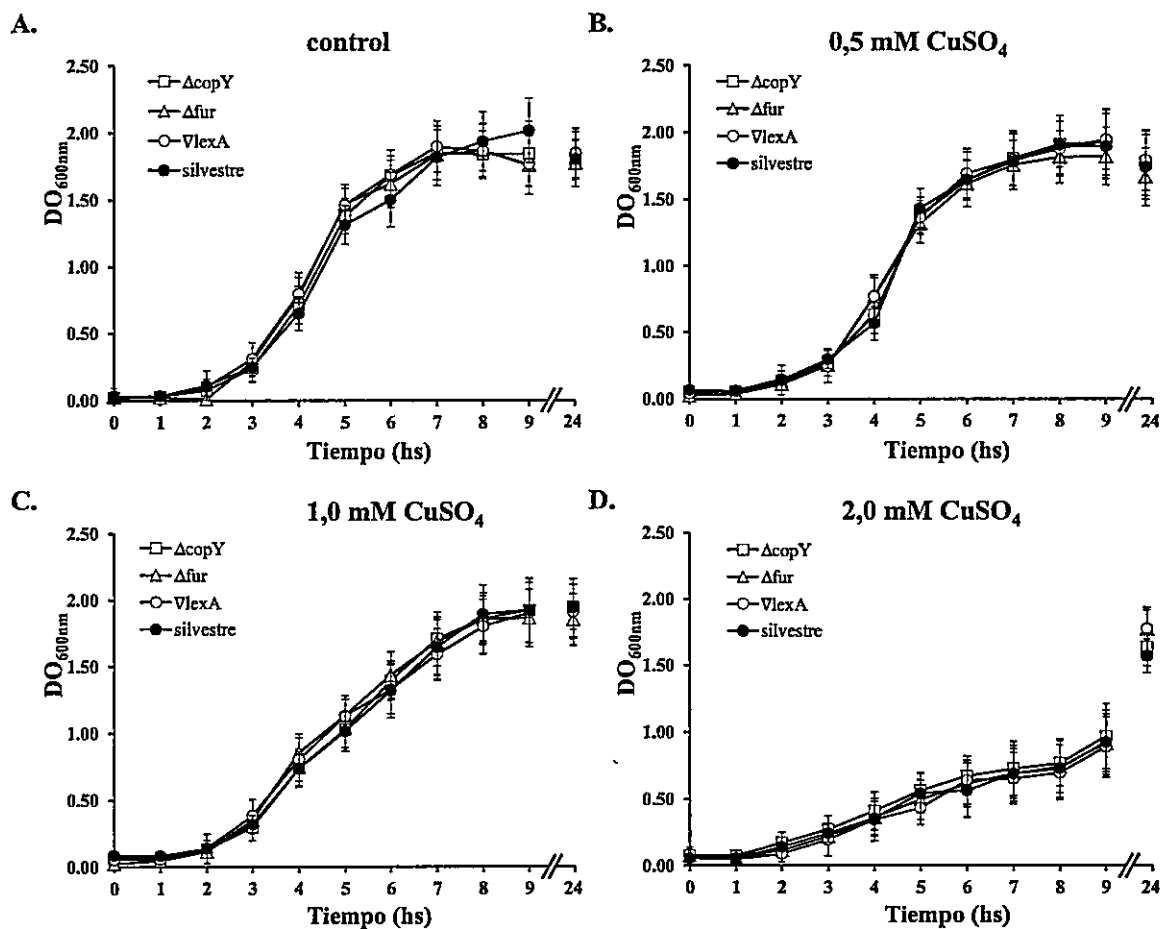


Figura 11. Curva de crecimiento para las cepas silvestre y mutantes para los genes *copY*, *fur* y *lexA* expuestas a concentraciones de Cu. A. Los paneles A-D presentan los valores de absorbancia (DO_{600nm}) obtenidos durante el crecimiento de las mutantes *copY* (cuadrados blancos), *fur* (triángulo blanco) y *lexA* (círculo blanco) y silvestre de *E. faecalis* OG1RF (círculos negros) en los medios: control (panel A), suplementados con Cu (paneles B, C y D). Cada punto representa el valor promedio de tres réplicas biológicas, las barras de error corresponden a la desviación estándar.

Tabla 9. Tiempo generacional para las cepas silvestre y mutantes para los genes *copY*, *fur* y *lexA* de *E. faecalis* expuestas a diferentes concentraciones de Cu. Para el cálculo de los tiempos generacionales se utilizó los valores obtenidos de absorbancia a los tiempos 3 y 5 h de cada curva de crecimiento de la figura 11, calculados mediante el uso de la ecuación 2 (ver sección 2.4.1 de MyM). Los valores corresponden al promedio de tres réplicas biológicas con su respectiva desviación estándar.

CuSO ₄ (mM)	Tiempo generacional (min)			
	silvestre	$\Delta copY$	Δfur	$\nabla lexA$
control	50,4 ± 4,2	48,1 ± 5,2	51,6 ± 4,4	53,9 ± 5,3
0,5	53,1 ± 5,1	51,5 ± 5,5	50,7 ± 4,9	48,7 ± 4,4
1,0	72,1 ± 6,0	67,6 ± 5,8	71,8 ± 5,0	77,6 ± 4,9
2,0	104,4 ± 4,8	114,4 ± 4,6	103,8 ± 5,2	105,2 ± 5,0

3.2. Contenido intracelular de Cu

En términos de resistencia bacteriana frente a la exposición a Cu, un aumento en las concentraciones intracelulares del metal puede romper la homeostasis celular generando especies ROS y posterior muerte celular¹. Previamente, se generó un protocolo capaz de cuantificar la concentración intracelular de Cu, definido operacionalmente como la cantidad del metal medido en el sobrenadante de un extracto, obtenido luego de la ruptura celular por sonicación y posterior centrifugación del lisado celular^{34,36}.

A diferencia de lo ocurrido en las curvas de crecimiento, si se logró observar una baja significativa en la concentración intracelular de Cu en la cepa mutante *copY* en relación a la silvestre (figura 12), posiblemente debido a un aumento en la cantidad de la ATPasa de eflujo de Cu CopA lo que estaría disminuyendo el contenido intracelular del metal. Por otra parte, los contenidos medidos en las mutantes que carecen de los reguladores LexA y Fur, al igual que lo ocurrido en los ensayos de curvas de crecimiento, no describen diferencias significativas en comparación a la cepa silvestre.

Considerando en conjunto estos resultados, el aumento significativo para los transcritos de los regulones LexA y Fur, interpretado a partir de los datos obtenidos en cepas bacterianas carentes de estos factores transcripcionales, no se condice con un efecto real tanto en la viabilidad como en el contenido de Cu, resultados que no nos permiten suponer una participación en resistencia a pesar del aumento en su nivel de expresión frente a la adición del metal.

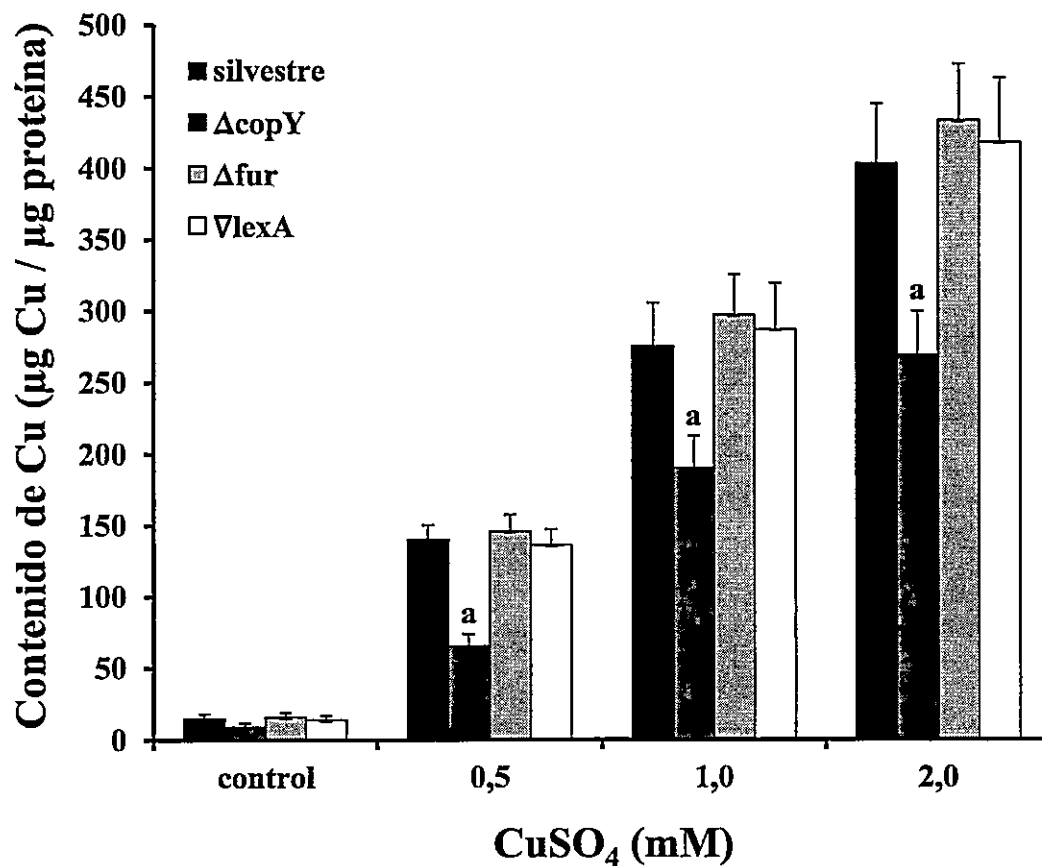


Figura 12. Contenido intracelular de Cu en las cepas silvestre y mutantes para los genes *copY*, *fur* y *lexA* de *E. faecalis*. El gráfico muestra la acumulación de Cu en el tiempo para la cepa silvestre de *E. faecalis* OG1RF (barras negras) y mutantes *copY* (barras plomo oscuras), *fur* (barras plomo claras) y *lexA* (barras blancas). El valor calculado declara el promedio de tres réplicas biológicas con tres réplicas técnicas de medición de absorbancia. La letra a corresponde a diferencias significativas con respecto a la cepa silvestre para el mismo tratamiento con Cu. ANOVA test $p < 0,05$.

DISCUSION

Cambios en la disponibilidad del Cu provocan diversos ajustes metabólicos cuyo propósito es mantener el correcto funcionamiento celular. Dentro de estos procesos, la activación y represión de genes juega un papel fundamental en controlar la abundancia de transcritos que codifican para diversos componentes involucrados en la homeostasis del metal. En cuanto a los mecanismos capaces de coordinar y regular a escala global estos procesos, así como la lógica de cómo operan estos sistemas en términos de magnitud, especificidad y temporalidad aún es un tema desconocido.

Con el propósito de entender el funcionamiento de estos mecanismos es necesario primero, identificar los componentes capaces de traducir de forma directa o indirecta el estímulo Cu con sus respectivos blancos de regulación. Es en este contexto, que la presente tesis exploró las posibles conectividades que se establecen entre factores de regulación transcripcional y genes inducidos en respuesta a Cu, basando el estudio en la construcción y verificación de una hipótesis de conectividades a escala global denominada red de regulación transcripcional activada por Cu. Tomando esto último en consideración, a continuación, se discutirán diversos aspectos referentes a la metodología y resultados obtenidos que permitieron avanzar en responder las preguntas y objetivos planteados.

Construcción de la red global de regulación transcripcional

El desarrollo de herramientas capaces de generar y recopilar una gran cantidad de información, ha permitido en estos últimos años preguntarnos a escala global, sobre el estado de la totalidad de los componentes celulares en un momento determinado. No obstante, es la integración de esta información la que ha posibilitado desde una visión teórica y experimental el poder planificar, comprender y modelar en términos cuantificables, las propiedades topológicas y dinámicas de las redes que contralan el comportamiento celular⁸¹.

De los diferentes tipos de redes biológicas estudiadas, las de regulación transcripcional permiten obtener conocimiento del circuito capaz de captar un estímulo determinado para luego traducirlo en el control de la expresión de genes implicados en la respuesta celular. En bacterias, este tipo de redes describen principalmente las interacciones entre los factores de transcripción y los genes que regulan, caracterizándose principalmente por responder a señales biológicas que en consecuencia logran cambiar la velocidad de transcripción, permitiendo que las células puedan sintetizar las proteínas necesarias en momentos y cantidades apropiadas⁹⁰.

En esta línea de investigación, este trabajo de tesis indagó sobre la existencia de redes de regulación transcripcional capaces de dar cuenta de conectividades entre genes activados por el Cu, tema no abordado hasta el momento a escala global. La identificación y posterior análisis de estas redes resulta de vital importancia para entender cuál es la articulación de los mecanismos transcripcionales involucrados en la coordinación de la expresión génica, proponiendo finalmente como podría ser la dinámica y especificidad de activación génica frente a la exposición del metal cuyo propósito en definitiva es mantener la homeostasis celular.

Primeramente, para poder identificar y caracterizar la configuración de redes de regulación transcripcional, es necesario primero, contar con un modelo que describa las posibles conectividades entre los factores de transcripción y los genes del genoma. Entre los modelos descritos en la actualidad, el primer organismo analizado fue la levadura *S. cerevisiae*²⁵. En este organismo, la estrategia de predicción y análisis de la red se construyó a partir de la información obtenida de cepas de la levadura (n = 106), que contienen diferentes factores de transcripción fusionado a un “tag” o epítipo reconocible por un anticuerpo. En extractos de células crecidas bajo diferentes condiciones de cultivo se realizaron inmunoprecipitaciones de cromatina (Chip) para las 106 cepas en estudio, los fragmentos de DNA obtenidos luego fueron hibridados en un microarreglo sembrado con todas las regiones promotoras de la levadura, obteniendo finalmente el conjunto de secuencias de los promotores reconocibles por los factores de transcripción en las diferentes condiciones experimentales. Si bien, los resultados permitieron construir una red compleja de interacciones entre factores de transcripción y regiones promotoras en *S. cerevisiae*, resulta extremadamente difícil replicar esta estrategia experimental en otros organismos.

Otra de las estrategias empleadas para generar modelos de redes transcripcionales, es la identificación de “clusters” o grupos de genes con perfiles de expresión similares mediante la asociación de resultados de microarreglos, esta aproximación supone que, genes que comparten las mismas regiones reguladoras eventualmente responden de la misma forma frente a la exposición a diferentes estímulos. Esta aproximación ha tenido resultados interesantes en organismos como *Arabidopsis thaliana* y *Drosophila melanogaster*^{91, 92}. Sin embargo, se requiere del uso de un elevado número de experimentos para obtener grupos de genes lo suficientemente diferenciados, lo cual se traduce en una seria limitante al momento de utilizar esta estrategia en organismos no modelos.

Finalmente, la tercera y más utilizada estrategia en la identificación de conectividades entre factores de transcripción y sus posibles blancos de regulación, consiste en aplicar herramientas *in silico* capaces de rastrear sitios de unión para factores de transcripción descritos en la literatura sobre las regiones promotoras del total del genoma⁹³. La búsqueda como tal, se realiza a través del uso de MPPs utilizadas como templado, estas matrices describen la probabilidad de encontrar un nucleótido en particular en una posición específica dentro de un alineamiento múltiple de sitios de unión para un factor o familia de factores determinados, aproximación que permite generar grados de flexibilidad durante la búsqueda, caso contrario al uso directo de una sola secuencia como templado que exige una alta identidad entre el sitio de unión utilizado como entrada y el genoma en cuestión.

En los últimos cinco años, una combinación de estas estrategias ha sido aplicada en forma exitosa en la construcción de modelos de redes de regulación transcripcional, tal es el caso del trabajo realizado en la bacteria *M. tuberculosis*²⁸ y el modelo recientemente publicado para *S. cerevisiae*⁹⁴. En ambos casos, el uso de un modelo *in silico* capaz de integrar datos de microarreglos, permitió identificar dentro de una red transcripcional global, sub-redes involucradas específicamente en diferentes respuestas frente a diversos estímulos^{28, 94}.

Es importante mencionar que las estrategias actuales para la generación de modelos de redes, consideran herramientas y elementos disponibles efectivamente en nuestro laboratorio para la bacteria *E. faecalis*, como por ejemplo: 1) contar con su genoma secuenciado y anotado, 2) una base de datos de sitios de unión para diferentes bacterias construida anteriormente y 3) ensayos de expresión a escala global de exposición a Cu. Por tal motivo, el uso de una estrategia *in silico* que permitiese proponer conectividades entre los factores de transcripción de la bacteria y sus potenciales genes blancos y que pudiese integrar la información de los microarreglos

previamente publicados, resultó a nuestro juicio, una estrategia adecuada para abordar la pregunta de tesis planteada.

Como punto inicial, este tipo de estrategia *in silico* demanda adicionalmente⁹⁵: i) tener conocimiento de los factores de transcripción y genes blanco de la regulación que van a constituir parte de la red y ii) dado el uso de información de sitios de unión caracterizados en otros genomas, estos elementos deben presentar características conservadas a nivel global con lo descrito previamente en otras bacterias. En este contexto, el uso de diversas herramientas bioinformáticas permitió identificar y clasificar los elementos que fueron utilizados en la construcción de la red de regulación transcripcional de *E. faecalis*. Los datos indican que la cantidad y tipo de factores de transcripción y genes blanco presentes en el genoma de *E. faecalis* concuerdan con datos reportados en la literatura para otras bacterias de genoma conocido, lo que permitió descartar la posibilidad de encontrarnos con un modelo cuya información disponible estuviese sesgada o fuese parcial, lo que limitaría o eventualmente imposibilitaría la aplicación de la estrategia *in silico*, como pudiese haber ocurrido si se hubiese elegido como organismo de estudio una bacteria cuyo genoma no se encuentre ensamblado correctamente o que su anotación presentase una baja precisión a nivel de asignación de unidades transcripcionales.

Seleccionados los nodos que van a constituir la red, se procedió a la utilización de la estrategia de construcción propiamente tal. Como se presentó en la sección de resultados, es importante mencionar que la ausencia de una base de datos con factores de transcripción y sitios de unión para el genoma de *E. faecalis*, condujo a la utilización de información presente en otras bacterias como templado de la búsqueda. Considerando variabilidades existentes entre los genomas bacterianos, las MPPs utilizadas integran información de familias de factores de transcripción y factores particulares, descritos en una o varias especies. Esto permitió definir

diferentes grados de flexibilidad dentro de la búsqueda, rangos que van desde las matrices que contienen sólo secuencias para un factor de transcripción codificado en una especie en particular (mayor estrictez) hasta las que contienen secuencias para una familia de factores de transcripción descritas en diferentes especies (menor estrictez).

La búsqueda de cada una de las MPPs se efectuó dentro de la región promotora de cada operon, definida operacionalmente como las 300 pb río arriba y 50 pb río abajo del inicio de la traducción. Tanto la longitud como la posición de la región promotora son utilizadas habitualmente en protocolos de búsqueda en este tipo de estrategia, la cual apunta a maximizar las opciones de encontrar alguna correspondencia con las matrices construidas. Si bien, la mayor cantidad de factores de transcripción reportados reconocen secuencias de unión al DNA dentro de las primeras 150 pb, existen familias de reguladores que no interactúan directamente con la maquinaria transcripcional²², tal es el caso de los factores de reposicionamiento (ejemplo MalT), los cuales pueden, por un lado mover al activador principal o alternativamente generar pliegues o loops en el DNA, en ambos casos el activador principal queda posicionado en un lugar donde es incapaz de activar la transcripción⁹⁶. Otro ejemplo son los factores de transcripción cuyo mecanismo de represión involucra la generación de loops de DNA (ejemplo GalR-R)⁷⁹, los cuales se ubican en regiones río abajo del inicio de la traducción impidiendo la interacción entre el promotor y la maquinaria transcripcional. Además, puede existir el problema técnico de una asignación automática equivocada del codón de inicio de la traducción durante la anotación del genoma de *E. faecalis*, lo que sumado a los puntos anteriores refuerzan la necesidad de considerar un rango de búsqueda desde los -300 hasta la posición +50 pb. Tanto la estrategia de asignación de elementos como la aplicación del programa fueron parte de las herramientas *in silico* utilizadas en un trabajo ya publicado⁵³ (ANEXO 2).

En cuanto a la posibilidad de encontrar sitios de unión al DNA en las regiones intergénicas comprendidas dentro de un mismo operon, en el momento de empezar el estudio no había información a nivel global que avalara esta posibilidad. Sin embargo, a finales del año 2009, Cho y colaboradores presentaron una nueva arquitectura de las unidades transcripcionales presentes en *E. coli*⁹⁷, mediante la integración de diferentes experimentos a escala global, lograron demostrar que bajo ciertas condiciones la maquinaria basal de la transcripción logra identificar regiones dentro del operon, lo que genera mas de 3500 nuevas unidades transcripcionales (>530 % de lo conocido a esa fecha). El trabajo de Cho y colaboradores⁹⁷ aporta evidencia sobre la existencia de regulación transcripcional dentro del operon, lo que permite suponer la presencia de mecanismos regulatorios (como sitios de unión para factores de transcripción) dentro de estas regiones. Si asumimos que lo observado en *E. coli* representa un mecanismo conservado en bacterias, entonces el número de unidades transcripcionales posibles en *E. faecalis* podría ascender a las 6000, número que aumenta notablemente el tiempo requerido para generar y curar hipótesis de redes de regulación transcripcional. Por tal motivo, como una proyección de este trabajo de tesis, el programa de búsqueda será reajustado para disminuir el tiempo de búsqueda de las matrices de peso y ampliar las zonas de búsquedas a las regiones intergénicas de los operones de *E. faecalis*.

Definidos los parámetros de búsqueda, el siguiente paso fue agrupar los datos de tal forma que pudiesen ser analizados manualmente. En un principio, los posibles sitios de unión encontrados en el genoma de *E. faecalis* se agruparon de acuerdo a la matriz utilizada como entrada, si bien esto permite realizar comentarios de la divergencia dentro de los sitios de unión entre las especies utilizadas como entrada y los sitios de unión encontrados en el genoma de interés, el total de elementos en *E. faecalis* ascendió a un total cercano a los 100.000 posibles

sitios de unión. Esta cantidad de opciones generalmente conlleva dificultades en términos de clasificación dado la redundancia entre los sitios detectados y la interpretación biológica de un número tan elevado de opciones de unión al DNA. En atención a lo expuesto, en este trabajo se tomó la decisión de seleccionar solo una fracción de las posibilidades generadas, específicamente se consideró solamente una secuencia nucleotídica de un conjunto de sitios de unión que compartían un nivel de identidad superior al 80% y que además, estaban asignados por nombre a un mismo factor o familia de factores de transcripción, omitiendo en definitiva su procedencia por matriz de peso probabilística utilizada en la búsqueda.

La figura 3 presenta la red de regulación transcripcional a escala global obtenida luego de la utilización de la estrategia *in silico*. Es interesante que tanto la distribución de operones (mono y policistrónicos) y las funciones de las proteínas codificadas por los genes que forman parte de la red mantienen una proporción similar a su proporción en el genoma total de la bacteria, evidenciando que no se produjeron sesgos en estos puntos durante la construcción, adjudicándole a la red características globales al modelo en estudio.

Previo al estudio de su topología propiamente tal y la discusión del funcionamiento del sistemas frente a la adición a Cu, es necesario contar con un modelo que sea capaz de reproducir en lo posible lo reportado en la literatura, información que permite tener una visión inicial de la calidad y veracidad de la red construida a nivel de las conectividades generadas y con ello disminuir la probabilidad de contar con una red ficticia. Con este propósito se verificó que: i) la calidad de los sitios de unión y el número de conectividades por factor o familia de factores de transcripción presenten algún grado de homología con lo descrito en otras bacterias y ii) logre recuperar en la mayor medida, los regulones descritos a la fecha en *E. faecalis*.

En el primer punto, no es esperable encontrar una concordancia absoluta entre las conectividades presentes en la red de *E. faecalis* y lo descrito en otras bacterias. Si bien, tanto en la asignación de familias como en la identificación de sitios de unión se utilizó información de otras especies, los programas utilizados entregan grados de libertad durante la búsqueda, cuyo objetivo finalmente es considerar la divergencia propia del organismo en estudio, tanto a nivel de la proteína, la secuencia del sitios de unión y la posición de este último dentro del genoma. Ahora bien, para familias de factores de transcripción que presentan un alto grado de conservación de estos elementos en diversas especies bacterianas, se espera un elevado nivel de homología de los regulones entre la red de *E. faecalis* y lo reportado. En atención a esto, sobre el 70% de los sitios de unión identificados se encuentran dentro de las primeras 150 pb rio arriba desde el inicio de la traducción, resultado acorde con lo caracterizado para la mayoría de los sistemas de regulación transcripcionales en bacterias. Esto da cuenta de la operación de los factores, los cuales generalmente interactúan directamente con la maquinaria de la RNA polimerasa, ya sea modificando la estabilidad del complejo molecular o afectando el reconocimiento del promotor comprendida en esta posición²².

Siguiendo con el análisis de los sitios de unión, como se aprecia en la figura 5 el Logo generado con las secuencias de cada uno de los regulones obtenidos en la red de *E. faecalis*, presenta un alto grado de homología con los consensos reportados en la literatura, información que sumada a la posición dentro de la región promotora, nos permitieron en esta etapa, suponer una alta posibilidad de encontrarnos con un alto número motivos funcionales reconocibles por los factores de transcripción asignados.

En términos de los blanco de regulación, la mayoría de los regulones presentes en la red transcripcional de *E. faecalis* exhiben un alto grado de conservación en bacterias. Tal es el caso

de los regulones para las familias LysR⁸⁰, GalR-S⁹⁸, CRP-FNR⁹⁹ y ArgR¹⁰⁰, involucrados en regular la expresión de genes que codifican para proteínas que participan en procesos metabólicos globales dentro de la célula, o los factores de la familia Fur (metales y estrés oxidativo)¹⁰¹, Fis (modificación del DNA)⁷⁶ y LexA (reparación de DNA)⁸⁸, todos ellos posiblemente regulando dentro del modelo, genes que codifican para proteínas con funciones vinculadas a estos mismos procesos.

En particular, para la familia Fur de *E. faecalis* se logró proponer regulones específicos para cada uno de sus componentes (Fur, Per y Zur). Este tipo de resultados con este nivel de depuración, se relaciona directamente con la calidad de las MPPs de entrada utilizadas durante la búsqueda. Nuestros resultados refuerzan la idea que en la medida que se disponga de mayor información de secuencias nucleotídicas reconocidas específicamente por un factor de transcripción particular, las matrices generadas logran identificar con gran precisión sitios de unión dentro de genomas bacterianos. Por el contrario, la baja disponibilidad de información con este nivel de depuración, impacta drásticamente en la cobertura de los reguladores contenidos en la red transcripcional de *E. faecalis* obtenida, la cual abarca solamente el 30% del total de factores de transcripción posiblemente codificados en la bacteria, siendo este el único sesgo presente durante la construcción de la red. Al respecto, en función de nuevos experimentos e información disponibles, luego de esta tesis, los próximos incluyen trabajar en incorporar nuevos sitios de unión a la búsqueda y con ello enriquecer la red transcripcional de *E. faecalis* aumentando su representatividad actual en términos del alcance de los elementos reguladores.

Relativo al segundo punto, previamente nuestro laboratorio presentó a la proteína CopY como el primer factor de transcripción de *E. faecalis* involucrado en regular la expresión de genes vinculados a homeostasis de Cu³⁶. Esta información fue utilizada siempre como control

positivo dentro de la búsqueda y aplicación de programas, exigiendo siempre como salida de los resultados encontrar al menos al operon *cop* como un posible blanco de regulación para CopY.

Recientemente se han caracterizado los regulones PerA¹⁰², SlyA¹⁰³ y Ers¹⁰⁴ en *E. faecalis* (relacionados principalmente con patogenicidad), siendo este último el único presente en la red de *E. faecalis*. El factor de transcripción Ers pertenece a la familia de reguladores CRP-FNR (3 posibles miembros en total codificados en *E. faecalis*), está involucrado en captar señales vinculadas a estrés oxidativo, controlando posteriormente la expresión de operones involucrados en patogenicidad (4 en total). Ninguno de estos blancos fue ligado directamente con el regulon CRP-FNR dentro de la red, esto puede explicarse ya que los autores del reporte, proponen la secuencia AACATTTGTTG como el consenso reconocido por Ers de *E. faecalis*¹⁰⁵. Como se puede observar, este consenso se aleja bastante del descrito para factores de la familia CRP-FNR (TGnATnnnnATnCA) y el obtenido dentro de la red para esta familia (TGnATnnnnATnCA, figura 5), lo cual sugiere que los sitios de unión identificados podrían pertenecer a uno de los otros dos componentes restantes de esta familia y no al factor de transcripción Ers. Este resultado abre nuevamente la posibilidad de generar matrices con las secuencia consenso de Ers, así como para los sitios de unión PerA y SlyA, los cuales serán incorporados como controles positivos en futuras búsquedas con el propósito de refinar los criterios y puntos de corte utilizados hasta el momento.

Hasta este punto, la discusión se ha enfocado principalmente en presentar y analizar información de forma descriptiva y comparativa entre la red de regulación transcripcional de *E. faecalis* y otros estudios descritos en la literatura. Evidentemente, el reciente desarrollo de esta línea de investigación, contenida dentro del área de la biología de sistemas complejos, aun no permite contar con herramientas y protocolos completamente efectivos en la construcción de

redes de regulación transcripcionales, es por esto que se vuelve necesario realizar todo este tipo de análisis previos a la interpretación en términos biológicos del funcionamiento de la red. Si bien, esta parte pudo presentarse como laboriosa dentro del desarrollo de la tesis, se logró entregar las bases necesarias para contar con una red confiable en términos del número y calidad de elementos en estudio.

Topología de la red

Los análisis topológicos aplicados entregaron información sobre la estructura y posible funcionamiento de la red transcripcional construida, dentro de ellos el tipo de elementos juega un papel fundamental en adjudicar a la red características particulares. Generalmente, estos pueden dividirse según su organización estructural en cuatro niveles⁸⁵.

El nivel más básico comprende a los elementos de la red, conformados en este caso por el conjunto de factores de transcripción, los sitios de unión y los blancos de regulación (operones). Los elementos reguladores pueden ser clasificados por un lado por su función (activador/represor) o por el número de elementos regulados. Sobre su capacidad de aumentar o reprimir la transcripción, un análisis por homología de secuencia no es suficiente para predecir o proponer una función específica para un regulador. Diversos estudios sugieren que dos factores de transcripción anotados bajo el mismo nombre en diferentes especies, los cuales comparten identidades sobre el 30% a nivel de secuencia aminoacídica, pueden cumplir funciones totalmente opuestas. Tal es el caso del regulador Fur, el cual en *E. coli* se ha descrito como represor de genes involucrados en homeostasis de Fe y que en *Neisseria gonorrhoea*, es capaz de activar la expresión de genes de metabolismo global y reparación evitando la acción de

represores de la maquinaria transcripcional¹⁰⁶. Tomando esto en consideración, nuestro nivel de análisis no nos permite predecir con exactitud cual sería su acción específica dentro de la red, por lo que no se establecerán discusiones al respecto.

En base al número de elementos regulados, podemos diferenciar dos tipos factores de transcripción, los del tipo global o maestro, cuyo sitio de unión se encuentra en diversos promotores permitiendo el control de un alto número de genes (sobre 20) y los de tipo local, capaces de regular solamente unos cuantos genes los cuales comparten funciones dentro de un mismo proceso celular. En la red construida de *E. faecalis*, genes que codifican para proteínas asociadas a procesos celulares acotados como la homeostasis de Cu (regulon CopY) y de Zn (regulon Zur) son controlados, según nuestros datos, por reguladores de características locales. El tercer factor de transcripción de carácter local identificado en *E. faecalis* fue la proteína Per, esta proteína ha sido caracterizada previamente en esta bacteria, como un componente posiblemente involucrado en el control de genes que codifican para protección contra ROS¹⁰⁷. En nuestra red de regulación, el regulon Per está formado por tres elementos sin función asignada (proteínas hipotéticas). Esta discrepancia puede explicarse por un lado, por un posible error durante nuestra búsqueda *in silico*. Si bien el logo obtenido para el regulon Per como se indica en la figura 5, conserva un alto grado de homología con el descrito en otras especies, cabe la posibilidad que la proteína Per de *E. faecalis* presente algún grado de divergencia en el dominio de unión al DNA que posiblemente se traduce en un reconocimiento de secuencias diferentes al consenso reportado. Sin embargo, una segunda opción para explicar esta discrepancia radica en la estrategia experimental utilizada en el estudio de la función de Per en *E. faecalis*¹⁰⁷. En este trabajo, los autores construyen una mutante para el factor Per, la cual en comparación a la cepa silvestre presenta diferencias en la expresión de genes vinculados a la respuesta a estrés

oxidativo bajo la exposición a peróxido de hidrógeno. Por lo tanto es posible que, los tiempos, concentraciones y las características del estímulo utilizado (capaz de generar diversas respuestas inespecíficas dentro de la célula), sumado a la no identificación de sitios de unión para Per en los genes diferencialmente expresados en respuesta a estrés, abre la posibilidad que se trate de una respuesta transcripcional tardía a la ausencia de la proteína y no un efecto directo de Per sobre la expresión génica. En el marco de la identificación del regulon Per, dado su posible relación con estrés oxidativo, se está trabajando hoy en caracterizar una nueva mutante para este factor y con ello poder determinar con exactitud sus blancos de regulación en experimentos de series de tiempo.

Según análisis en redes de *E. coli*¹⁰⁸, los factores locales corresponden a la mayor fracción de casos dentro de una red transcripcional, observación que se corresponde con nuestra red en *E. faecalis*. Hay que recordar que en nuestra red los nodos que representan a los factores de transcripción están integrados, en la mayoría de los casos, por una familia de factores y no necesariamente corresponden a una proteína en particular. Debido a esto, podemos suponer que cualquier progreso en precisar conectividades entre genes blanco y factores de transcripción individuales llevaría a un aumento en la cantidad de factores con características locales. En cuanto a los factores con características globales, en la red se destacan los factores LysR y ArgR, los cuales previamente han sido descritos en diversos organismos con estas propiedades^{80, 100}, regulando un gran número de genes que participan en diversas funciones dentro del metabolismo celular.

La posibilidad de contar con estos dos tipos de factores, le confiere al sistema el poder controlar de forma fina la transcripción, así como ampliar la respuesta transcripcional de un gen

hacia otros estímulos⁸⁶. Esta propiedad se determina principalmente por la combinatoria de más de un tipo de sitio de unión en la región promotora.

Por otra parte y sobre la base de la presencia de sitios de unión en la región promotora, se han clasificado cuatro arquitecturas diferentes¹⁰⁹: A) regulador único, B) motivo repetido, C) regulador múltiple y D) co-regulador. Los primeros dos se caracterizan por la presencia de uno o más sitios de unión al DNA para un único factor de transcripción, mientras que los otros dos contienen sitios para dos o más reguladores. En la red en estudio, los motivos del tipo regulador único se encuentran más representados (sobre el 80%), sugiriendo un sistema con un alto grado de especificidad de respuesta transcripcional frente a un estímulo. La combinatoria dentro de la red de *E. faecalis* se representa principalmente con los represores ArgR y LysR, para quienes más del 50% de sus blancos comparten regulación con otro factor de transcripción. La presencia de este tipo de reguladores además de permitir la combinatoria, permite expandir señales hacia un gran número de blancos de regulación de forma rápida (en especial para estímulos con características globales, como la presencia de estrés oxidativo) y por otro lado, cuando la bacteria se encuentra en una situación de reposo transcripcional (sin perturbaciones en el medio), logran mantener apagado un gran número de transcritos ahorrándole energía al sistema.

El tercer y último elemento basal corresponde a los operones que conforman la red, tal como se mencionó en la sección de resultados, la distribución de funciones y tipo de operones en la red mantiene las proporciones con el genoma completo. Conjuntamente, la estrategia de construcción de redes empleada puede ser utilizada para clasificar o proponer posibles funciones a genes que codifican para proteínas hipotéticas o sin función asociada. Al respecto, nuestro modelo logró integrar un total de 477 regulones activados por estímulos que además codifican para proteínas con funciones conocidas involucradas en diversos procesos celulares,

destacándose los regulones comandados por factores de transcripción locales, como el caso de Per. Los blancos hipotéticos de estos genes, actualmente están siendo analizados con el propósito de poder identificar su posible participación en la homeostasis de metales.

El segundo nivel de análisis apunta a los patrones recurrentes establecidos por las unidades básicas, las cuales aparecen frecuentemente dentro de la red. Dentro de ellos, destacan los 4 orígenes correspondientes a CopY (homeostasis de Cu), DnaA (inicio de la transcripción)¹¹⁰, Per (estrés oxidativo) y Zur (homeostasis de Zn), cuya característica principal dentro del modelo es la de traducir el estímulo de forma directa, sin la presencia de otro factor que pueda modular su expresión (sistema en cadena), propiedad del sistema que adjudica un alto grado de especificidad sobre control de la transcripción. Estos 4 orígenes además modulan la expresión de genes en *E. faecalis*, que codifican para proteínas con funciones directamente relacionadas con el estímulo, lo que supone un requerimiento particular y específico para manejarlo.

Los loops de alimentación son los patrones más recurrentes en redes transcripcionales, permiten controlar de forma fina la temporalidad y niveles expresión transcripcional así como ampliar la capacidad del sistema para responder a más de un estímulo^{27, 85}. Considerando el punto que la red de *E. faecalis* en su estado actual, no es capaz de precisar las funciones de los reguladores, se hace imposible clasificar estos patrones en sub-grupos particulares, lo que impide en definitiva especular sobre el posible funcionamiento de la red con un mayor grado de exactitud.

El tercer patrón que hemos considerado en nuestro análisis de la red regulación de *E. faecalis* se refiere a los módulos de entradas singulares. Su función principal es la coordinar un

gran número de genes que comparten una función en particular. Dentro de la red de *E. faecalis*, resaltan los factores de tipo global LysR y ArgR ya analizados en sección anterior.

Retomando la clasificación de las redes, en el tercer y cuarto nivel se presentan las agrupaciones o “clusters” capaces de diferenciarse dentro de la red global (semi-independientes), unidades denominadas módulos y sus conexiones con otros módulos. En general, las redes tienden a estar altamente interconectadas, con pequeños módulos separados del grueso de la red. Estos niveles pueden abordarse en función de los parámetros derivados del grafo de conectividades (figura 3).

De acuerdo a los parámetros globales de análisis, la red de *E. faecalis* se presenta con características de pequeño mundo a escala libre. Esta clasificación según Alberts⁸³, describe tanto en modelos de redes de regulación transcripcional de eucariontes como procariontes, el potencial de los factores de transcripción para regular un gran número de genes, donde la distribución del grado a diferencia de otras redes presenta una función exponencial más restringida, ilustrando que la regulación por combinatoria de factores ocurre en menor grado que la regulación de varios blancos de transcripción por el mismo regulador (tal como se presenta en el grafo de la figura 3), lo que en definitiva sugiere que distintas señales ambientales tienden a iniciar distintas respuestas transcripcionales en *E. faecalis* (diferentes sub-redes dependientes del estímulo).

La identificación y caracterización de la estructura de red, permiten determinar la configuración estática de los elementos con que cuenta la bacteria para responder frente a un estímulo. Diversos estudios han demostrado que, los organismos son propensos a utilizar distintas sub-redes transcripcionales reguladas por un conjunto de sensores especializados capaces de detectar aspectos específicos a partir de estímulos ambientales complejos⁸¹.

En este contexto, el siguiente paso fue abordar el funcionamiento del modelo por el cual varias partes de esta red estática de regulación funcionan de forma dinámica bajo una condición dada, con esto, es posible predecir cuales son los genes afectados por una perturbación del medio, determinando las rutas transcripcionales que dan cuenta de la propagación de la señal. Considerando el volumen y la variedad de combinaciones posibles dentro del modelo, así como de la diversidad de estímulos capaces de activar la red transcripcional, se vuelve imposible poner a prueba la capacidad total de la red en estudio. Es por esto, que la tesis se centró principalmente en usar el modelo en estudio para poder identificar sub-redes transcripcionales involucradas en la adaptación celular a Cu, micronutriente esencial para los sistemas biológicos que sin embargo, cuando su homeostasis es perturbada, se generan graves daños al correcto funcionamiento celular.

Sub-red de regulación transcripcional inducida en respuesta a Cu

Diversos han sido los esfuerzos para determinar ¿cuál es el efecto del Cu sobre *E. faecalis*?, en su mayoría enfocados a identificar y caracterizar componentes involucrados directamente en la homeostasis del metal. Experimentos de microarreglos realizados en nuestro laboratorio para *E. faecalis*, indican que genes que codifican para estos procesos solo cubren el 10% del total de cambios transcripcionales ocurridos en una situación de exposición a Cu, datos que coinciden con lo reportado en otros organismos¹¹¹⁻¹¹³.

Utilizando los datos de expresión génica en *E. faecalis* expuesta a Cu, se identificó en la red global de interacciones, aquellos genes que aumentaron su abundancia a nivel transcripcional, lo que fue definido en esta tesis como la red activada por el metal (figura 7). Este

nuevo grafo de conectividades da cuenta de una sub-red dentro del modelo global, cuya particularidad es presentar la respuesta transcripcional con sus posibles mecanismos regulatorios capaces de responder frente al estímulo Cu.

Es importante definir que la red activada se restringe a los datos de microarreglos, donde la cantidad y tiempo de exposición al metal utilizada buscaron no afectar la tasa de duplicación celular y aumentara significativamente la expresión del operon *cop* y el contenido intracelular de Cu. Con esto se logró mantener un escenario de homeostasis en la bacteria en el cual se maximizó la posibilidad de obtener cambios transcripcionales vinculados al regulador CopY.

Como es sabido, dentro de los grandes problemas actuales en la utilización de protocolos bioinformáticos es la generación de falsos positivos dentro del universo de datos a trabajar. Esto quedó manifestado al comparar las dos redes activadas generadas durante el desarrollo de la tesis (figura 7 y 10). La primera dio cuenta de una primera visión utilizando solo los datos globales de expresión y la red global construida *in silico*, la segunda, refleja el resultado final luego de la aplicación de diversos filtros los cuales permitieron curar *in vivo* tres regulones dentro del universo en estudio.

Con el propósito de entender como la bacteria estaría coordinando en términos globales su expresión para poder mantener la homeostasis celular, así como identificar cuales podrían ser los mecanismos que se están activando con sus respectivas características topológicas, a pesar del mayor grado de depuración para los tres regulones verificados *in vivo*, la discusión de la red activada consideró efectiva todas las conectividades presentes en la figura 10.

Como podemos observar dentro de ésta red activada, se destacan tres módulos de regulones de respuesta:

i) Homeostasis de Cu: Confinado esencialmente al operon *cop*. Los genes dentro de este módulo codifican exclusivamente para componentes involucrados directamente en la homeostasis de Cu. El represor de tipo local CopY es capaz de coordinar directamente la señal Cu, lo que se traduce en una des-represión del operon *cop* y con ello el aumento principalmente de la ATPasa de Cu CopA, involucrada en la salida del exceso de Cu intracelular.

ii) Daño al DNA y respuesta a estrés: Dentro de la red activada podemos encontrar principalmente dos regulones involucrados en estos procesos, el regulon LexA, el cual estaría controlando la expresión de los genes vinculados directamente al daño al DNA y el regulon Fur, que podría estar respondiendo a un efecto secundario del Cu como la generación de estrés oxidativo. Tanto el regulador CopY, LexA y Fur, serán abordados en detalle en la sección final de la discusión.

iii) Metabolismo basal: Este módulo se presenta como el más representado dentro de la red, conectando mas del 80% de los nodos activados que en su mayoría codifican para componentes involucrados directamente en la generación de energía y síntesis de moléculas basales para la recuperación de estructuras o síntesis de proteínas, principalmente regulados por los factores de transcripción ArgR y LysR.

La topología de estos tres módulos nos permite suponer que la bacteria presenta diferentes grados o niveles de especificidad en su respuesta, considerando por una parte aquellos cambios basados directamente con el efecto del Cu y por otra, la participación funcional de las proteínas codificadas por los genes activados. Interesantemente, el módulo vinculado directamente al control de la homeostasis de Cu se presentan de forma independiente al resto de la red, confinado exclusivamente al operon *cop*, el cual logra activarse de forma directa sin la

necesidad de otros reguladores para traducir la señal, comandado únicamente por el represor CopY, otorgándole al sistema un alto grado de especificidad en la respuesta.

En un segundo nivel de especificidad, se encuentra el posible efecto secundario ocurrido durante el tratamiento con el metal, basado principalmente en la generación de estrés oxidativo. En un primer nivel, encontramos al regulon Fur, la activación de sus genes podría inducir un aumento en la cantidad intracelular del Fe. Diversas enzimas que participan en la defensa contra estrés oxidativo como catalasas y peroxidasas utilizan a este metal como un cofactor¹¹⁴, lo que explicaría la activación de los genes de entrada de Fe. Este regulon se muestra independiente de la red global, otorgándole a este módulo al igual que lo ocurrido para CopY, una capacidad de respuesta específica basada directamente con el estímulo Fe.

El segundo nivel además está formado por el regulon LexA, el cual mantiene un grado de dependencia con el global de red a través de su conectividad a través de su factor de transcripción. La posibilidad de que LexA además sea regulado por otro factor de transcripción (en este caso dependiente de CRP-FNR o LysR, ambos de respuesta a estrés en general), le confiere la posibilidad de responder a otro estímulo que no sea directamente estrés oxidativo, así como controlar la activación o represión de LexA y con ello producir un segundo paso de control sobre la expresión de genes de reparación al DNA. Por otro lado, si bien LexA mantiene conectividades con el centro general de la red, este vínculo se establece a nivel de su propia regulación, mediada posiblemente por LysR y no por sus blancos, lo que eventualmente lo clasifica como otro módulo específico dentro de esta red activada por la exposición a Cu encargado de la reparación del DNA frente a estrés oxidativo.

Por último, la tercera nivel de especificidad en la respuesta da cuenta del posible requerimiento de energía o de nuevas moléculas para el correcto funcionamiento de la

homeostasis del metal y la reparación frente al daño oxidativo. Como se puede observar, los genes presentes en este módulo presentan el mayor grado de conectividad y combinatoria dentro de la red. Esta configuración guarda directa relación con el requerimiento de los productos metabólicos generados por las proteínas codificadas por los genes de estos regulones. Dentro de este módulo, el regulon PurR presenta el mayor grado de regulación dentro de la red. Su posible función es la de controlar la expresión de genes relacionados con el metabolismo de purinas¹¹⁵. El alto requerimiento de estos nucleótidos en diversos procesos celulares, posiblemente explica que PurR tenga un alto grado de conectividad. En el caso particular de la exposición a Cu y generación de ROS, dentro de las bases nitrogenadas las guaninas presentan el mayor grado de ataque (producción de oxo-guaninas)¹¹⁶, la sobre expresión del regulon PurR por lo tanto, podría dar cuenta de una renovación de estas moléculas producto del daño generado por el metal.

En definitiva, podemos decir que la topología presente en esta red activada presenta una posible configuración directamente relacionada con la especificidad de la respuesta frente a la adición de Cu, donde sus componentes se encuentran conectados de acuerdo a su requerimiento funcional en diferentes procesos metabólicos. Esta hipótesis se funda principalmente en un momento transcripcional ocurrido en la bacteria durante el proceso de adaptación a cambios a una concentración extracelular única y aun tiempo fijo en la disponibilidad del metal. Con el propósito de abordar preguntas como la magnitud, la temporalidad y la especificidad de la respuesta a Cu será necesario contar con datos a partir de otros experimentos de expresión global para ser integrados dentro de la red global y así poder tener otras sub-redes activadas por Cu.

En este contexto, el año 2011 Abrantes y colaboradores publicaron un estudio de expresión génica a escala global en *E. faecalis* en el cual se expuso a la bacteria a una concentración de 0,05 mM de CuSO₄ (un orden de magnitud inferior al usado en este trabajo de

tesis, utilizando prácticamente el mismo tiempo de exposición)¹¹⁷. La lista de genes diferencialmente expresados fue cruzada con nuestra red transcripcional global, obteniendo una segunda sub-red activada por Cu (ver ANEXO 3). En términos de número de elementos, este nuevo modelo de red se conforma por 16 operones (41 genes) conectados entre sí por 22 sitios de unión para 6 familias de factores de transcripción. El tamaño de esta red es menor a la red activada de la figura 10, esto se explica puesto que la cantidad de operones activados a 0,05 mM de CuSO₄ (n = 44) es cerca de la mitad en comparación al microarreglo realizado por nuestro laboratorio (n = 84).

El análisis de la red activada por una baja concentración de Cu, muestra dos módulos independientes, el primero compuesto por el regulon CopY (homeostasis de Cu), único módulo activado en ambas redes, lo que resalta la alta sensibilidad de este componente en la respuesta a Cu junto con su configuración particular dentro de la red. El segundo módulo da cuenta de una alta conectividad entre genes que codifican principalmente para metabolismo basal y regulación transcripcional, resaltando elementos conservados entre ambas redes para el transporte de aminoácidos (EF0891-EF0893), generación de energía (EF1492-EF1500) y metabolismo de la pared celular (EF3244-EF3245) regulados por los factores ArgR y LysR.

La ausencia de activación de los módulos Fur y LexA, así como la de otros mecanismos antioxidantes, hace suponer que la adición de 0,05 mM de CuSO₄ posiblemente no fue suficiente para generar un desbalance capaz de inducir estrés oxidativo dentro de la célula. Tal situación se correlaciona directamente con el menor número de componentes activados en la red de Abrantes, los cuales posiblemente dan cuenta de una respuesta vinculada al efecto directo del metal más que a su capacidad de formar especies reactivas.

De este análisis se desprende que *E. faecalis* es capaz de ajustar su maquinaria transcripcional de acuerdo a la magnitud y posible efecto de la exposición a Cu. Hacia una baja concentración de exposición al metal, la bacteria activa redes vinculadas directamente con el manejo del metal y sistemas capaces de generar los metabolitos necesarios para cumplir con este proceso. Cuando la concentración extracelular aumenta, además de estos procesos, se incorporan nuevas redes (módulos dependientes e independientes) capaces de manejar los efectos secundarios tóxicos y/o relacionados directamente a protección contra el daño oxidativo.

En la nomenclatura de clasificación de sub-redes transcripcionales, el análisis de estos dos modelos nos permite definir por un lado, nodos activados de forma independiente a las fluctuaciones del estímulo (módulos invariables), vinculados a la homeostasis del Cu de forma directa (chaperonas, ATPasas, etc.) e indirecta (metabolitos necesarios para la generación de energía), y nodos activados de forma dependiente, relacionados principalmente a la protección contra estrés oxidativo respuesta toxica, los cuales se activan cuando las concentraciones de exposición a Cu se acercan a generar un cambio en la viabilidad de la bacteria (0,5 mM de CuSO_4).

Esto, pone de manifiesto que la activación de redes transcripcionales de *E. faecalis* presenta umbrales de respuesta directamente relacionados con la cantidad de Cu adicionado en el medio. Ahora bien, aún queda la interrogante de la temporalidad en la ocurrencia de estos eventos, lo cual basado en los datos disponibles es imposible de abordar. No obstante, en términos generales se puede especular que regulones capaces de traducir la señal directamente (tipo CopY) podrían responder de forma más temprana que los que requieren otro tipo de estímulos para modificar su capacidad de activar o reprimir la transcripción. En cuanto a la respuesta secundaria inducida por el metal, basándonos en la activación de mecanismos de

reparación al DNA (regulon LexA) y el no observar aumento en transcritos de enzimas relacionadas directamente con mecanismos enzimáticos antioxidantes (super-óxido dismutasas y catalasas) o no enzimáticos (síntesis de glutatión), abre dos posibilidades. Una opción es que el estrés oxidativo producido en primera instancia, posterior síntesis de transcritos para componentes antioxidantes y manejo de las especies reactivas, ocurrió de forma temprana (previo a la fase exponencial media del crecimiento bacteriano). La segunda alternativa, es que nunca ocurrió un desbalance entre las ROS y los mecanismos antioxidantes presente, por lo que no se requirió un aumento en la transcripción de estos productos para mantener la homeostasis celular y que el daño observado fue consecuencia del daño directo por la toxicidad del metal, como el producido por desplazamiento de metales que forman parte de enzimas y estructuras celulares⁵. Para poder abordar este tipo de interrogantes, es necesario complementar nuestro estudio, por ejemplo con experimentos de expresión génica en series de tiempo, a fin de extraer a diferentes tiempos de exposición muestras de RNA durante una curva de crecimiento (trabajo en actual desarrollo dentro del laboratorio), con esto, podríamos tener una visión de cuáles son las sub-redes activadas en función del tiempo y así poder tener una aproximación de la dinámica de los eventos sucedidos.

El siguiente punto, radica en la especificidad de la respuesta observada. De igual manera que lo discutido en el punto anterior, el alcance de nuestro análisis no permite definir de forma clara si la activación de estas sub-redes son particulares del tratamiento con Cu. Sin embargo, considerando que estímulos como estrés oxidativo pueden generarse bajo diversas condiciones, así como que el requerimiento de metabolitos encargados de generar los productos necesarios para el correcto funcionamiento de celular, como síntesis de moléculas basales, ATP, transporte de aminoácidos y iones, son demandados bajo diferentes perturbaciones del medio, podemos

suponer que los módulos relacionados con estos procesos como los regulones Fur, LexA, LysR y ArgR podrían no ser específicos para la respuesta a Cu y por el contrario, mecanismos involucrados directamente con la homeostasis de Cu como el regulon CopY, podrían tener un mayor grado de especificidad de respuesta hacia el metal.

Con el propósito de precisar la especificidad de la red, recientemente nuestro grupo publicó un microarreglo donde la bacteria fue expuesta a una concentración no letal de Fe (ANEXO 4)¹¹⁸, además en el mismo trabajo de Abrantes¹¹⁷, se presentaron datos de la respuesta transcripcional global de *E. faecalis* frente a un tratamiento con Zn. En el marco de la publicación en progreso referente al trabajo de tesis, estos experimentos de igual forma que lo realizado para los ensayos con Cu, están actualmente siendo incorporados en la red global para poder tener comentarios sobre la especificidad de las sub-redes activadas por metales en la bacteria, identificando mecanismos conservados e independientes de respuesta en todas las combinatorias posibles para los diferentes metales.

La construcción y análisis de redes vistas desde un punto de vista sistémico, permiten estudiar y entender posibles configuraciones presentes en un organismo que le brindan al organismo la posibilidad de modular su expresión génica en respuesta de diversas perturbaciones del medio. Ahora bien, es importante considerar que dentro de este tipo de estudios, una aproximación abordada solamente desde análisis *in silico*, conlleva una gran cantidad de elementos ficticios los cuales no necesariamente están presentes en la bacteria. Producto del alto volumen de información, no es fácilmente abordable en un corto plazo la verificación *in vivo* todas las conectividades predichas *in silico*. Es por esto, que la parte siguiente del trabajo, discutirá la selección de tres regulones dentro de la red, cuyo propósito principal fue la de tener

una aproximación del grado de veracidad y alcance de la red de regulación transcripcional construida mediante el uso de diversos protocolos bioinformáticos.

Regulones CopY, Fur y LexA

Los filtros utilizados condujeron a la elección de los regulones dependientes de CopY, Fur y LexA para su verificación *in vivo*. Con este propósito, se eligió la estrategia de generar cepas mutantes de *E. faecalis* de cada uno de los factores de transcripción que comandan estos regulones, lo que permitió en definitiva determinar si la ausencia del represor genera cambios en la abundancia transcripcional de los blancos de regulación predichos *in silico* en ausencia de Cu.

El mecanismo de acción por parte de CopY, previamente ha sido descrito en diferentes organismos^{119, 120}. En una condición de represión, el factor coordina dos átomos de Zn lo que estabiliza el complejo DNA-proteína impidiendo la unión de la maquinaria de transcripción en sus genes blanco. El motivo de coordinación a Zn presenta una mayor afinidad por átomos de Cu, lo que genera que frente a un aumento de este metal en la bacteria, la probabilidad de intercambio de un átomo de Zn por Cu aumenta. La formación del complejo CopY-Cu posee una baja afinidad con el DNA, lo que conduce a su liberación permitiendo finalmente la expresión génica.

Los resultados obtenidos para el regulon CopY de *E. faecalis*, sugieren que su factor de transcripción posee un alto grado de especificidad por la secuencia TACAxxTGTA, dado que variaciones en un solo nucleótido (tabla 8), impactan en la afinidad de la proteína por el DNA anulando su acción represiva. La presencia de esta secuencia concreta solo en la región

promotora del operon *cop*, le adjudica al sistema de homeostasis de Cu un control transcripcional específico y único, el cual depende directamente de la disponibilidad de Cu en el medio.

En la bacteria *L. lactis*, la cual pertenece al mismo orden *Lactobacillales* que *E. faecalis*, este factor además de regular genes vinculados a la homeostasis de Cu, regula otros 5 operones relacionados con funciones desconocidas (posibles reductasas de CopY), todos ellos con la misma secuencia TACAxxTGTA presente en su región promotora¹²¹. Una búsqueda bioinformática manual en el genoma de *E. faecalis*, solo logró identificar un operon de este grupo, el cual no presentó la secuencia nucleotídica específica para CopY en su promotor. Este dato permitió suponer que la búsqueda global *in silico* de sitios para CopY no generó falsos negativos dentro de sus resultados.

En cuanto a la estrictez del sitio de unión reconocido, una situación similar se presentó para el regulon Fur. El análisis dio cuenta que la proteína de *E. faecalis*, al parecer es capaz de reconocer un palíndromo bastante estricto conformado principalmente por las letras TGAG/TAAxxxTTCTCA. En el caso particular del gen EF3086, si bien presenta esta secuencia en su región promotora, no se observaron cambios para su transcrito en la mutante Fur en comparación a la cepa silvestre. Esto puede explicarse puesto que la distancia hacia el inicio de la traducción del sitio de unión identificado supera las 200 pb, ubicando al represor lejos de la maquinaria transcripcional, observación que brinda información fundamental para Fur sobre la importancia de la cercanía de posición de su sitio de unión dentro de la región promotora.

Sobre la potencia en la represión del factor, los resultados sugieren que existe una relación directa entre el tamaño del palíndromo (entre 10, 12 y 14 pb) y el cambio de abundancia transcripcional entre la situación silvestre y mutante. Como se puede observar el gen EF3082-EF3085 el cual presenta una mayor diferencia a nivel transcripcional contiene el palíndromo de

mayor tamaño. Para el resto de los casos, podemos decir que el programa logró recuperar con un alto grado de efectividad este regulon, puesto que de los 5 posibles componentes 4 de ellos mostraron cambios en su abundancia transcripcional en la mutante.

En cuanto al mecanismo de como opera Fur, a la fecha se han propuesto al menos 4 tipos posibles de acción capaces de activar y reprimir de forma directa e indirecta genes principalmente vinculados con la homeostasis de Fe^{101} . En su acción clásica como represor, en presencia de una alta concentración de Fe, Fur a través de sus motivos ricos en cisteínas coordina átomos de este metal lo que aumenta su afinidad por su sitio de unión en el promotor evitando la transcripción génica de genes de captación de Fe^{122} .

Considerando el fenotipo resultante en la mutante de *E. faecalis*, al menos para los 4 operones que integran el regulon, Fur estaría actuando como un represor transcripcional. En cuanto a la posible des-represión observada en los tratamientos con Cu, dado el mecanismo de acción de Fur, cabe la posibilidad que las cisteínas de coordinación de Fe estén siendo oxidadas por la acción de ROS, probablemente como consecuencia de la adición del metal, lo que conduciría finalmente a la activación de este regulon.

Cabe destacar, que además de Fur, existen otros dos factores de transcripción que comparten la misma familia que esta proteína (Per y Zur). El sitio de unión consenso para esta familia conserva un núcleo o centro de nucleótidos común compuesto por las letras ATxAT, característica que podría generar una regulación cruzada en esta familia. En este escenario, Per y Zur podrían regular los genes que son activados en respuesta al Cu del regulon Fur; para abordar esta posibilidad, en estos momentos se está trabajando en la construcción de mutantes para estos otros dos factores. Contar con estas nuevas cepas, se podrá verificar además si las conectividades



predichas para regulones Per y Zur presentes en la red global, comparten las mismas características observadas para Fur.

Dentro de los regulones mejor descritos para Fur encontramos el de *Y. pestis*¹²³, compuesto por un total de 34 blancos de regulación que codifican para proteínas que participan en diversos procesos celulares, clasificando a este factor con características globales. Al igual que lo ocurrido para CopY, Fur en *E. faecalis* estaría confinado exclusivamente a regular genes vinculados al mismo proceso celular, en este caso homeostasis de Fe. Estas evidencias ponen de manifiesto que al parecer, la bacteria en estudio presenta mecanismos extremadamente específicos vinculados al control de las concentraciones de metales disponibles en el medio, modulados por factores directamente por la disponibilidad del estímulo. No obstante, la presencia de estímulos capaces de modificar la estructura de la proteína como por ejemplo vía oxidación de residuos, abre la posibilidad de que estos sistemas sean capaces de responder a otras situaciones, modificando las concentraciones de ambos micronutrientes frente a diversos requerimientos.

En cuanto al aumento de los transcritos vinculados a la captación de Fe, supone indirectamente un requerimiento del metal para mantener la homeostasis celular frente al exceso del metal, posiblemente vinculado a su acción como cofactor de enzimas con funciones antioxidantes. Sin embargo, diversas evidencias declaran que un aumento en la concentración de Fe intracelular puede gatillar un efecto tóxico en la célula a través de la generación de ROS. Experimentos publicados por nuestro laboratorio mostraron que cuando la concentración intracelular de Fe aumenta 10 veces sobre el contenido medido en medio basal, se disminuye significativamente la viabilidad de la bacteria. Tomando esto en cuenta, se cuantificó el

contenido intracelular de Fe en la cepa silvestre y mutante para Fur expuestas ambas a 0,5 mM de CuSO_4 (ANEXO 5).

El gráfico del anexo 5, describe un aumento significativo en el contenido de Fe para ambas cepas cuando la concentración de exposición a Cu aumenta, fenotipo explicable posiblemente por la sobre-expresión de los mecanismos de captación del metal. Este resultado además sugiere que la bacteria es capaz de controlar de forma efectiva este aumento de concentración de Fe intracelular, no superando el umbral tóxico. En cuanto al contenido entre la cepa silvestre y mutante Fur, si bien, se observa una pequeña diferencia entre la cepa silvestre y mutante en todos los tratamientos, esta no logra diferenciarse significativamente. En el caso particular de los mecanismos involucrados con la captación de Fe (sobre-expresados en la mutante de Fur), estos requieren la presencia de sideróforos los cuales van a ser transportados por estas proteínas con el propósito de capturar el Fe extracelular para su posterior incorporación, sistemas que en general son activados cuando las concentraciones de Fe son limitantes¹²⁴. Por lo tanto, si la cantidad de sideróforos se mantiene constante entre la cepa mutante y silvestre, aunque la probabilidad de encuentro sea mayor cuando hay mayor cantidad de transportadores, esta abundancia podría no ser suficiente para ejercer un efecto significativo en la captación del metal.

El tercer y último regulon verificado corresponde al factor LexA. Esta proteína forma un sistema acoplado con RecA, en diversas bacterias coordina la expresión de más de 30 genes, vinculados principalmente con sistemas de reparación de DNA¹²⁵. El mecanismo de cómo opera LexA ya fue presentado en la sección de resultados, destacándose su acción como represor transcripcional.

Según los resultados obtenidos en *E. faecalis*, al parecer esta proteína también controlaría genes relacionados con el daño a DNA. El análisis del sitio de unión a diferencia de lo mostrado por CopY y Fur, permite un mayor grado de variabilidad en su secuencia palindrómica, lo cual podría impactar directamente en la capacidad represiva del factor sobre la transcripción¹²⁶. Tal como lo muestran los datos de qPCR para los tres regulones en estudio, mientras mayor es la divergencia del palíndrome menor es la diferencia de abundancia de transcrito entre la cepa mutante y silvestre.

En *E. coli*, se ha demostrado que la capacidad de LexA de reprimir su propia expresión y con ello disminuir la abundancia de su transcrito, de esta forma permite un aumento en la velocidad de activación de sus blanco de regulación¹²⁷, mecanismo adaptado especialmente para una respuesta potenciada y rápida de genes involucrados directamente en la supervivencia de la bacteria, situación que eventualmente podría estar ocurriendo en *E. faecalis*, donde al parecer también podría estar auto-reprimiéndose. Este loop de auto-regulación denominado coherente, no ocurre en el caso del operon *cop*. La activación del operon genera un aumento del represor CopY, dando cuenta de un sistema incoherente, mecanismos utilizados por los sistemas biológicos para controlar de forma fina la expresión del gen. Ahora bien, cuando existe un requerimiento elevado de las otras proteínas Cop, en particular una sobre-expresión de la ATPasa CopA cuando los niveles de Cu son extremadamente tóxicos, existen otros mecanismos no transcripcionales capaces de degradar o inactivar a las proteínas Cop¹²⁸.

Si comparamos las razones de abundancia obtenidas por qPCR para los tres regulones en estudio, las diferencias obtenidas por la adición del metal son menores que las reportadas en la cepa mutante (tabla 7 y 8). Esto se explica por el balance que existe en la cepa silvestre, entre el factor de transcripción en su forma activa (con capacidad para reprimir) e inactiva, balance que

al adicionar Cu eventualmente se desplaza hacia la des-represión. En una cepa mutante, este balance es inexistente puesto que no existe la capacidad de reprimir al sistema.

En términos generales, la estrategia de generación de mutantes para factores de transcripción permite verificar y curar con un alto grado de efectividad posibles blancos transcripcionales. En modelos eucariontes donde la transcripción es controlada por un gran número de elementos reguladores, llegando a tener 20 elementos reguladores de la transcripción interactuando de forma combinatoria en un solo promotor, muchas veces la ausencia de un factor en particular no genera fenotipos detectables¹²⁹. En bacterias, los modelos transcripcionales conocidos describen que la combinatoria es baja, por lo que la ausencia de un regulador genera cambios significativos de expresión, no obstante la presencia de un segundo factor de transcripción incapaz de responder a Cu, podría explicar que no todos los operones contenidos para un mismo regulon dentro de la red global se activaron en presencia del metal. Como en el caso del regulon LexA, que de un total de 33 posibles regulones solo 5 mostraron un aumento a nivel de transcrito cuando se agregó Cu al medio. Además, dentro de este regulon activado, los operones EF0955-EF0957 y EF3027 no presentaron cambios en la cepa mutante. Esto puede explicarse dado que la secuencia palindrómica tipo LexA que presentan estos dos elementos es la de menor tamaño dentro del regulon. Cabe la posibilidad que la técnica de utilización de mutantes y qPCR no tenga el grado de resolución suficiente para descartar la unión de LexA sobre estas secuencias. Para este tipo de situaciones, a la fecha existen diferentes aproximaciones experimentales que complementan con mayor información sobre la interacción entre el factor de transcripción y el sitio de unión, lamentablemente la mayoría de ellas permiten estudiar solo un componente, implicando en la mayoría de los casos, purificar un factor de transcripción en particular (EMSA, termoforesis a microescala o manipulación de moléculas individuales) o

construir diferentes vectores reporteros para cada uno de los sitios de unión predichos, lo que demanda mucho tiempo y dinero si se quiere abordar una red en términos globales. Recientemente, Michael Berger y Martha Bulyk presentaron una nueva estrategia denominada microarreglos de unión de proteína (PBM)¹³⁰, si bien logra determinar a gran escala la capacidad de unión de un factor de transcripción sobre más de 1000 posibles sitios de unión, su principal limitante es que sólo permite analizar un regulador el cual debe purificarse expresarse y purificarse previamente.

En otro aspecto, nuestros datos indican que los sitios de unión para *E. faecalis* presentan un alto grado de homología con los consensos reportados en otras bacterias (figura 5). Es importante mencionar que el evento de unión efectiva entre el factor de transcripción y el DNA, no solo depende de la secuencia nucleotídica del sitio de unión presente en el promotor, sino también la proteína debe tener residuos específicos y una configuración tridimensional capaz de reconocer un sitio de unión particular y no otro. Tomando esto en cuenta, se generaron tres modelos tridimensionales de las proteínas CopY, Fur y LexA de *E. faecalis* generados por homología utilizando datos cristalográficos disponibles en la literatura (ANEXOS 6 y 7). Los modelos mostraron que las tres proteínas presentan un algo grado de identidad en su conformación tridimensional con los cristales de referencia y que, los residuos determinantes de la unión al DNA se encuentran en las posiciones y distancias observadas en el cristal, lo que apoya el hecho de que el reconocimiento por los sitios de unión al DNA predichos son reconocidos en *E. faecalis*.

En términos del tipo de elementos regulados, de igual forma a lo observado en otros organismos, CopY y Fur, LexA también regulan un alto número de operones vinculados con distintos procesos celulares. Estas evidencias suponen un alto grado de divergencia en términos

de la capacidad y especificidad que presentan los factores de transcripción de regular operones dentro del genoma. Para poder avanzar en la comprensión de este fenómeno, estamos trabajando en generar una mayor cantidad de información para bacterias del orden *Lactobacillale* (ANEXO 8). Con la información generada se pretende evaluar si es una propiedad característica del orden o propia de *E. faecalis*, estudiando principalmente cuáles son las divergencias que se establecen en términos de sus elementos básicos (factores de transcripción, sitios de unión y blancos de regulación) y sus parámetro topológicos a nivel de la configuración de la red global.

El tercer y último objetivo de la tesis, mediante ensayos microbiológicos buscó estudiar si la ausencia de la regulación transcripcional de nodos activados por el Cu, vinculados directamente con la homeostasis de Cu y posiblemente con la homeostasis de Fe y resistencia a estrés, impacta sobre la viabilidad celular y el contenido intracelular del metal.

Basado en la posible función de las proteínas codificadas en estos operones, vinculadas principalmente a la resistencia contra un exceso del metal, como esperado para los ensayos de viabilidad celular era encontrar una mayor tolerancia a Cu para las mutantes en comparación a la cepa silvestre, fenotipo no observado en ninguno de los experimentos. Este resultado desde un punto de vista sistémico puede explicarse por; que si bien el proceso de transcripción y traducción en bacterias es un mecanismo que puede estar acoplado¹³¹, lo que en la mayoría de los caso supone que un aumento en la cantidad de transcrito se correlaciona con una mayor cantidad de proteína, cabe la posibilidad de que ésta no este activa o no posea los sustratos suficientes para su función (como el caso de los mecanismos de entrada de Fe comandados por Fur y la cantidad de sideróforos disponibles en el medio). Se señaló que ciertos módulos activados por Cu están encargados, al parecer, de generar la energía necesaria para el correcto funcionamiento de sistemas involucrados en la resistencia contra el metal. En una cepa mutante, parte del ATP

podría estar siendo utilizado en la formación de estas proteínas expresadas en forma anormal, disminuyendo eventualmente la disponibilidad de esta molécula para procesos de resistencia, como los transportadores de Fe y mecanismos de reparación de DNA, ambos dependientes de la molécula de ATP para su correcto funcionamiento^{132, 133}.

En el caso particular de la mutante de CopY, se pudo observar una disminución en el contenido intracelular del metal, explicado posiblemente por el aumento de la actividad ATPasa de salida CopA producto de la sobre-expresión del operon *cop*. Esta disminución en todos los casos estudiados, no es suficiente para disminuir los efectos tóxicos del metal, los cuales no sólo afectan moléculas en el citoplasma de la bacteria, se ha descrito que altas concentraciones de exposición a Cu logran dañar componentes de la membrana y pared celular de forma extracelular¹³⁴.

Otros experimentos actualmente se están llevando a cabo para caracterizar estas mutantes en términos de su relación con la resistencia al metal. Como alternativas, se pretende exponer a la bacteria en una situación limitante de Fe en un medio con concentraciones crecientes de Cu, en esta situación esperamos que la mutante de Fur tenga un mejor desempeño que la cepa silvestre (mayor viabilidad) en este este escenario. Del mismo modo, aumentar la probabilidad de generación de daño oxidativo producto del Cu, como por ejemplo adicionando un agente prooxidante al medio (H_2O_2), podría evidenciar una diferencia en la viabilidad entre la mutante LexA y la cepa silvestre.

Los resultados en general hacen suponer que la red global actual contiene un alto número de falsos positivos dentro de sus operones blanco. En este aspecto, el análisis de mutante nos permitió depurar el sitio de unión para los regulones CopY, Fur y LexA, información que actualmente esta siendo incorporada en la búsqueda y generación de MPP para refinar la

predicción de genes blanco en *E. faecalis*. Es importante hacer notar, que el alcance de las conclusiones obtenidas durante este estudio posee diferentes limitantes. Primero encontramos la cantidad de microarreglos incorporados, el disponer de un bajo número de muestras entrega una alta probabilidad de que alguno de los cambios transcripcionales sean producto de una simple casualidad y por lo tanto no representen necesariamente un efecto significativo en la adaptación hacia el estímulo. Las sub-redes analizadas no dan cuenta de los genes no activados (no presentaron cambio transcripcional o su expresión fue hacia la baja), descartados del análisis por diversas razones expuestas anteriormente. El número de elementos que considera la red aborda alrededor de un 30% del total de genes y factores de factores de transcripción contenidos en el genoma de *E. faecalis*. Pensando integrar un mayor número de elementos, dado el aumento exponencial de genomas bacterianos disponibles en la literatura en este último tiempo¹³⁵, actualmente el Laboratorio de Bioinformática y Matemática del Genoma en colaboración con nuestro grupo nos encontramos re-anotando el genoma de *E. faecalis* V583, secuenciada en el año 2003 por el grupo de Paulsen⁴³. Esta nueva información permitirá precisar en gran medida los sitios de inicio de la transcripción y traducción de los operones así como, dentro de las proteínas con funciones desconocidas o hipotéticas (cerca del 33% del total del genoma) se podrán identificar nuevos factores de transcripción los cuales eventualmente podrán formar parte de la red transcripcional desarrollada. El contar con un mayor número de elementos, podría mejorar significativamente la comprensión del fenómeno de adaptación y funcionamiento de la red. La red en su estado actual además, no considera otros mecanismos independientes de los factores de transcripción los cuales pueden modificar la expresión génica, como los factores sigma, pequeños ligandos (ppGpp) y sRNA^{22, 136}.

Considerando estas limitaciones, nuestro modelo actual de redes logró presentar a nivel sistémico que *E. faecalis* cuenta con módulos de activación, que en su conjunto constituyen dentro de una red global de interacciones, una sub-red transcripcional activada por el Cu. Logramos proponer que la bacteria frente a la adición del metal, presenta la activación de: i) módulos independientes con un alto grado de especificidad como es el regulon CopY, el cual se acota exclusivamente a controlar genes involucrados en la homeostasis del Cu, y ii) módulos altamente conectados, capaces de procesar señales secundarias más complejas dentro la red activada, como los regulones LysR y ArgR. Además, utilizando datos de otros microarreglos, logramos detectar que si la señal Cu es baja (red de Abrantes, 0,05 mM de CuSO₄), la bacteria activa esencialmente una menor cantidad de regulones, vinculados principalmente a la respuesta directa frente al metal mientras que el resto de la red, cuyo potencial puede ser activado por una mayor concentración de Cu. De tal forma, cuando la señal Cu se intensifica (red activada a 0,5 mM de CuSO₄), se integran nuevos regulones con características específicas a los efectos indirectos del metal, en este caso resistencia frente a estrés oxidativo (regulones Fur y LexA), evidenciando que la activación de las redes transcripcionales tiene una directa relación con los efectos producidos por el metal, lo que guarda estrecha relación con la existencia de unidades de procesamiento de señales que funcionan de forma específica y simple, resguardando el ahorro energético para la célula.

Finalmente, las evidencias hasta el momento suponen que la resistencia hacia el Cu es un proceso que no depende exclusivamente de operones vinculados directamente a la homeostasis del metal, donde los efectos tóxicos generados por el metal son tratados por la bacteria a un nivel sistémico, modulando su expresión activando diferentes componentes que le adjudican adaptabilidad al sistema. En este contexto, la hipótesis del trabajo propuso que la respuesta

transcripcional de la bacteria frente a la exposición a Cu se establece a través de factores comunes capaces de coordinar la expresión génica.

Hoy podemos decir que la secuencia del genoma determina la potencialidad de un organismo para responder frente a un estímulo, no obstante del como la célula logra a nivel sistémico, controlar su expresión génica aún es un tema que esta lejos de ser entendido. Donde la pregunta como postula Balázsi sigue siendo⁸¹: ¿Es posible predecir *a priori* qué parte de la red es capaz de responder a un estímulo determinado? La respuesta es probablemente sí, si conocemos los sensores para el estímulo y la topología de red completa. Es en este campo de investigación, la presente tesis presentó un modelo global de redes de regulación transcripcional, cuyo propósito principal fue el de entregar información suficiente que nos permitiese entender como puede ser la adaptabilidad de un organismos en términos de traducción de la señal sobre la coordinación y regulación de su expresión génica frente a la exposición a Cu. Esperamos que este trabajo sirva a futuro para comprender la respuesta transcripcional a otros estímulos en la bacteria, brindando a la comunidad científica de un modelo capaz de integrar diferente información desde una visión de la biología de sistemas complejos.

CONCLUSIONES

La red de regulación transcripcional propuesta para *E. faecalis* presenta una topología con propiedades globales conservadas con otros organismos lo cual sugiere un alto grado de confiabilidad para el conjunto de conectividades predichas entre los factores de transcripción y operones codificados en la bacteria.

El modelo logró incorporar información de la respuesta transcripcional global inducida de *E. faecalis* frente a la exposición a Cu, lo que permitió identificar un conjunto de regulones interconectados que pasaron a constituir el universo de sub-redes activadas por el metal.

La construcción y caracterización de mutantes para un grupo particular de factores de transcripción permitieron por un lado, depurar el modelo *in silico* de redes así como, sugerir que la ausencia de estos reguladores a pesar de los cambios fenotipos observados, no impacta significativamente en la resistencia contra el metal, medido como cambios en la viabilidad y proliferación celular.

La integración de todos los resultados sugieren que *E. faecalis* presenta módulos transcripcionales específicos, los cuales se activan dependiendo de las concentraciones de Cu presentes en el medio, cuya características a nivel del grado de conectividad y capacidad de repuesta, depende fuertemente de la especificidad funcional de los componentes codificados en los operones.

Es importante acentuar que el estado actual del modelo requiere la integración de otros procesos celulares que influyen en la regulación génica, como también la constante adición de nueva información para su verificación y cobertura.

Finalmente, las capacidades generadas durante el desarrollo de la tesis no solo brindaron un modelo de estudio con potencialidad para entender como es la coordinación, dinámica y especificidad a nivel global de la respuesta transcripcional de *E. faecalis* frente a diferentes estímulos, además se entregó un protocolo de construcción de redes extrapolable para otras bacterias con genoma anotado.

REFERENCIAS BIBLIOGRAFICAS

1. Linder, M.C. & Hazegh-Azam, M. Copper biochemistry and molecular biology. *Am J Clin Nutr* 63, 797S-811S (1996).
2. Uauy, R., Olivares, M. & Gonzalez, M. Essentiality of copper in humans. *Am J Clin Nutr* 67, 952S-959S (1998).
3. Halliwell, B. & Gutteridge, J.M. Free radicals, lipid peroxidation, and cell damage. *Lancet* 2, 1095 (1984).
4. Yoshida, Y., Furuta, S. & Niki, E. Effects of metal chelating agents on the oxidation of lipids induced by copper and iron. *Biochim Biophys Acta* 1210, 81-88 (1993).
5. Predki, P.F. & Sarkar, B. Effect of replacement of "zinc finger" zinc on estrogen receptor DNA interactions. *J Biol Chem* 267, 5842-5846 (1992).
6. Kim, B.E., Nevitt, T. & Thiele, D.J. Mechanisms for copper acquisition, distribution and regulation. *Nat Chem Biol* 4, 176-185 (2008).
7. Bird, A.J. Metallosensors, the ups and downs of gene regulation. *Adv Microb Physiol* 53, 231-267 (2008).
8. Permina, E.A., Kazakov, A.E., Kalinina, O.V. & Gelfand, M.S. Comparative genomics of regulation of heavy metal resistance in Eubacteria. *BMC Microbiol* 6, 49 (2006).
9. Bertini, I. & Rosato, A. Menkes disease. *Cell Mol Life Sci* 65, 89-91 (2008).
10. Bingham, M.J., Ong, T.J., Summer, K.H., Middleton, R.B. & McArdle, H.J. Physiologic function of the Wilson disease gene product, ATP7B. *Am J Clin Nutr* 67, 982S-987S (1998).
11. Bartee, M.Y. & Lutsenko, S. Hepatic copper-transporting ATPase ATP7B: function and inactivation at the molecular and cellular level. *Biometals* 20, 627-637 (2007).
12. Squitti, R. Metals in Alzheimer's disease: a systemic perspective. *Front Biosci* 17, 451-472 (2012).
13. Zheng, Y., Li, X.K., Wang, Y. & Cai, L. The role of zinc, copper and iron in the pathogenesis of diabetes and diabetic complications: therapeutic effects by chelators. *Hemoglobin* 32, 135-145 (2008).
14. Gonzalez, M., Reyes, A., Suazo, M., Jo, W. & Vulpe, C.D. Expression of copper-related genes in response to copper load. *Am J Clin Nutr* 88 (2008).
15. Winge, D.R. Copper-regulatory domain involved in gene expression. *Prog Nucleic Acid Res Mol Biol* 58, 165-195 (1998).
16. Perego, P. & Howell, S.B. Molecular mechanisms controlling sensitivity to toxic metal ions in yeast. *Toxicol Appl Pharmacol* 147, 312-318 (1997).
17. De Freitas, J., Wintz, H., Kim, J.H., Poynton, H., Fox, T. & Vulpe, C. Yeast, a model organism for iron and copper metabolism studies. *Biometals* 16, 185-197 (2003).
18. Odermatt, A. & Solioz, M. Two trans-acting metalloregulatory proteins controlling expression of the copper-ATPases of *Enterococcus hirae*. *J Biol Chem* 270, 4349-4354 (1995).
19. Otsuka, F., Ohno, S., Suzuki, K., Takahashi, K., Ohsawa, M. & Koizumi, S. Mechanism of metallothionein gene activation mediated by heavy-metal dependent transcription factor MTF-1. *Yakugaku Zasshi* 127, 675-684 (2007).
20. Latchman, D.S. Transcription factors: an overview. *Int J Biochem Cell Biol* 29, 1305-1312 (1997).

21. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* 386, 569-577 (1997).
22. Browning, D.F. & Busby, S.J. The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2, 57-65 (2004).
23. Nam, J., Dong, P., Tarpine, R., Istrail, S. & Davidson, E.H. Functional cis-regulatory genomics for systems biology. *Proc Natl Acad Sci U S A* 107, 3930-3935 (2010).
24. Davidson, E. & Levin, M. Gene regulatory networks. *Proc Natl Acad Sci U S A* 102, 4935 (2005).
25. Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Murray, H.L., Gordon, D.B., Ren, B., Wyrick, J.J., Tagne, J.B., Volkert, T.L., Fraenkel, E., Gifford, D.K. & Young, R.A. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 298, 799-804 (2002).
26. Janga, S.C., Salgado, H., Martinez-Antonio, A. & Collado-Vides, J. Coordination logic of the sensing machinery in the transcriptional regulatory network of *Escherichia coli*. *Nucleic Acids Res* 35, 6963-6972 (2007).
27. Alon, U. Network motifs: theory and experimental approaches. *Nat Rev Genet* 8, 450-461 (2007).
28. Balazsi, G., Heath, A.P., Shi, L. & Gennaro, M.L. The temporal response of the *Mycobacterium tuberculosis* gene regulatory network during growth arrest. *Mol Syst Biol* 4, 225 (2008).
29. Kitano, H. Systems biology: a brief overview. *Science* 295, 1662-1664 (2002).
30. Okuda, S. & Yoshizawa, A.C. ODB: a database for operon organizations, 2011 update. *Nucleic Acids Res* 39, D552-555 (2011).
31. Rensing, C. & Grass, G. *Escherichia coli* mechanisms of copper homeostasis in a changing environment. *FEMS Microbiol Rev* 27, 197-213 (2003).
32. Solioz, M. & Stoyanov, J.V. Copper homeostasis in *Enterococcus hirae*. *FEMS Microbiol Rev* 27, 183-195 (2003).
33. Reyes, A., Leiva, A., Cambiazo, V., Mendez, M.A. & Gonzalez, M. Cop-like operon: structure and organization in species of the Lactobacillales order. *Biol Res* 39, 87-93 (2006).
34. Latorre, M., Olivares, F., Reyes-Jara, A., Lopez, G. & Gonzalez, M. CutC is induced late during copper exposure and can modify intracellular copper content in *Enterococcus faecalis*. *Biochem Biophys Res Commun* 406, 633-637 (2011).
35. Murray, B.E., Singh, K.V., Ross, R.P., Heath, J.D., Dunny, G.M. & Weinstock, G.M. Generation of restriction map of *Enterococcus faecalis* OG1 and investigation of growth requirements and regions encoding biosynthetic function. *J Bacteriol* 175, 5216-5223 (1993).
36. Reyes-Jara, A., Latorre, M., Lopez, G., Bourgoigne, A., Murray, B.E., Cambiazo, V. & Gonzalez, M. Genome-wide transcriptome analysis of the adaptive response of *Enterococcus faecalis* to copper exposure. *Biometals* 23, 1105-1112 (2010).
37. Garsin, D.A., Urbach, J., Huguet-Tapia, J.C., Peters, J.E. & Ausubel, F.M. Construction of an *Enterococcus faecalis* Tn917-mediated-gene-disruption library offers insight into Tn917 insertion patterns. *J Bacteriol* 186, 7280-7289 (2004).
38. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166, 557-580 (1983).

39. Leenhouts, K., Buist, G., Bolhuis, A., ten Berge, A., Kiel, J., Mierau, I., Dabrowska, M., Venema, G. & Kok, J. A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol Gen Genet* 253, 217-224 (1996).
40. Messing, J., Crea, R. & Seeburg, P.H. A system for shotgun DNA sequencing. *Nucleic Acids Res* 9, 309-321 (1981).
41. Nallapareddy, S.R., Singh, K.V. & Murray, B.E. Construction of improved temperature-sensitive and mobilizable vectors and their use for constructing mutations in the adhesin-encoding *acm* gene of poorly transformable clinical *Enterococcus faecium* strains. *Appl Environ Microbiol* 72, 334-345 (2006).
42. Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. & Leunissen, J.A. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* 35, W71-74 (2007).
43. Paulsen, I.T., Banerjee, L., Myers, G.S., Nelson, K.E., Seshadri, R., Read, T.D., Fouts, D.E., Eisen, J.A., Gill, S.R., Heidelberg, J.F., Tettelin, H., Dodson, R.J., Umayam, L., Brinkac, L., Beanan, M., Daugherty, S., DeBoy, R.T., Durkin, S., Kolonay, J., Madupu, R., Nelson, W., Vamathevan, J., Tran, B., Upton, J., Hansen, T., Shetty, J., Khouri, H., Utterback, T., Radune, D., Ketchum, K.A., Dougherty, B.A. & Fraser, C.M. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299, 2071-2074 (2003).
44. van Bakel, H., Huynen, M. & Wijnenga, C. Prokaryotic diversity of the *Saccharomyces cerevisiae* Atx1p-mediated copper pathway. *Bioinformatics* 20, 2644-2655 (2004).
45. Che, D., Li, G., Mao, F., Wu, H. & Xu, Y. Detecting uber-operons in prokaryotic genomes. *Nucleic Acids Res* 34, 2418-2427 (2006).
46. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, 3389-3402 (1997).
47. Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. & Lopez, R. InterProScan: protein domains identifier. *Nucleic Acids Res* 33, W116-120 (2005).
48. Sonnhammer, E.L., Eddy, S.R. & Durbin, R. Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* 28, 405-420 (1997).
49. Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M.D., Durbin, R., Falquet, L., Fleischmann, W., Gouzy, J., Hermjakob, H., Hulo, N., Jonassen, I., Kahn, D., Kanapin, A., Karavidopoulou, Y., Lopez, R., Marx, B., Mulder, N.J., Oinn, T.M., Pagni, M., Servant, F., Sigrist, C.J. & Zdobnov, E.M. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res* 29, 37-40 (2001).
50. Rodriguez-Tome, P., Stoehr, P.J., Cameron, G.N. & Flores, T.P. The European Bioinformatics Institute (EBI) databases. *Nucleic Acids Res* 24, 6-12 (1996).
51. Thijs, G., Moreau, Y., De Smet, F., Mathys, J., Lescot, M., Rombauts, S., Rouze, P., De Moor, B. & Marchal, K. INCLUSive: integrated clustering, upstream sequence retrieval and motif sampling. *Bioinformatics* 18, 331-332 (2002).
52. Benitez-Bellon, E., Moreno-Hagelsieb, G. & Collado-Vides, J. Evaluation of thresholds for the detection of binding sites for regulatory proteins in *Escherichia coli* K12 DNA. *Genome Biol* 3, RESEARCH0013 (2002).

53. Hodar, C., Moreno, P., di Genova, A., Latorre, M., Reyes-Jara, A., Maass, A., Gonzalez, M. & Cambiazo, V. Genome wide identification of *Acidithiobacillus ferrooxidans* (ATCC 23270) transcription factors and comparative analysis of ArsR and MerR metal regulators. *Biometals* 25, 75-93 (2012).
54. Huerta, A.M., Salgado, H., Thieffry, D. & Collado-Vides, J. RegulonDB: a database on transcriptional regulation in *Escherichia coli*. *Nucleic Acids Res* 26, 55-59 (1998).
55. Perez, A.G., Angarica, V.E., Vasconcelos, A.T. & Collado-Vides, J. Tractor_DB (version 2.0): a database of regulatory interactions in gamma-proteobacterial genomes. *Nucleic Acids Res* 35, D132-136 (2007).
56. Makita, Y., Nakao, M., Ogasawara, N. & Nakai, K. DBTBS: database of transcriptional regulation in *Bacillus subtilis* and its contribution to comparative genomics. *Nucleic Acids Res* 32, D75-77 (2004).
57. Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A. & Barrell, B. Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944-945 (2000).
58. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498-2504 (2003).
59. derechos, M.C.R.t.l. Microsoft Office 2010 (2010).
60. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *International Journal on Complex Systems ICCS2006* (2006).
61. Hirose, M., Totoki, Y., Hoshida, M. & Ishikawa, M. Comprehensive study on iterative algorithms of multiple sequence alignment. *Comput Appl Biosci* 11, 13-18 (1995).
62. Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. WebLogo: a sequence logo generator. *Genome Res* 14, 1188-1190 (2004).
63. Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Wheeler, D.L. GenBank. *Nucleic Acids Res* 35, D21-25 (2007).
64. Zhou, D. & Yang, R. Global analysis of gene transcription regulation in prokaryotes. *Cell Mol Life Sci* 63, 2260-2290 (2006).
65. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29, e45 (2001).
66. Pfaffl, M.W., Horgan, G.W. & Dempfle, L. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30, e36 (2002).
67. Jothi, R., Balaji, S., Wuster, A., Grochow, J.A., Gsponer, J., Przytycka, T.M., Aravind, L. & Babu, M.M. Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. *Mol Syst Biol* 5, 294 (2009).
68. Charoensawan, V., Wilson, D. & Teichmann, S.A. Genomic repertoires of DNA-binding transcription factors across the tree of life. *Nucleic Acids Res* 38, 7364-7377 (2010).
69. Perez-Rueda, E., Collado-Vides, J. & Segovia, L. Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea. *Comput Biol Chem* 28, 341-350 (2004).
70. Gama-Castro, S., Salgado, H., Peralta-Gil, M., Santos-Zavaleta, A., Muniz-Rascado, L., Solano-Lira, H., Jimenez-Jacinto, V., Weiss, V., Garcia-Sotelo, J.S., Lopez-Fuentes, A., Porron-Sotelo, L., Alquicira-Hernandez, S., Medina-Rivera, A., Martinez-Flores, I., Alquicira-Hernandez, K., Martinez-Adame, R., Bonavides-Martinez, C., Miranda-Rios, J., Huerta, A.M., Mendoza-Vargas, A., Collado-Torres, L., Taboada, B., Vega-Alvarado,

- L., Olvera, M., Olvera, L., Grande, R., Morett, E. & Collado-Vides, J. RegulonDB version 7.0: transcriptional regulation of Escherichia coli K-12 integrated within genetic sensory response units (Sensor Units). *Nucleic Acids Res* 39, D98-105 (2011).
71. Perez-Rueda, E., Gralla, J.D. & Collado-Vides, J. Genomic position analyses and the transcription machinery. *J Mol Biol* 275, 165-170 (1998).
 72. Lozada-Chavez, I., Janga, S.C. & Collado-Vides, J. Bacterial regulatory networks are extremely flexible in evolution. *Nucleic Acids Res* 34, 3434-3445 (2006).
 73. Li, H., Rhodius, V., Gross, C. & Siggia, E.D. Identification of the binding sites of regulatory proteins in bacterial genomes. *Proc Natl Acad Sci U S A* 99, 11772-11777 (2002).
 74. Yoshida, K.I., Shibayama, T., Aoyama, D. & Fujita, Y. Interaction of a repressor and its binding sites for regulation of the Bacillus subtilis iol divergon. *J Mol Biol* 285, 917-929 (1999).
 75. Hansen, F.G., Christensen, B.B., Nielsen, C.B. & Atlung, T. Insights into the quality of DnaA boxes and their cooperativity. *J Mol Biol* 355, 85-95 (2006).
 76. Stella, S., Cascio, D. & Johnson, R.C. The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. *Genes Dev* 24, 814-826 (2010).
 77. Fuangthong, M. & Helmann, J.D. Recognition of DNA by three ferric uptake regulator (Fur) homologs in Bacillus subtilis. *J Bacteriol* 185, 6348-6357 (2003).
 78. Lloyd, G.S., Godfrey, R.E. & Busby, S.J. Targets for the MalI repressor at the divergent Escherichia coli K-12 malX-malI promoters. *FEMS Microbiol Lett* 305, 28-34 (2010).
 79. Semsey, S., Tolstorukov, M.Y., Virnik, K., Zhurkin, V.B. & Adhya, S. DNA trajectory in the Gal repressosome. *Genes Dev* 18, 1898-1907 (2004).
 80. Schell, M.A. Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol* 47, 597-626 (1993).
 81. Balazsi, G. & Oltvai, Z.N. Sensing your surroundings: how transcription-regulatory networks of the cell discern environmental signals. *Sci STKE* 2005, pe20 (2005).
 82. Luscombe, N.M., Babu, M.M., Yu, H., Snyder, M., Teichmann, S.A. & Gerstein, M. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* 431, 308-312 (2004).
 83. Albert, R. Scale-free networks in cell biology. *J Cell Sci* 118, 4947-4957 (2005).
 84. Brinkrolf, K., Schroder, J., Puhler, A. & Tauch, A. The transcriptional regulatory repertoire of Corynebacterium glutamicum: reconstruction of the network controlling pathways involved in lysine and glutamate production. *J Biotechnol* 149, 173-182 (2010).
 85. Babu, M.M., Luscombe, N.M., Aravind, L., Gerstein, M. & Teichmann, S.A. Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* 14, 283-291 (2004).
 86. Schroder, J. & Tauch, A. Transcriptional regulation of gene expression in Corynebacterium glutamicum: the role of global, master and local regulators in the modular and hierarchical gene regulatory network. *FEMS Microbiol Rev* 34, 685-737 (2010).
 87. Biswas, T., Small, J., Vandal, O., Odaira, T., Deng, H., Ehrt, S. & Tsodikov, O.V. Structural insight into serine protease Rv3671c that Protects M. tuberculosis from oxidative and acidic stress. *Structure* 18, 1353-1363 (2010).
 88. Butala, M., Zgur-Bertok, D. & Busby, S.J. The bacterial LexA transcriptional repressor. *Cell Mol Life Sci* 66, 82-93 (2009).

89. Benaroudj, N., Lee, D.H. & Goldberg, A.L. Trehalose accumulation during cellular stress protects cells and cellular proteins from damage by oxygen radicals. *J Biol Chem* 276, 24261-24267 (2001).
90. Thieffry, D., Huerta, A.M., Perez-Rueda, E. & Collado-Vides, J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* 20, 433-440 (1998).
91. Ruan, J., Perez, J., Hernandez, B., Lei, C., Sunter, G. & Sponsel, V.M. Systematic identification of functional modules and cis-regulatory elements in *Arabidopsis thaliana*. *BMC Bioinformatics* 12 Suppl 12, S2 (2011).
92. Marbach, D., Roy, S., Ay, F., Meyer, P.E., Candeias, R., Kahveci, T., Bristow, C.A. & Kellis, M. Predictive regulatory models in *Drosophila melanogaster* by integrative inference of transcriptional networks. *Genome Res* 22, 1334-1349 (2012).
93. Rodionov, D.A. Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. *Chem Rev* 107, 3467-3497 (2007).
94. Gormley, M., Akella, V.U., Quong, J.N. & Quong, A.A. An integrated framework to model cellular phenotype as a component of biochemical networks. *Adv Bioinformatics* 2011, 608295 (2011).
95. Rodionov, D.A., Novichkov, P.S., Stavrovskaya, E.D., Rodionova, I.A., Li, X., Kazanov, M.D., Ravcheev, D.A., Gerasimova, A.V., Kazakov, A.E., Kovaleva, G.Y., Permina, E.A., Laikova, O.N., Overbeek, R., Romine, M.F., Fredrickson, J.K., Arkin, A.P., Dubchak, I., Osterman, A.L. & Gelfand, M.S. Comparative genomic reconstruction of transcriptional networks controlling central metabolism in the *Shewanella* genus. *BMC Genomics* 12 Suppl 1, S3 (2011).
96. Richet, E., Vidal-Ingigliardi, D. & Raibaud, O. A new mechanism for coactivation of transcription initiation: repositioning of an activator triggered by the binding of a second activator. *Cell* 66, 1185-1195 (1991).
97. Cho, B.K., Zengler, K., Qiu, Y., Park, Y.S., Knight, E.M., Barrett, C.L., Gao, Y. & Palsson, B.O. The transcription unit architecture of the *Escherichia coli* genome. *Nat Biotechnol* 27, 1043-1049 (2009).
98. Semsey, S., Krishna, S., Sneppen, K. & Adhya, S. Signal integration in the galactose network of *Escherichia coli*. *Mol Microbiol* 65, 465-476 (2007).
99. Sawers, G., Kaiser, M., Sirko, A. & Freundlich, M. Transcriptional activation by FNR and CRP: reciprocity of binding-site recognition. *Mol Microbiol* 23, 835-845 (1997).
100. Lu, C.D., Yang, Z. & Li, W. Transcriptome analysis of the ArgR regulon in *Pseudomonas aeruginosa*. *J Bacteriol* 186, 3855-3861 (2004).
101. Lee, J.W. & Helmann, J.D. Functional specialization within the Fur family of metalloregulators. *Biomaterials* 20, 485-499 (2007).
102. Maddox, S.M., Coburn, P.S., Shankar, N. & Conway, T. Transcriptional regulator PerA influences biofilm-associated, platelet binding, and metabolic gene expression in *Enterococcus faecalis*. *PLoS One* 7, e34398 (2012).
103. Michaux, C., Sanguinetti, M., Reffuveille, F., Auffray, Y., Posteraro, B., Gilmore, M.S., Hartke, A. & Giard, J.C. SlyA is a transcriptional regulator involved in the virulence of *Enterococcus faecalis*. *Infect Immun* 79, 2638-2645 (2011).
104. Riboulet-Bisson, E., Le Jeune, A., Benachour, A., Auffray, Y., Hartke, A. & Giard, J.C. Ers a Crp/Fnr-like transcriptional regulator of *Enterococcus faecalis*. *Int J Food Microbiol* 131, 71-74 (2009).

105. Riboulet-Bisson, E., Sanguinetti, M., Budin-Verneuil, A., Auffray, Y., Hartke, A. & Giard, J.C. Characterization of the Ers regulon of *Enterococcus faecalis*. *Infect Immun* 76, 3064-3074 (2008).
106. Yu, C. & Genco, C.A. Fur-mediated activation of gene transcription in the human pathogen *Neisseria gonorrhoeae*. *J Bacteriol* 194, 1730-1742 (2012).
107. Verneuil, N., Rince, A., Sanguinetti, M., Posteraro, B., Fadda, G., Auffray, Y., Hartke, A. & Giard, J.C. Contribution of a PerR-like regulator to the oxidative-stress response and virulence of *Enterococcus faecalis*. *Microbiology* 151, 3997-4004 (2005).
108. Ma, H.W., Kumar, B., Ditges, U., Gunzer, F., Buer, J. & Zeng, A.P. An extended transcriptional regulatory network of *Escherichia coli* and analysis of its hierarchical structure and network motifs. *Nucleic Acids Res* 32, 6643-6649 (2004).
109. Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., Jennings, E.G., Zeitlinger, J., Pokholok, D.K., Kellis, M., Rolfe, P.A., Takusagawa, K.T., Lander, E.S., Gifford, D.K., Fraenkel, E. & Young, R.A. Transcriptional regulatory code of a eukaryotic genome. *Nature* 431, 99-104 (2004).
110. Ozaki, S. & Katayama, T. DnaA structure, function, and dynamics in the initiation at the chromosomal origin. *Plasmid* 62, 71-82 (2009).
111. Armendariz, A.D., Gonzalez, M., Loguinov, A.V. & Vulpe, C.D. Gene expression profiling in chronic copper overload reveals upregulation of Prnp and App. *Physiol Genomics* 20, 45-54 (2004).
112. Kershaw, C.J., Brown, N.L., Constantinidou, C., Patel, M.D. & Hobman, J.L. The expression profile of *Escherichia coli* K-12 in response to minimal, optimal and excess copper concentrations. *Microbiology* 151, 1187-1198 (2005).
113. Jin, Y.H., Dunlap, P.E., McBride, S.J., Al-Refai, H., Bushel, P.R. & Freedman, J.H. Global transcriptome and deletome profiles of yeast exposed to transition metals. *PLoS Genet* 4, e1000053 (2008).
114. Agranoff, D.D. & Krishna, S. Metal ion homeostasis and intracellular parasitism. *Mol Microbiol* 28, 403-412 (1998).
115. Cho, B.K., Federowicz, S.A., Embree, M., Park, Y.S., Kim, D. & Palsson, B.O. The PurR regulon in *Escherichia coli* K-12 MG1655. *Nucleic Acids Res* 39, 6456-6464 (2011).
116. White, B., Tarun, M.C., Gathergood, N., Rusling, J.F. & Smyth, M.R. Oxidised guanidinohydantoin (Ghox) and spiroiminodihydantoin (Sp) are major products of iron- and copper-mediated 8-oxo-7,8-dihydroguanine and 8-oxo-7,8-dihydro-2'-deoxyguanosine oxidation. *Mol Biosyst* 1, 373-381 (2005).
117. Abrantes, M.C., Lopes Mde, F. & Kok, J. Impact of manganese, copper and zinc ions on the transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583. *PLoS One* 6, e26519 (2011).
118. Lopez, G., Latorre, M., Reyes-Jara, A., Cambiazo, V. & Gonzalez, M. Transcriptomic response of *Enterococcus faecalis* to iron excess. *Biometals* 25, 737-747 (2012).
119. Cantini, F., Banci, L. & Solioz, M. The copper-responsive repressor CopR of *Lactococcus lactis* is a 'winged helix' protein. *Biochem J* 417, 493-499 (2009).
120. Strausak, D. & Solioz, M. CopY is a copper-inducible repressor of the *Enterococcus hirae* copper ATPases. *J Biol Chem* 272, 8932-8936 (1997).
121. Magnani, D., Barre, O., Gerber, S.D. & Solioz, M. Characterization of the CopR regulon of *Lactococcus lactis* IL1403. *J Bacteriol* 190, 536-545 (2008).

122. Dian, C., Vitale, S., Leonard, G.A., Bahlawane, C., Fauquant, C., Leduc, D., Muller, C., de Reuse, H., Michaud-Soret, I. & Terradot, L. The structure of the *Helicobacter pylori* ferric uptake regulator Fur reveals three functional metal binding sites. *Mol Microbiol* 79, 1260-1275 (2011).
123. Gao, H., Zhou, D., Li, Y., Guo, Z., Han, Y., Song, Y., Zhai, J., Du, Z., Wang, X., Lu, J. & Yang, R. The iron-responsive Fur regulon in *Yersinia pestis*. *J Bacteriol* 190, 3063-3075 (2008).
124. Miethke, M. & Marahiel, M.A. Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* 71, 413-451 (2007).
125. da Rocha, R.P., Paquola, A.C., Marques Mdo, V., Menck, C.F. & Galhardo, R.S. Characterization of the SOS regulon of *Caulobacter crescentus*. *J Bacteriol* 190, 1209-1218 (2008).
126. Balleza, E., Lopez-Bojorquez, L.N., Martinez-Antonio, A., Resendis-Antonio, O., Lozada-Chavez, I., Balderas-Martinez, Y.I., Encarnacion, S. & Collado-Vides, J. Regulation by transcription factors in bacteria: beyond description. *FEMS Microbiol Rev* 33, 133-151 (2009).
127. Camas, F.M., Blazquez, J. & Poyatos, J.F. Autogenous and nonautogenous control of response in a genetic network. *Proc Natl Acad Sci USA* 103, 12718-12723 (2006).
128. Magnani, D. & Solioz, M. Copper chaperone cycling and degradation in the regulation of the cop operon of *Enterococcus hirae*. *Biometals* 18, 407-412 (2005).
129. Macneil, L.T. & Walhout, A.J. Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. *Genome Res* (2011).
130. Berger, M.F. & Bulyk, M.L. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat Protoc* 4, 393-411 (2009).
131. Gowrishankar, J. & Harinarayanan, R. Why is transcription coupled to translation in bacteria? *Mol Microbiol* 54, 598-603 (2004).
132. Braun, V. & Hantke, K. Recent insights into iron import by bacteria. *Curr Opin Chem Biol* 15, 328-334 (2011).
133. Adikesavan, A.K., Katsonis, P., Marciano, D.C., Lua, R., Herman, C. & Lichtarge, O. Separation of recombination and SOS response in *Escherichia coli* RecA suggests LexA interaction sites. *PLoS Genet* 7, e1002244 (2011).
134. Lebedev, V.S., Veselovskii, A.V., Deinega, E. & Fedorov Iu, I. [The role of reactive oxygen species in copper-induced permeability of plasma membranes in *Escherichia coli*]. *Biofizika* 47, 295-299 (2002).
135. Liolios, K., Chen, I.M., Mavromatis, K., Tavernarakis, N., Hugenholtz, P., Markowitz, V.M. & Kyrpides, N.C. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 38, D346-354 (2010).
136. Shioya, K., Michaux, C., Kuenne, C., Hain, T., Verneuil, N., Budin-Verneuil, A., Hartsch, T., Hartke, A. & Giard, J.C. Genome-wide identification of small RNAs in the opportunistic pathogen *Enterococcus faecalis* V583. *PLoS One* 6, e23948 (2011).

ANEXOS

ANEXO 1. Tabla de partidores en la construcción de mutantes y ensayos de qPCR. Todos los partidores utilizados en la tesis tienen una temperatura de alineamiento de 60 °C. En rojo se indica la secuencia reversa complementaria utilizada para unir ambos fragmentos de amplificación. La secuencia subrayada demarca los sitios de corte para las enzimas NotI (GCGGCCGC) y Sall (GTCGAC).

Código	Secuencia (5' → 3')
mutantes	
fur_up_s	CGCGCGGCCGC <u>TTGTGGATGTTGGCTTACCA</u>
fur_up_a	<u>ACCATGAAA</u> TACAGCTCCTTGTTTAGAAATGATTACA
fur_down_s	<u>AGGAGCTGTA</u> TTTTTCATGGTATTTGTCAAAGTTGT
fur_down_a	CGCG <u>TGCAC</u> ACTAGCCCAACAATGTTGGC
dfur_s	CAGCGATGATGCGTTAATGA
dfur_a	ATAATCGGCACGACAGAAGG
dlexA_s	CGATAACATCAAAATCCCTCCT
dlexA_a	GGAAGCTCAAAAGGCTATTC
dcopY_s	CGACCTTAGTTGCACCGTTC
dcopY_a	CTGGGCTAATCGCCTTGTAG
qPCR	
EF0102_s	CCGATTA AACGCCAAACG
EF0103_s	GTCGCACAAGCAACGATTTTC
EF1979_s	TGGAGCCACTACCGAAAATC
EF2195-EF2198_s	CCTCGGGTTTATGATTGCAG
EF2756_s	ATCGTAAAGTCCGCGGTATC
EF3027_s	CTTCCGGCAACAATGGTAAC
EF3052_s	CCGCCCAAAAATCCTTAGTG
EF0191-EF0193_s	CTGGCTTAGCAGATGTTTTTCG
EF0475-EF0477_s	CGAATTGCGCTGACAGATAC
EF1543_s	TGATGATCATGGGATTGTCC
EF0717-EF0719_s	GCTTGGTGGTACTTTGAAGCTC
EF1556-EF1559_s	GCAATCGATGCAAACAGAAG
EF2379-EF2380_s	AGCGTAACGGCAAAGAAGAG
EF2385-EF2387_s	AAAACAGCGTTGAGCGTAGG
EF2623-EF2624_s	AAACCGGGACATCAACAGTC
EF2927-EF2928_s	AGTAGAAAATGCGGCGGTAG
EF1655_s	GCAAAGTGAACAAGCAGCAG
EF2857_s	AAAGTGGCGCATTACGTAGC
EF3082-EF3085_s	TTAGGTTTGACGCTGGTTGC
EF3086_s	CGATGAACCTGTGATTGACG
EF0066-EF0067_s	AACGAAGTTCCCAGGTGTTG

Continuación tabla ANEXOS 1.

EF0955-EF0957_s	AGATGACGAAGCCCAACAAG
EF1579_s	TCACGGACATTTAGCTCGTC
EF3166-EF3172_s	GTGGAACAGCCGCTTTTATC
EF0188_s	TTCTTACTCGCAGCTTGTGG
EF2740_s	GAAGGACAAGCAGCTGAGATTC
EF0634_s	AATCGTTCGTGGCTTAGCAG
EF0102_a	TTCAGCTCACGAATGTCACG
EF0103_a	CTCGCCAGACTCATCTAATGC
EF1979_a	GCTTCTTGAATCGCTTGCTG
EF2195-EF2198_a	CGTATCCGTAATTGGCACAC
EF2756_a	AACCCCTGCTGCTAATTGTC
EF3027_a	TGCGCTTTATCTACGACGTG
EF3052_a	AGACAATCCAGCCAATCACC
EF0191-EF0193_a	CCAGCTCCATTTCCATCAAC
EF0475-EF0477_a	CCAAGGAGACCCAAACTTTC
EF1543_a	TTAACACATTCTCGCGTTTCG
EF0717-EF0719_a	GGCGTCGTTAAACCAAACCTC
EF1556-EF1559_a	GACCTCTTCCGTCAAACGAG
EF2379-EF2380_a	TATCCTGAATACCGCCCATC
EF2385-EF2387_a	TGCATTGAATTGAGGGTCAC
EF2623-EF2624_a	AAAGAGCCGTCACCATTAC
EF2927-EF2928_a	AAGACGTTCCATTGAGCTG
EF1655_a	TTCCATTACACAGGCCTTC
EF2857_a	CACGGTCATTTTCGTGCATAG
EF3082-EF3085_a	CGCAAATCCCAGTTAACCAC
EF3086_a	GCTGTTGGCTCTTCTTTTGG
EF0066-EF0067_a	GCCAATGCTTGATTGCTTG
EF0955-EF0957_a	ATCCCAATATGTGGCACCTC
EF1579_a	CTTTTCCAAGCCTTCTGGTG
EF3166-EF3172_a	ACTAAGGCATCGGCAATCTC
EF0188_a	CGGGGACGGTTACTTTATTG
EF2740_a	CATTA AAAACGCCCAAC
EF0634_a	ATTGAGTCAACGGCACCTTC

Genome wide identification of *Acidithiobacillus ferrooxidans* (ATCC 23270) transcription factors and comparative analysis of ArsR and MerR metal regulators

Christian Hödar · Pablo Moreno · Alex di Genova ·
Mauricio Latorre · Angélica Reyes-Jara · Alejandro Maass ·
Mauricio González · Verónica Cambiazo

Received: 24 February 2011 / Accepted: 21 July 2011 / Published online: 10 August 2011
© Springer Science+Business Media, LLC. 2011

Abstract *Acidithiobacillus ferrooxidans* is a chemolithoautotrophic acidophilic bacterium that obtains its energy from the oxidation of ferrous iron, elemental sulfur, or reduced sulfur minerals. This capability makes it of great industrial importance due to its applications in biomining. During the industrial processes, *A. ferrooxidans* survives to stressing circumstances in its environment, such as an extremely acidic pH and high concentration of transition metals. In order to gain insight into the organization of *A. ferrooxidans*

regulatory networks and to provide a framework for further studies in bacterial growth under extreme conditions, we applied a genome-wide annotation procedure to identify 87 *A. ferrooxidans* transcription factors. We classified them into 19 families that were conserved among diverse prokaryotic phyla. Our annotation procedure revealed that *A. ferrooxidans* genome contains several members of the ArsR and MerR families, which are involved in metal resistance and detoxification. Analysis of their sequences revealed known and potentially new mechanism to coordinate gene-expression in response to metal availability. *A. ferrooxidans* inhabit some of the most metal-rich environments known, thus transcription factors identified here seem to be good candidates for functional studies in order to determine their physiological roles and to place them into *A. ferrooxidans* transcriptional regulatory networks.

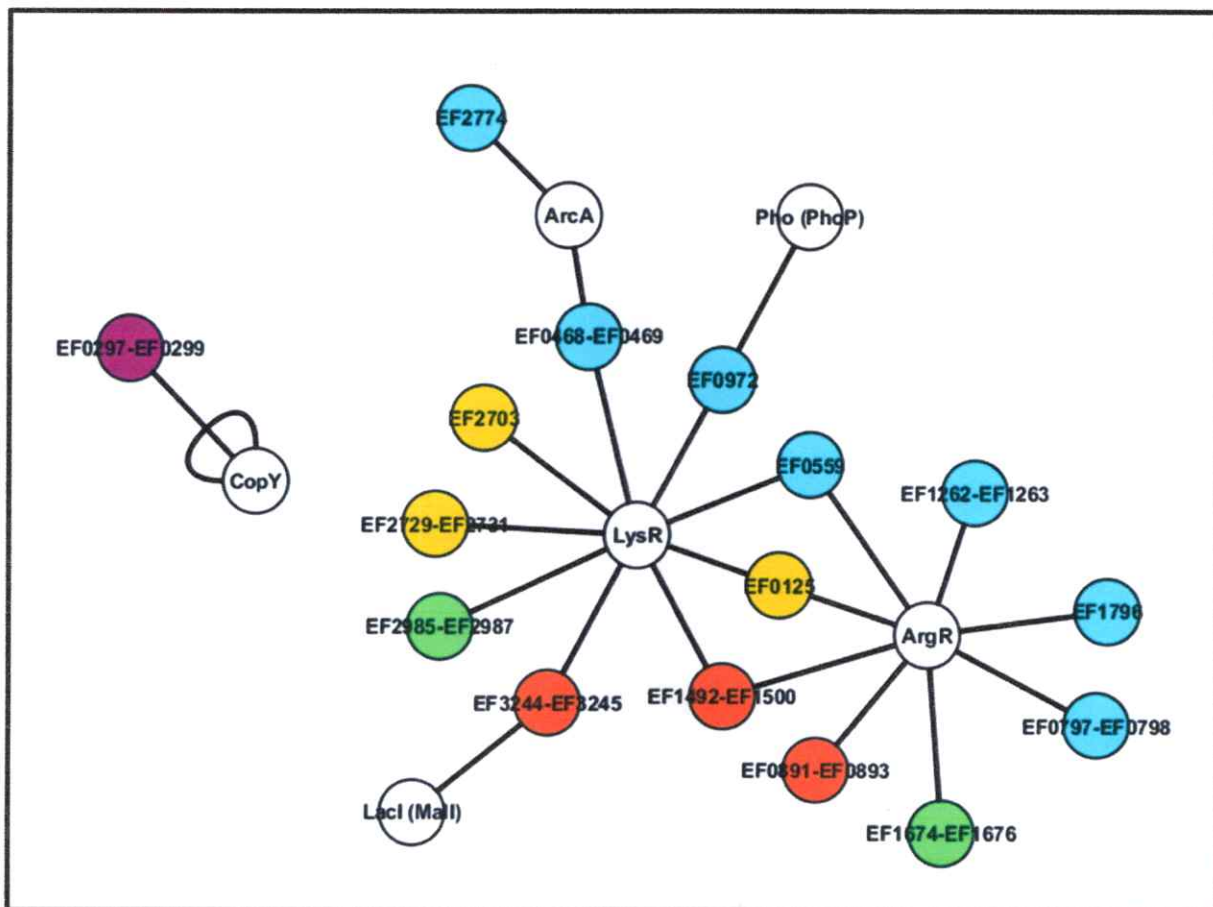
C. Hödar and P. Moreno contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10534-011-9484-8) contains supplementary material, which is available to authorized users.

C. Hödar · M. Latorre · A. Reyes-Jara ·
M. González · V. Cambiazo (✉)
Laboratorio de Bioinformática y Expresión Génica,
INTA, Universidad de Chile, El Líbano 5524,
Santiago, Chile
e-mail: vcambiaz@inta.cl

Keywords *Acidithiobacillus ferrooxidans* ·
Transcriptional regulators · Metal resistance ·
ArsR family · MerR family

ANEXO 3. Sub-red de regulación transcripcional activada por 0,05 mM de CuSO₄ (Red de Abrantes). El grafo contiene 22 nodos (6 reguladores o familias de factores de transcripción y 16 blancos de regulación u operones) conectados a través de 22 enlaces (sitios de unión). Simbología de colores (agrupados por posibles funciones) para nodos (esferas): i) blancos de regulación (542), amarillo transcripción, traducción y replicación; morado transporte y señalización; naranja metabolismo basal y generación de energía; verde otras funciones; celeste sin función asignada o proteínas hipotéticas; ii) elementos reguladores en blanco (13 familias de factores de transcripción). Los enlaces se indican con flechas negras.



ANEXO 4. Publicación.

Biometals (2012) 25:737–747
DOI 10.1007/s10534-012-9539-5

Transcriptomic response of *Enterococcus faecalis* to iron excess

Guadalupe López · Mauricio Latorre ·
Angélica Reyes-Jara · Verónica Cambiazo ·
Mauricio González

Received: 30 December 2011 / Accepted: 7 March 2012 / Published online: 24 March 2012
© Springer Science+Business Media, LLC. 2012.

Abstract Iron is an essential nutrient for sustaining bacterial growth; however, little is known about the molecular mechanisms that govern gene expression during the homeostatic response to iron availability. In this study we analyzed the global transcriptional response of *Enterococcus faecalis* to a non-toxic iron excess in order to identify the set of genes that respond

to an increment of intracellular iron. Our results showed an up-regulation of transcriptional regulators of the Fur family (PerR and ZurR), the cation efflux family (CzcD) and ferredoxin, while proton-dependent Mn/Fe (MntH) transporters and the universal stress protein (UspA) were down-regulated. This indicated that *E. faecalis* was able to activate a transcriptional response while growing in the presence of an excess of non-toxic iron, assuring the maintenance of iron homeostasis. Gene expression analysis of *E. faecalis* treated with H₂O₂ indicated that a fraction of the transcriptional changes induced by iron appears to be mediated by oxidative stress. A

Guadalupe López, Mauricio Latorre and Angélica Reyes-Jara contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s10534-012-9539-5) contains supplementary material, which is available to authorized users.

G. López · M. Latorre · A. Reyes-Jara ·
V. Cambiazo · M. González (✉)
Laboratorio de Bioinformática y Expresión Génica,
INTA, Universidad de Chile, El Líbano 5524,
Macul, Santiago, Chile
e-mail: mgonzale@inta.uchile.cl

M. Latorre
e-mail: mlatorre@inta.uchile.cl

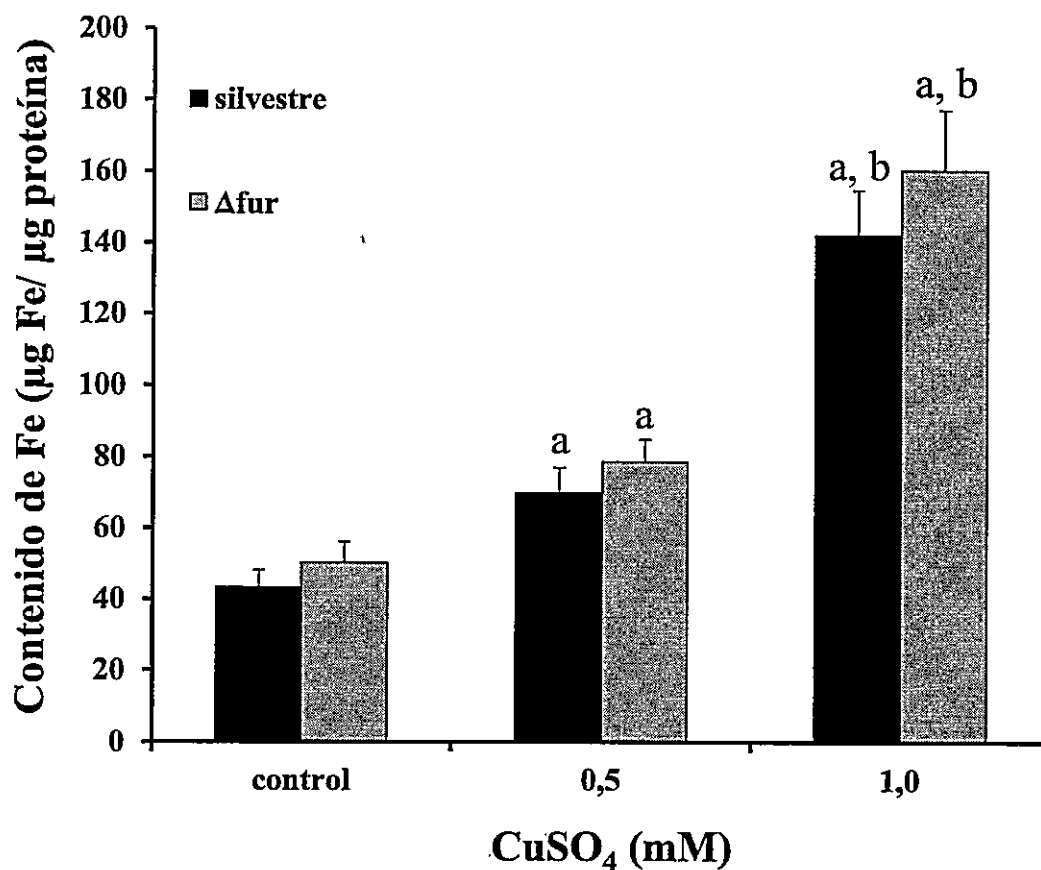
G. López
Laboratorio de Nutrición Molecular, Universidad
Autónoma del Estado de Hidalgo, Sahagún,
Pachuca, Mexico
e-mail: glopez@inta.uchile.cl

A. Reyes-Jara
Laboratorio de Microbiología y Probióticos, INTA,
Universidad de Chile, El Líbano 5524, Macul,
Santiago, Chile
e-mail: areyes@inta.uchile.cl

V. Cambiazo · M. González
Fondap-Center of Genome Regulation, Universidad de
Chile, Santiago, Chile
e-mail: vcambiaz@inta.uchile

M. González
Laboratorio de Bioinformática y Matemática del Genoma,
Centro de Modelamiento Matemático (UMI 2807, CNRS)
Facultad de Ciencias Físicas y Matemáticas, Universidad
de Chile, 8370459 Santiago, Chile

ANEXO 5. Contenido intracelular de Fe en cepas silvestre y mutantes *fur* de *E. faecalis* expuestas a 0,5 mM de CuSO_4 . El gráfico muestra la acumulación de Fe en el tiempo para la cepa silvestre de *E. faecalis* OG1RF (barras negras) y mutante *fur* (barras plomo claras). El valor calculado declara el promedio tres réplicas biológicas con tres réplicas técnicas de medición de absorbancia. La letra a y b corresponde a diferencias significativas con respecto a la situación control y 0,5 mM de CuSO_4 respectivamente. ANOVA test $p < 0,05$.



ANEXO 6. Resumen congreso internacional “Tema 14-Trace Elements in Man and Animals”. Realizado entre los días 19 - 24 de septiembre 2011 en la ciudad de Enshi, China.

Transcriptional regulatory network activated by copper and iron in *Enterococcus faecalis*.

Latorre M.¹, Maass A.², González M.¹.

mlatorre@inta.cl

¹LBEG-INTA, Universidad de Chile. ²LBMG-CMM, Universidad de Chile.

Introduction

Copper (Cu) and Iron (Fe) are micronutrient required from prokaryotes to eukaryotes organism, nevertheless excess of both metals could produce cellular toxicity by generation of reactive oxygen species (ROS). Global transcript abundance profiling, shows different groups of genes able to respond commonly and specifically when cells were exposed to Cu and Fe. In this context, this work aims to identify potential transcriptional regulatory networks (TRN), capable to coordinate the transcriptional response of *E. faecalis* against these metals.

Materials and Methods

Microarrays: Total *E. faecalis*V583 genome (Nimblegen N°A4359-00-01) was hybridized independently (2 biological replicates) with RNA extracted from cells grown at mid-log phase under three conditions, control (without metal addition), Cu (0.5mM CuSO₄) and Fe (0.5mM FeCl₃-NTA). Data analysis DNASTARArrayStarv2.0.

TRN: Consensus/Patser-MotifSampler/Scan algorithms were used to search upstream *E. faecalis* operons (-300/+50bp), bacterial binding site (BS) previously assigned to transcription factor families (TF). Microarray data and TNR was integrated by Cytoscape software.

Results

A total of 159 (81up/78down) and 480 (269up/211down) operons change his expression profiles under Cu and Fe exposed respectively, being 43 (17up/26down) shared in both condition. The 50% of the total of these metal response operons present at least one putative BS. These elements establish the TRN activated, composed by 20TF/360operons (nodes) and 374BSs (vertex). In general, the most of TF connect the common and specifically response, highlighting as the most represented LysR/ArgR/FNR/Fur/LexA families, however Abr/FruR/Fis connect only Fe respond operons.

Discussion

The TRN activated described a putative connectivity between different Cu and Fe response operons, suggesting common TF able to sense both metals direct or indirectly, probably linked to ROS effects. On the other hand, the TRN allow us to group, operons which codified for unknown proteins (n=100) with operons classified into different metabolic process, projecting to these method like a tool for operon classifications.

Financing

This work was financed by Fondecyt N°1030618 and FONDAP N°15090007 projects. M. Latorre is a recipient of a Conicyt doctoral fellowship.

ANEXO 7. Resumen congreso internacional “Fourth International Workshop on Iron and Copper Homeostasis”. Realizado entre los días 9 – 13 de noviembre 2011 en la ciudad de Pucon, Chile.

Conserved transcriptional regulatory mechanism involved in copper and iron homeostasis in *Lactobacillales* order species.

Mauricio Latorre¹, Jessica Galloway-Peña^{2,3}, Jung Rho², Barbara E. Murray^{2,3}, Alejandro Maass^{4,5}, Mauricio Gonzalez^{1,4}

1. Laboratorio de Bioinformática y Expresión Génica, INTA-Universidad de Chile, Santiago, Chile., 2. Department of Internal Medicine, Division of Infectious Diseases, Center for the Study of Emerging and Reemerging Pathogens, Houston, USA., 3. Department of Microbiology and Molecular Genetics, University of Texas Medical School at Houston, Houston, USA., 4. Center of Genome Regulation, Santiago, Chile., 5. Laboratorio de Bioinformática y Matemática del Genoma, CMM-Universidad de Chile, Santiago, Chile.

Key words: Transcriptional regulatory mechanism, copper, iron, homeostasis, *Lactobacillales* order.

Copper and iron are important micronutrients involved in different cellular processes. However, an excess of these metals can generate free radicals toxic to the cell. Therefore, homeostasis requires the presence of several proteins implicated in controlling the concentration of these trace elements. In previous work, we characterized proteins involved in copper and iron homeostasis and their corresponding transcriptional regulatory mechanisms in the bacterial model *Enterococcus faecalis*. This work aims to study the conservation of these mechanisms in the *Lactobacillales* order. Blastp was used to search for the *E. faecalis* copper (3 genes, 1 operon, regulated by CopY) and iron (11 genes, 4 operons, regulated by Fur) homeostasis protein homologs in 11 species of the *Lactobacillales* order. We identified putative homologous of these components in all the genomes analyzed. We then generated a CopY and Fur 3D molecular model for each species to identify functional motifs previously characterized in DNA binding and metal coordination (SWISS-MODEL) and we searched for its corresponding CopY and Fur nucleotide binding site (in the promoter regions for each operon (MOTIF-SEARCH)). Molecular modeling showed that CopY and Fur homologs have conserved protein folding and functional motifs in all 11 species. In addition, the corresponding CopY and Fur nucleotide binding sites are also conserved, highlighting a strict palindrome sequence inside the first 100bp upstream of the translation start site. These data show for first time two metal transcriptional regulons extremely conserved in a bacterial order, suggesting the remarkable control of the expression of copper and iron homeostasis genes.

Financing: Fondecyt 1110427; FONDAF 15090007; MECESUP UCH0713, Vicerrectoría de Asuntos Académicos Universidad de Chile and Conicyt fellowship (ML).