



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA QUÍMICA,
BIOTECNOLOGÍA Y MATERIALES

**IDENTIFICACIÓN DE ARNS NO CODIFICANTES DE RESPUESTA A COBRE
Y ANTIBIÓTICOS EN *ENTEROCOCCUS FAECALIS***

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN QUÍMICA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL EN BIOTECNOLOGÍA

JAVIERA ANDREA PINO GAETE

PROFESOR GUÍA:
ÁLVARO OLIVERA NAPPA

PROFESOR CO-GUÍA:
MAURICIO LATORRE MORA

MIEMBROS DE LA COMISIÓN:
IRENE MARTÍNEZ BASTERRECHEA
ORIANA SALAZAR AGUIRRE

Este trabajo ha sido financiado por:
Fondecyt Regular N°1190742, año 2021 y Proyecto Anillo ACT210004, año 2021

Este trabajo ha sido realizado en cotutela con la Universidad de O'Higgins

SANTIAGO DE CHILE

2023

RESUMEN DE LA TESIS PARA OPTAR AL
GRADO DE MAGÍSTER EN CIENCIAS DE
LA INGENIERÍA, MENCIÓN QUÍMICA Y
MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERA CIVIL EN BIOTECNOLOGÍA
POR: JAVIERA ANDREA PINO GAETE
FECHA: ENERO 2023
PROF. GUÍA: ÁLVARO OLIVERA NAPPA

IDENTIFICACIÓN DE ARNS NO CODIFICANTES DE RESPUESTA A COBRE Y ANTIBIÓTICOS EN *ENTEROCOCCUS FAECALIS*

Enterococcus faecalis es una bacteria Gram positiva de gran interés debido a su capacidad intrínseca y adquirida de resistencia a antibióticos y antimicrobianos. Este microorganismo es responsable de brotes de infecciones nosocomiales en centros hospitalarios, produciendo complicaciones en pacientes difíciles de tratar. El control de esta bacteria es un desafío debido a las estrategias utilizadas para su supervivencia, pues se ha demostrado que posee mecanismos específicos que facilitan su resistencia a antibióticos y metales pesados con propiedades antimicrobianas como el cobre. En términos generales, se han descrito en bacterias diversos mecanismos involucrados en la resistencia a antimicrobianos, tales como factores de transcripción, proteínas, operones, ribointerruptores y recientemente, ARN no codificantes o ARNnc, moléculas capaces de regular a diferentes niveles, ya sea la expresión de genes o modificando interacciones moleculares de componentes de resistencia. En este trabajo, el objetivo principal fue identificar ARNnc diferencialmente expresados bajo la exposición a Cu y antibióticos en *E. faecalis*. Para esto, primero mediante un análisis *in silico* de datos masivos de RNA-seq, se identificó, clasificó y seleccionó un conjunto de 586 ARNnc candidatos. En base a diferentes criterios de conservación y función regulatoria, dos de ellos (nc3340 y nc3683) mostraron estar presentes en los tres datos de RNA-seq estudiados, con una abundancia transcripcional mayor con respecto a un estado basal de crecimiento en al menos dos condiciones, y cuyas secuencias poseían blancos de regulación codificantes para proteínas hipotéticas. En base a este análisis, ambos finalmente fueron seleccionados para su validación por RT-qPCR. Interesantemente, los ARNnc respondieron transcripcionalmente en *E. faecalis* frente a la adición de vancomicina, ceftriaxona y cobre, evidenciando cambios de abundancia a nivel transcripcional de ARNnc frente a la exposición de estos estresores. Estos ARNnc podrían a futuro ser blancos de estudio para el diseño de nuevos fármacos y estrategias biotecnológicas para el control de la bacteria patógena *E. faecalis*.

IDENTIFICATION OF NON-CODING RNAs IN RESPONSE TO COPPER AND ANTIBIOTICS IN *ENTEROCOCCUS FAECALIS*

Enterococcus faecalis is a Gram-positive bacterium of great interest due to its intrinsic and acquired capacity for resistance to antibiotics and antimicrobials. This microorganism is responsible for outbreaks of nosocomial infections in hospitals, causing complications in patients that had difficult treatments. Controlling this bacterium is a challenge due to the strategies used for its survival, it has specific mechanisms that facilitate its resistance to antibiotics and heavy metals with antimicrobial properties, such as copper. In general terms, various mechanisms involved in antimicrobial resistance have been described in bacteria, such as transcription factors, proteins, operons, riboswitches and recently, non-coding RNA or ncRNA, molecules able of regulating at different levels, either the expression of genes or modifying molecular interactions of resistance components. In this work, the main aim was to identify differentially expressed ncRNAs under Cu and antibiotics exposure in *E. faecalis*. Consequently, first by an *in silico* analysis of massive RNA-seq data, a set of 586 candidate ncRNAs were identified, classified and selected. Based on conservation criteria and regulatory function, two of them (nc3340 and nc3683) showed to be present in the three RNA-seq data studied, with a major transcriptional abundance compared with a basal state of growth in at least two conditions, and whose sequences had regulatory targets encoding hypothetical proteins. Based on this analysis, both were finally selected for validation by RT-qPCR. Interestingly, the ncRNAs had a transcriptional response in *E. faecalis* to the addition of vancomycin, ceftriaxone, and copper, evidencing changes in abundance at the transcriptional level of ncRNAs when exposed to these stressors. In the future, these ncRNAs could be studied as targets for the design of new drugs and biotechnological strategies for the control of the pathogenic bacterium *E. faecalis*.

A mi familia y a la ciencia

Agradecimientos

Quisiera agradecer a mis papás por estar incondicionalmente, apoyarme en todas mis aventuras, formarme, enseñarme de la vida y hacer que sea quien soy hoy. A mi hermano por regalarme tantas alegrías, cariño y risas (y rabias también). A mis abuelos que me cuidan desde lejos, por darme amor infinito y consentirme hasta el último de sus días. A mis abuelas por enseñarme de la vida, darme amor, regalónearme y malcriarme menos. A mis tíos y primos por hacer que mi familia sea siempre tan unida y especial, su apoyo y amor ha sido muy importante para mí, siempre. A mis amigos de la U, Benja P., Benja S., Cris, Claudio, Dani, Dieguito, Franco, Talo y Vane, gracias por hacer mi paso por el departamento tan lindo y por seguir compartiendo juntos. A mis amigos de taller, Diego, Pali, Seba y Pame, por haber confiado en mi trabajo, soportarme bajo estrés y haber formado una bella amistad. A mis amigos de IQBM, Gabito, por siempre darme tanto cariño, cafecito, una buena conversación y estar siempre; Martín, por compartir tantas risas, abrazos, conversaciones ñoñas y no tan ñoñas; Jose, por hacer mis primeros años en el departamento tan lindos y lleno de buenos momentos; Mati, por apañar hasta las tantas en el laboratorio, estudiando o hablando de la vida. A mis amigos de plan común, Gabo, Isa, Javi, Tere, Marya y Seba por el apoyo en momentos difíciles y las aventuras compartidas. A mis compañeras de colegio, Carito, Claire, Anto, Carola y Vale, por estar siempre y darme ánimo, por tantas risas, anécdotas y por compartir hasta el día de hoy. A quienes conocí en la UOH, Profe Mauricio por aportar en mi formación y por el apoyo constante como profesional, a Feña por enseñarme, ayudarme, acompañarme durante la tesis y por la amistad construida, a Víctor por toda la ayuda computina en mi tesis y su gran disposición, a Gabo, Jimmy, Seba y Jorge por las risas y orientación cuando lo necesité. A Fede, por apoyarme en todo este proceso y por el cariño compartido. A mi mejor amiga Romi, por las carcajadas de todos los días, las aventuras, las lloraditas, el apoyo de siempre y el HDM. Al Profe Álvaro por compartir sus consejos profesionales y personales cuando lo necesité, por orientarme y estar siempre presente. Al Profe Garrido por enseñarme la biotecnología que tanto me encanta.

Tabla de Contenido

1. Antecedentes generales	1
1.1. Estado del arte	1
1.1.1. Infecciones nosocomiales	1
1.1.2. <i>Enterococcus faecalis</i>	2
1.1.3. Mecanismos de resistencia a antibióticos en <i>E. faecalis</i>	2
1.1.4. Respuesta transcripcional global de <i>E. faecalis</i> frente a antibióticos . .	6
1.1.5. Mecanismos de resistencia y respuesta transcripcional global a cobre en <i>E. faecalis</i>	8
1.1.6. ARNs no codificantes	11
1.1.7. ARNs no codificantes en <i>E. faecalis</i>	13
2. Motivación	15
3. Objetivos	16
3.1. Objetivo general	16
3.2. Objetivos específicos	16
4. Materiales y Metodología	17
4.1. Materiales	17
4.1.1. Reactivos	17
4.1.2. Instrumentos y elementos	18
4.1.3. Equipos	19
4.1.4. Herramientas Computacionales	20
4.1.5. Microorganismo	20
4.1.6. Composición del Medio de Cultivo	21
4.2. Metodología	21
4.2.1. Etapa 1: Recopilación bibliográfica y uso de herramientas bioinformáticas	21
4.2.1.1. Búsqueda bibliográfica	21
4.2.1.2. Descarga y análisis de datos de RNA-seq	22
4.2.1.3. Comparación de la expresión de transcritos	23
4.2.1.4. Clasificación y selección de los ARNnc candidatos	23
4.2.1.5. Predicción de estructuras de los ARNnc, conservación y codificación de genes blanco	24

4.2.2.	Etapa 2: Validación <i>in vitro</i> y análisis de resultados	24
4.2.2.1.	Crecimiento bacteriano bajo condiciones control	24
4.2.2.2.	Determinación de concentraciones de agentes estresores	24
4.2.2.3.	Extracción de ARN	25
4.2.2.4.	PCR en tiempo real (qPCR)	26
4.2.2.5.	Procesamiento de resultados de qPCR	27
4.2.2.6.	Análisis estadístico	28
5.	Resultados y Discusión	29
5.1.	Identificación <i>in silico</i> de ARNnc	29
5.1.1.	Análisis de datos de RNA-seq	29
5.1.2.	Identificación y comparación de la expresión diferencial de ARNs no codificantes de respuesta a Cu y antibióticos	30
5.1.3.	Selección final de ARNnc	33
5.1.4.	Posición en el genoma de los ARNnc seleccionados con sus blancos de regulación	35
5.1.5.	Análisis de la estructura y conservación de los ARNnc seleccionados	37
5.2.	Validación de abundancia de los ARNnc	41
5.2.1.	Determinación de las condiciones de tratamiento	41
5.2.2.	Extracción de ARN total y diseño de partidores	44
5.2.3.	Cuantificación de los ARNnc en cultivos de <i>E. faecalis</i> expuestos a Cu, vancomicina y ceftriaxona	46
6.	Conclusiones	49
6.1.	Proyecciones	50
	Bibliografía	53
	Anexos	65
A.	Análisis <i>in silico</i>	65
A.1.	Programas utilizados	65
A.2.	Código	65
A.2.1.	Expresión diferencial	77
A.3.	Conservación de los ARNnc	78
B.	Validación de abundancia de los ARNnc	88
B.1.	Diseño de partidores o <i>primers</i>	88

Índice de Tablas

4.1.	Reactivos y kits usados durante el trabajo en laboratorio.	17
4.2.	Instrumentos utilizados en el trabajo de laboratorio.	18
4.3.	Equipos utilizados durante el trabajo de tesis.	19
4.4.	Herramientas computaciones utilizadas durante el trabajo de tesis.	20
4.5.	Composición del Medio N líquido.	21
5.1.	Expresión diferencial de los ARNnc en exposición a Cu, vancomicina y ceftriaxona.	32
5.2.	Función de proteínas codificadas por genes blanco de los ARNnc identificados en Cu, vancomicina y ceftriaxona.	34
5.3.	Conservación de los transcritos según BLASTN.	39
5.4.	Concentración y calidad de los ARN purificados.	45
5.5.	Partidores diseñados para la síntesis del ADNc de hebra simple a partir de cada ARNnc.	45
5.6.	Partidores diseñados para las reacciones de amplificación y cuantificación mediante qPCR.	46
B.1.	Secuencia de los ARN no codificantes en estudio.	89

Índice de Ilustraciones

1.1.	Acción de los antibióticos sobre la estructura de la red de peptidoglicano. . . .	4
1.2.	Mecanismo de acción del operón <i>vanHAX</i> para adquirir resistencia a la vancomicina.	5
1.3.	Respuesta transcripcional global de <i>E. faecalis</i> tratadas con vancomicina. . . .	7
1.4.	Modelo del operón <i>cop</i> para la homeostasis de Cu en <i>E. faecalis</i>	9
1.5.	Respuesta transcripcional diferencial global de <i>E. faecalis</i> tratadas con Cu. . .	10
4.1.	Placa de 96 pocillos con diluciones en serie.	25
5.1.	Diagrama de Venn para ARNs no codificantes identificados en los RNA-seq de Cu, vancomicina y ceftriaxona.	30
5.2.	Contexto genómico de los ARNnc seleccionados.	37
5.3.	Predicción de la estructura secundaria de los ARNs no codificantes seleccionados	38
5.4.	Predicción de la estructura secundaria de los ARNnc según su conservación en distintas especies.	40
5.5.	Curvas de crecimiento para cultivos de <i>E. faecalis</i> OG1RF expuestos a Cu, vancomicina y ceftriaxona.	43
5.6.	Curvas de crecimiento para cultivos de <i>E. faecalis</i> OG1RF de 15 mL expuestos a estresores.	44
5.7.	Abundancia transcripcional de los transcritos de ARNnc en estudio, obtenida por qPCR.	46
A.1.	Alineamiento de ARN nc3340 en <i>Enterococcus</i> sp. PLM3.	79
A.2.	Alineamiento de ARN nc3340 en <i>Enterococcus</i> sp. LX10.	80
A.3.	Alineamiento de ARN nc3340 en <i>Lactococcus garvieae</i> JJJN1.	81
A.4.	Alineamiento de ARN nc3340 en <i>Lactococcus petauri</i> LG4.	82
A.5.	Alineamiento de ARN nc3340 en <i>Lactococcus garvieae</i> JJJN1.	83
A.6.	Alineamiento de ARN nc3683 en <i>Enterococcus</i> sp. LX10.	84
A.7.	Alineamiento de ARN nc3683 en <i>Enterococcus</i> sp. PLM3.	85
A.8.	Conservación de la secuencia del ARN nc3340 en varias especies.	86
A.9.	Conservación de la secuencia del ARN nc3683 en varias especies.	87

Capítulo 1

Antecedentes generales

1.1. Estado del arte

1.1.1. Infecciones nosocomiales

Las infecciones intrahospitalarias (IIH), también conocidas como infecciones nosocomiales o infecciones asociadas a la atención en salud (IAAS), son procesos infecciosos generales o localizados en ciertas regiones anatómicas, adquiridos durante la permanencia o concurrencia de un paciente al hospital o centro de salud [1]. Se clasifican dentro de esta categoría cualquier infección que se presente durante la hospitalización de un enfermo y que no se encontraba vigente o en incubación en el momento del ingreso. Dentro de las IAAS más frecuentes se encuentran infecciones del tracto digestivo, del tracto respiratorio, infecciones asociadas a catéter vascular, e infecciones urinarias [1].

En 1982, en Chile se creó un programa nacional de prevención de estas infecciones, obligatorio tanto para clínicas y hospitales públicos o privados del país, dirigido por el Ministerio de Salud [2]. El programa se ha ido actualizando durante el paso de los años; consiste principalmente en normas y regulaciones para detectar las infecciones, supervisarlas y prevenirlas, mediante vigilancia epidemiológica, capacitaciones del personal de salud, racionalización del uso de antimicrobianos, desinfección y esterilización de espacios e instrumentos, aislamiento de pacientes, coordinación entre centros de salud, entre otras [3][4].

A nivel mundial, a pesar de que las medidas implementadas han demostrado ser exitosas con la disminución y detención de la propagación de IAAS en centros de salud [5], se ha sumado otro obstáculo en el control de las enfermedades nosocomiales relacionado con la resistencia a antimicrobianos de algunas cepas de bacterias, principalmente *Enterococos* resistentes al antibiótico vancomicina (o VRE por sus siglas en inglés) [6]. Esto ha desencadenado la necesidad de la búsqueda de nuevos antibióticos eficientes contra este tipo de bacterias, ya que dicha resistencia impide erradicar las infecciones mediante terapia antibiótica en los pacientes. Chile no escapa a este problema, pues según reportes del Instituto de Salud Pública (ISP) existe un incremento gradual de aislados de enterococos resistentes a

vancomicina [7], dando cuenta de la importancia de abordar este tema de forma preventiva buscando alternativas al uso de este antibiótico.

1.1.2. *Enterococcus faecalis*

Dentro del género *Enterococcus* se encuentran alrededor de 49 especies descritas, donde 26 de ellas han sido asociadas a infecciones en humanos [8]. Dos de estas especies son comúnmente encontradas en el intestino de algunos mamíferos, incluyendo a los humanos: *Enterococcus faecalis* y *Enterococcus faecium*. Estas bacterias forman parte de la lista de microorganismos responsables de las infecciones nosocomiales, pues cobran importancia gracias a su capacidad intrínseca de resistencia a antibióticos, incluyendo gentamicina, estreptomicina y vancomicina. Se ha reportado que la resistencia a este último antibiótico en los años 90 era menor al 1%, mientras que para el 2018 era cercana al 30% [9]. Dada la disponibilidad del genoma de la cepa V583 de *E. faecalis* resistente a vancomicina, se ha demostrado que más de un cuarto de su genoma es ADN foráneo o móvil, lo que evidencia la habilidad de adquirir genes que le permiten adaptarse a condiciones de estrés [10]. Sumado a que la tolerancia antibiótica facilita y precede a la resistencia [11], esta capacidad hace que *E. faecalis* se convierta en un problema para la búsqueda de nuevos antimicrobianos eficientes para tratar una infección con este patógeno [12].

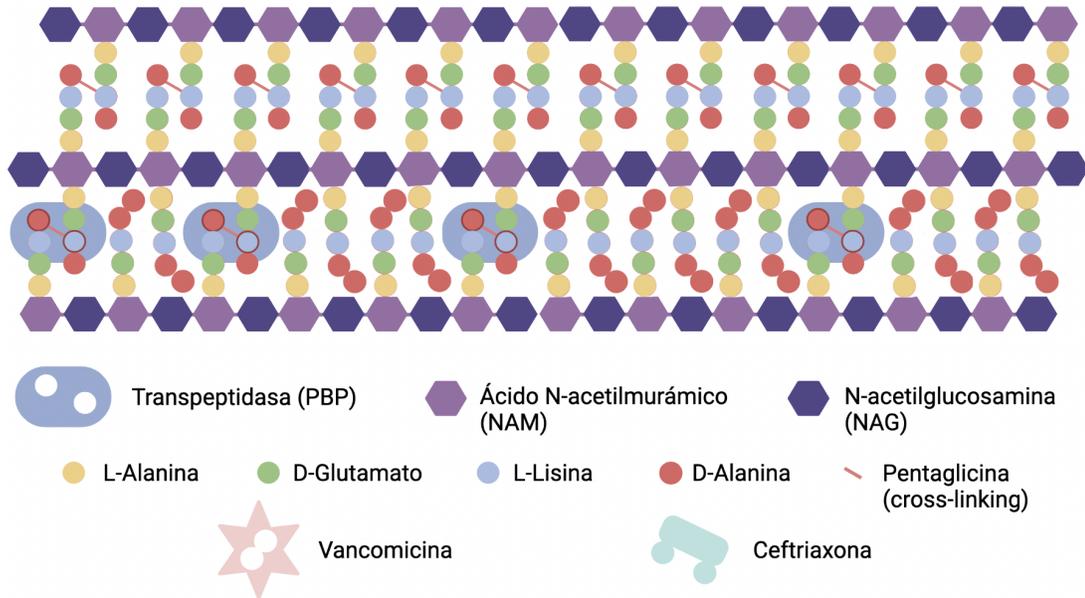
En particular, se ha estudiado en profundidad algunos mecanismos de acción de *E. faecalis*, bacteria Gram positiva anaerobia facultativa conocida por ser causa frecuente de infecciones intrahospitalarias que causan enfermedades o complicaciones como endocarditis, septicemia, infecciones al sistema nervioso central o neumonía. A pesar de que carece de factores de virulencia clásicos como toxinas proinflamatorias o moduladores inmunitarios, su principal característica radica en la capacidad de resistir en entornos de estrés, especialmente a antibióticos del tipo betalactámicos, glucopéptidos y aminoglucósidos [13]. Esta propiedad, sumado a una resistencia intrínseca comandada por la capacidad de adquirir material genético foráneo (tales como genes de resistencia), le permiten una mejor adaptación frente a la acción de diversos antibióticos, transformándose en un foco de estudio a nivel mundial en el desarrollo de nuevas estrategias antimicrobianas [14][15][16][17].

1.1.3. Mecanismos de resistencia a antibióticos en *E. faecalis*

Es posible identificar a nivel de resistencia, mecanismos específicos que presenta *E. faecalis* para degradar o impedir el efecto de los antibióticos. Para los glucopéptidos la resistencia está asociada a mutaciones en la ruta de síntesis de peptidoglicano, la que se observa en la Figura 1.1.A, específicamente en la mutación de sus precursores [18]. Un ejemplo clave es el mecanismo de resistencia a la vancomicina, cuyo blanco son los precursores pentapeptídicos del peptidoglicano terminados en D-Ala-D-Ala, como los que se muestran en la Figura 1.1.B. Al unirse la vancomicina a dichos aminoácidos, la transpeptidasa encargada de crear los enlaces entre cadenas de peptidoglicano y hacer el *cross-linking* entre ellas no logra unirse

a los residuos y por lo tanto no se estabiliza la pared de peptidoglicano, haciendo que la bacteria no posea una barrera celular funcional. Por lo tanto, como mecanismo de resistencia ante el antibiótico en cuestión, suelen mutarse los pentapéptidos, que pueden ser del tipo D-Ala-D-Lac o D-Ala-D-Ser final, reduciendo la afinidad con la vancomicina dependiendo del tipo de residuo mutado [19].

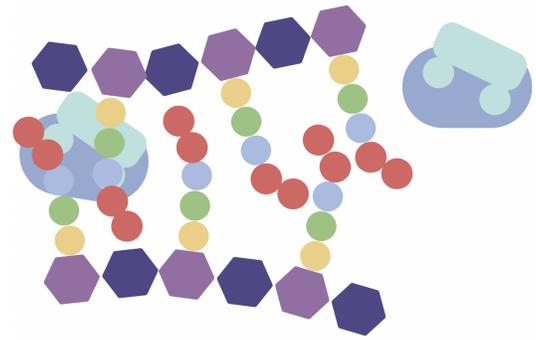
Se han descrito varios operones relacionados con la resistencia a glucopéptidos, donde el más común es el llamado operón *vanHAX* que se presenta en la Figura 1.2. Posee 7 genes transcritos por 2 promotores, donde 6 de ellos tienen función conocida para glucopéptidos: 3 responsables directamente de la resistencia a glucopéptidos, 2 responsables de la regulación de la resistencia y 1 responsable de eliminar los precursores no modificados de la pared celular [20]. El mecanismo de acción consta de varias etapas: primero, la regulación se lleva a cabo por un sistema de dos componentes (TCS por sus siglas en inglés), donde el primer operón transcribe los genes *vanS* y *vanR*, sensor de quinasa (HK) y regulador de respuesta (RR), respectivamente. Estos factores inducen la transcripción de los siguientes genes río abajo en el operón, controlados por un segundo promotor [21]. Cada uno de esos genes tiene un rol distinto en la modificación del pentapéptido precursor del peptidoglicano, como se muestra en la Figura 1.2. En definitiva, la acción de este sistema provoca que la vancomicina no pueda unirse al precursor de peptidoglicano y por lo tanto la síntesis de la pared celular de la bacteria no se ve interrumpida [18].



(A)



(B)



(C)

Figura 1.1: Acción de los antibióticos sobre la estructura de la red de peptidoglicano. La figura muestra principalmente un modelo de la creación de los enlaces entre cadenas de NAM-NAG (*cross-linking*) en la pared celular de una bacteria, comandado por la acción de transpeptidasas. (A) La transpeptidasa logra enlazar cadenas y generar un *cross-linking* exitoso, por lo que la pared de peptidoglicano es estable. (B) La acción de la vancomicina impide que la enzima genere enlaces entre las cadenas, pues el antibiótico se une a la terminación D-Ala-D-Ala del precursor de peptidoglicano. (C) Las moléculas de ceftriaxona se unen a las transpeptidasas disponibles, evitando que la enzima pueda unirse a NAM-NAG y formar los enlaces que le otorgan estabilidad a la pared. Creación en base a [19][22].

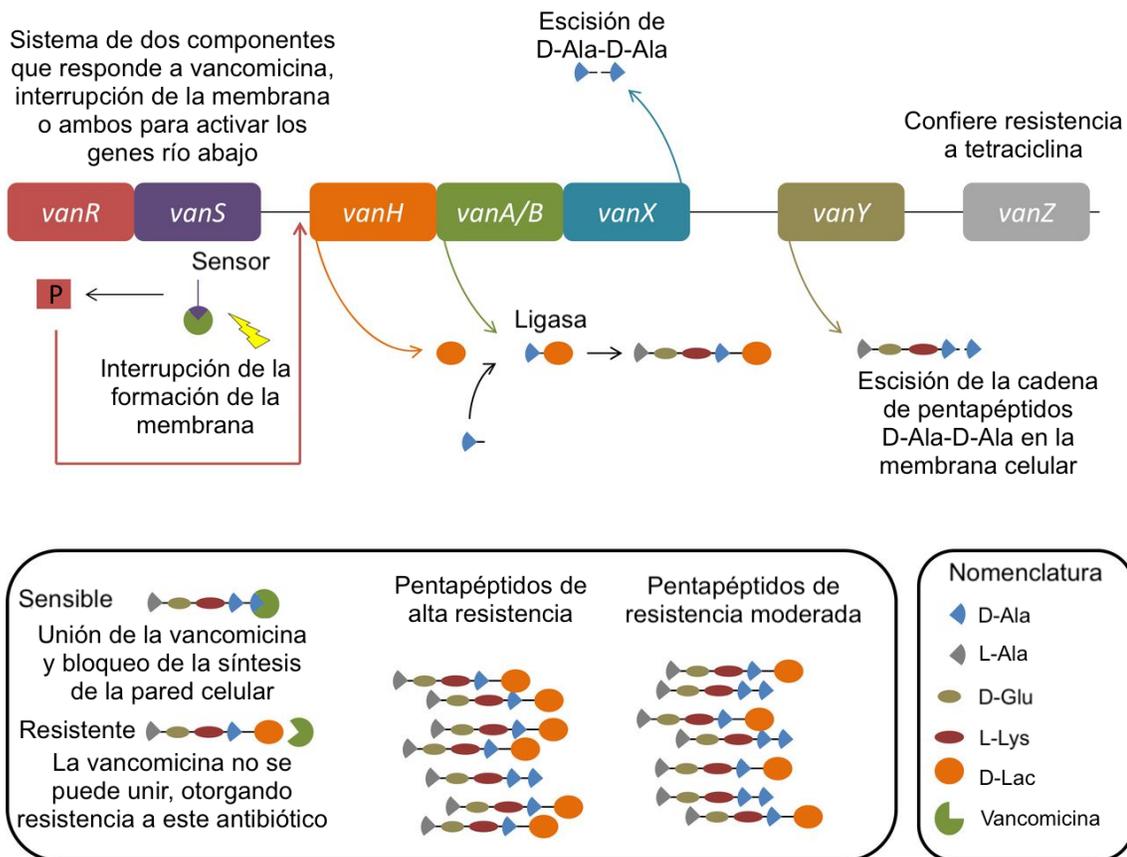


Figura 1.2: Mecanismo de acción del operón *vanHAX* para adquirir resistencia a la vancomicina. La figura muestra la función de cada gen del operón *vanHAX* activado por la presencia de vancomicina y las modificaciones hechas en los péptidos para adquirir resistencia al antibiótico. Figura adaptada [20].

La ceftriaxona es un antibiótico del tipo cefalosporina de tercera generación, cuyo mecanismo de acción se basa en la unión de moléculas del antibiótico a la transpeptidasa, específicamente a los sitios de unión de los péptidos de las cadenas de NAM-NAG (ácido N-acetilmurámico y N-acetilglucosamina, respectivamente), evitando así que se produzca el *cross-linking* entre las cadenas de peptidoglicano y desestabilizando la pared celular [22]. El detalle de la distribución de moléculas en la formación del peptidoglicano, se puede observar en la Figura 1.1.C. Para las cefalosporinas en general se ha encontrado que existen sistemas de resistencia a este tipo de antibióticos, como 1) la presencia de proteínas de unión a penicilina (*penicilin binding proteins* o PBPs por sus siglas en inglés) que poseen menos afinidad a las cefalosporinas y por lo tanto no pierde su función y se sintetiza la pared celular, 2) un sistema de regulación de dos componentes llamado CroRS que regularía la expresión de genes que otorgan resistencia, esencial para la tolerancia y resistencia en *E. faecalis* [23][24], pues las mutantes $\Delta croRS$ son susceptibles a antibióticos como la ceftriaxona, 3) un gen codificante de quinasa Ser/Thr predicha (*ireK*), que cuya delección reduce la resistencia a las cefalosporinas [25]. Para el caso de los antibióticos betalactámicos, la resistencia suele estar asociada principalmente a la producción de β -lactamasa por parte de la bacteria, mutaciones

en los aminoácidos que se encuentran en el sitio activo de las proteínas de unión a penicilina o bien a la represión de su producción, pues éstas son el principal blanco del antimicrobiano [26][27], dando cuenta de una respuesta altamente específica hacia el antibiótico.

Estos antecedentes dan cuenta de la existencia de mecanismos específicos que presenta *E. faecalis* para degradar o impedir la acción de los diversos antibióticos, no obstante, la respuesta transcripcional frente al antimicrobiano puede incluir otros genes involucrados, dando cuenta de la complejidad del fenómeno de resistencia.

1.1.4. Respuesta transcripcional global de *E.faecalis* frente a antibióticos

Se han realizado importantes esfuerzos para identificar la respuesta transcripcional global de *E. faecalis* frente a distintos tipos de crecimiento y la exposición a diversos antibióticos. En el trabajo publicado el año 2017 por el grupo de Seneviratne y colaboradores [28], se comparó la respuesta transcripcional y capacidad de resistencia de *E. faecalis* creciendo de forma libre y formando biopelículas con y sin exposición a una concentración no letal de vancomicina. Para ello, tomaron diferentes aislados de *E. faecalis* en pacientes con tratamientos conductos fallidos y se estudiaron las células crecidas bajo las condiciones de estudio. El primer análisis realizado mostró que la formación de biopelículas permite un incremento en la resistencia frente a vancomicina en relación con el crecimiento en vida libre, no obstante, si bien algunos aislados formaron biopelículas más gruesas que otras, el tamaño de ésta no guarda relación con la resistencia al antibiótico. Posteriormente se extrajo ARN en cada condición y se realizó un RNA-seq para determinar cambios transcripcionales. Se detectó que 163 genes fueron diferencialmente expresados al comparar el crecimiento en suspensión versus el estado de biopelículas, mientras que 101 fueron regulados diferencialmente para biopelículas tratadas con vancomicina versus biopelículas sin tratamiento. Estos genes fueron clasificados por categorías según la función de las proteínas codificadas, como se muestra en la Figura 1.3, indicando que diferentes procesos metabólicos son activados o reprimidos producto de la adición del antibiótico, dando cuenta de una respuesta transcripcional global que involucra mecanismos de regulación transcripcional altamente coordinados. Particularmente, en bacterias tratadas con vancomicina existe un aumento de transcritos que codifican para diferentes clases de transportadores tipo ABC de eflujo de fármacos y proteínas de unión a penicilina, involucradas en la síntesis de la pared celular y en la resistencia a antibióticos betalactámicos. También se destacan cambios en la abundancia transcripcional de genes de resistencia a metales, tales como el arsénico, correspondientes a reguladores transcripcionales del operón *ars* que contienen genes de respuesta al estresor.

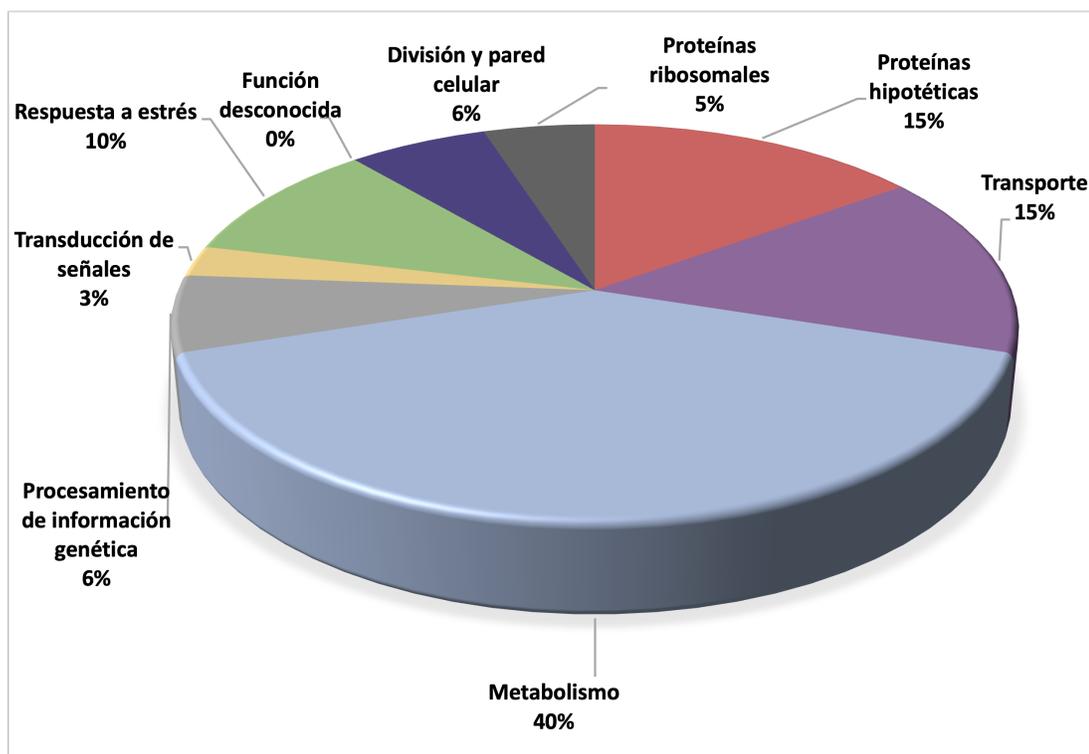


Figura 1.3: Respuesta transcripcional global de *E. faecalis* tratadas con vancomicina. El gráfico representa la clasificación de los datos provenientes de un ensayo de RNA-seq. Los datos se agruparon en base a los procesos celulares y distribución de abundancia de transcritos diferencialmente expresados frente a la exposición a vancomicina con respecto a la muestra control sin exposición. Figura adaptada [28].

En esta línea, otras investigaciones han analizado la respuesta transcripcional de *E. faecalis* en respuesta a otros antibióticos [23]. Se ha demostrado que genes involucrados a biogénesis de moléculas basales, división de la pared celular y transporte destacan dentro de los más expresados frente a un tratamiento con linezolid. En particular, transcritos que codifican para proteínas de biogénesis y división de la pared celular aumentan sus niveles en respuesta a los antimicrobianos del tipo teixobactina, respuesta directamente relacionada con la acción de ambos estresores los cuales afectan la formación de la pared celular.

Es así como queda demostrado que la resistencia a antibacterianos es compleja, pues el mecanismo de supervivencia involucra múltiples cambios en la expresión génica y no únicamente los directamente relacionados con genes de resistencia al estresor, reafirmando la idea de la presencia de diversos mecanismos regulatorios a nivel transcripcional capaces de controlar la expresión de estos genes. En este contexto, diversos trabajos han investigado y caracterizado reguladores transcripcionales relacionados a la virulencia y resistencia a estrés en *E. faecalis*, como son los casos mencionados a continuación: el regulador transcripcional SlyA asociado a la virulencia en presencia de sales biliares [29][30], la proteína Ers relacionada con la patogénesis y respuesta al estrés [31], el regulador transcripcional HypR en respuesta al estrés oxidativo y supervivencia intracelular [32], entre otros reguladores transcripcionales

ya descritos [33][34]. Si bien, estos reguladores evidencian la presencia de mecanismos de control de la expresión involucrados con la resistencia a antibióticos, considerando el total de genes que responden frente a la exposición a estos estresores, es de suponer la existencia de otros mecanismos que complementen la regulación a nivel transcripcional.

1.1.5. Mecanismos de resistencia y respuesta transcripcional global a cobre en *E. faecalis*

En términos generales, dado su acción como cofactor enzimático, el Cobre (Cu) es un micronutriente esencial que participa en diversos procesos metabólicos, tales como la respiración aeróbica, respiración anaeróbica y eliminación de especies tóxicas reactivas de oxígeno [35]. Por el contrario, se ha demostrado que el Cu en altas concentraciones es tóxico contra algunos microorganismos como virus y bacterias, incluyendo a *E. faecalis*, siendo específicamente los iones de Cu quienes producen daño en el ADN, membranas celulares, proteínas y otras moléculas a través de la generación de radicales libres debido a la reacción de Fenton y Haber-Weiss [36]. Esta propiedad tóxica ha permitido en el último tiempo la aplicación exitosa del Cu en algunas superficies y materiales, mostrando ser efectivo para disminuir la transmisión de infecciones debido a la contaminación cruzada, principalmente en recintos hospitalarios [37].

En relación con *E. faecalis* y Cu, se ha estudiado el mecanismo de regulación y homeostasis de este metal a nivel intracelular. La homeostasis de Cu en esta bacteria está regulada por los productos del operón *cop*, conformado por 3 genes: *copA*, *copY* y *copZ*, y que codifican para una ATPasa tipo P transportadora de Cu, un regulador transcripcional dependiente de Cu y una proteína chaperona de Cu, respectivamente, como se muestra en la Figura 1.4 [38]. La expresión del operón *cop* es inducida por la cantidad del metal. En ausencia de Cu el regulador transcripcional CopY actúa como represor, uniéndose al promotor del operon *cop* evitando la transcripción de sus genes. Frente a un aumento en el contenido de Cu intracelular, la chaperona CopZ le entrega Cu a CopY. El complejo CopY-Cu presenta una menor afinidad por el ADN conduciendo a la liberación del promotor permitiendo la transcripción de los genes río abajo, incluyendo al gen *copA* que codifica para la proteína de eflujo ATPasa tipo P, la cual transporta el Cu desde el citoplasma hacia el medio extracelular reduciendo la toxicidad del metal en el interior de la bacteria [38].

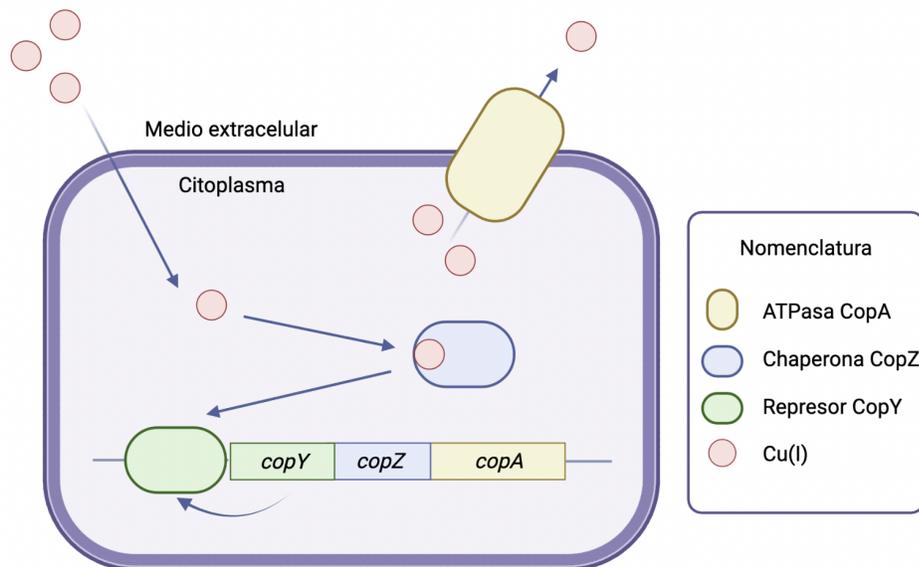


Figura 1.4: Modelo del operón *cop* para la homeostasis de Cu en *E. faecalis*. La figura muestra el ingreso a la célula, transporte y eflujo de Cu. Se indica según los colores verde, azul y amarillo los genes y proteínas respectivas codificadas. Modelo adaptado [39][38].

Del mismo modo como ocurre con la respuesta a antibióticos, el Cu además de inducir la expresión de genes involucrados directamente con su homeostasis, también genera una respuesta transcripcional a escala global en *E. faecalis*. En la investigación publicada en el año 2010 por Reyes-Jara y colaboradores [39], se estudió la respuesta adaptativa de *E. faecalis* a Cu y el papel de regulación de CopY en ella, analizando el transcriptoma a nivel genómico de la bacteria silvestre (*wild-type*) y una cepa mutante para *copY* ($\Delta copY$), ambas expuestas a niveles no tóxicos de exceso de Cu. Los resultados indicaron primero, que la respuesta transcripcional del operón *cop* de homeostasis de Cu está directamente relacionada con la presencia del regulador CopY, siendo al parecer el único sistema regulado por esta proteína, lo cual confiere al sistema un alto nivel de especificidad en la respuesta frente al metal. Por otro lado, la exposición a Cu, al igual que lo que ocurre con los tratamientos de antibióticos, es capaz de activar y reprimir un gran número de genes. Al respecto, un total de 287 genes responden a la exposición a Cu, 145 de ellos fueron regulados de forma positiva (incluyendo a los 3 miembros del operón *cop*) y 142 regulados de forma negativa. Además de los genes del operón *cop*, se encontraron otros genes asociados con diferentes procesos celulares, como el procesamiento de la información genética, transporte, pared celular, proteínas de unión a penicilina, proteínas hipotéticas, entre otros, como se muestra en la Figura 1.5. En resumen, existen varios componentes sumados a los de homeostasis de Cu que pueden complementar la respuesta al metal, lo que también supone la existencia de otros mecanismos regulatorios que operan en paralelo a CopY.

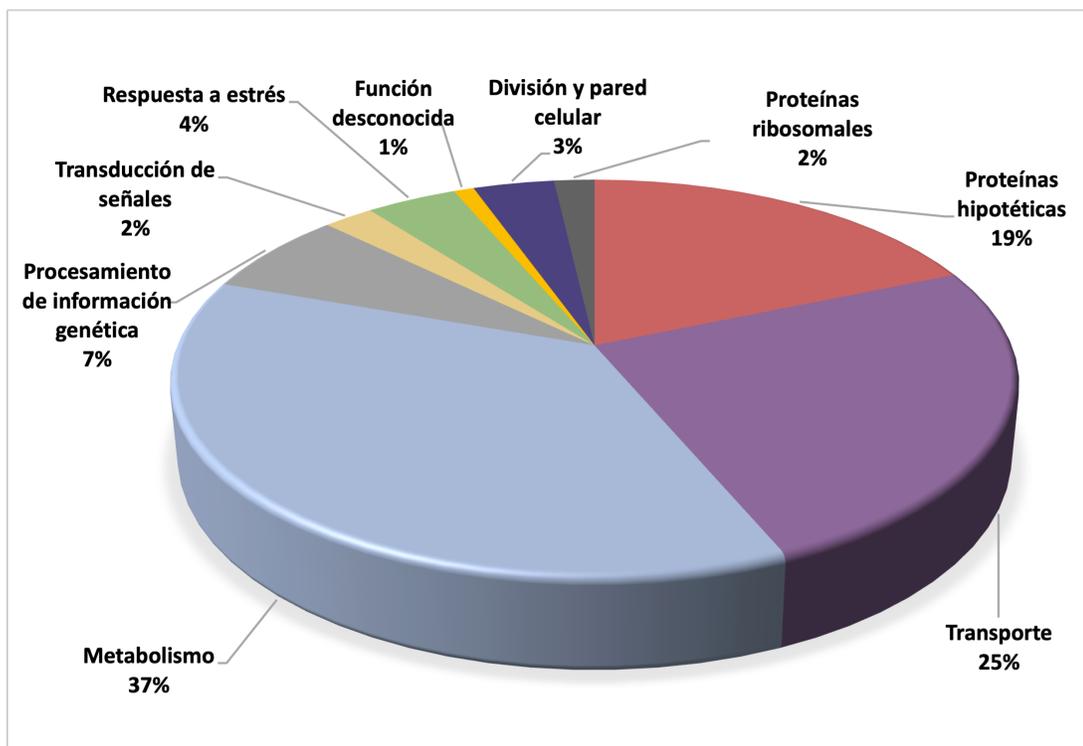


Figura 1.5: Respuesta transcripcional diferencial global de *E. faecalis* tratadas con Cu. El gráfico representa la clasificación de la función de proteínas codificadas por genes diferencialmente expresados en presencia de Cu (0.5 mM de CuSO₄ por 3 horas) con respecto a la muestra control sin exposición. Los datos se agruparon en base a los procesos celulares y distribución de abundancia de transcritos. Figura adaptada [39].

Interesantemente, se ha demostrado que en cepas de *Enterococcus* existe una correlación entre la capacidad de resistencia a antibióticos y la tolerancia a la toxicidad del Cu [40][41]. Un estudio realizado por Hasman y colaboradores en el año 2006, analizó la resistencia a Cu, macrólidos y glucopéptidos de algunas bacterias presentes en cerdos daneses, con el fin de buscar una correlación entre la presencia de genes de resistencia a Cu y antibióticos en dichas bacterias [42]. En un principio, se tomaron cepas de *E. faecium*, caracterizando si en ellas se encontraban genes de resistencia a Cu y antibióticos, como lo son *tcrB* para la homeostasis de Cu (también reportado en *E. faecalis*), *erm(B)* para la resistencia a macrólidos y *vanA* para vancomicina. Se evidenció que antes de la exposición a altos niveles de Cu, menos del 40% de ellas presentaba resistencia el gen de resistencia al metal, mientras que luego de suplementar la dieta de los lechones con altas concentraciones de Cu por 28 días, la presencia del gen de resistencia fue encontrada en el 94% de ellas. Del mismo modo, el gen de resistencia a macrólidos también se encontró en más bacterias, donde todas ellas también poseían resistencia a Cu. Finalmente, se encontró además que el gen de resistencia a vancomicina *vanA* también se encontraba en cepas de las bacterias presentes en los cerdos suplementados con dietas altas en Cu, a diferencia de aquellos que recibían concentraciones suplementarias menores. Este experimento mostró que los altos niveles de Cu administrados a los lechones como suplemento alimenticio no solo seleccionan *E. faecium* resistentes al metal,

sino que también condicionan un aumento de cepas resistentes a macrólidos y glicopéptidos.

En esta línea, un experimento similar al anterior fue llevado a cabo en China en el año 2017 por Zou y colaboradores, donde se mostró que al alimentar a lechones con altas cantidades de Cu por un mes, existe un aumento de especies de enterococos y lactobacilos obtenidas desde las heces de los animales, que presentan una mayor tolerancia al Cu y con ello además un aumento en la resistencia antibiótica para vancomicina, penicilina, enrofloxacina y estreptomicina [43].

Con estos estudios es posible pensar que existe un vínculo entre la presencia de genes de resistencia a Cu y antibióticos en bacterias expuestas únicamente a Cu. A nivel molecular, estudios han revelado que en suelos contaminados a largo plazo con metales pesados y metaloides como cadmio, arsénico, plomo, zinc y Cu, algunos genes de resistencia a metales presentes en bacterias (MRGs por su nombre en inglés), tienen una co-ocurrencia con la presencia de genes de resistencia a antibióticos (ARGs por sus siglas en inglés), revelando que algunos de los MRGs que más coexistían con ARGs eran *copA*, *copB* y *pcoA* [44]. Además, múltiples estudios han demostrado que la contaminación de Cu a largo plazo en diferentes tipos de suelos se correlaciona directamente con un aumento en la prevalencia de ARGs [45], estableciendo así posibles redes y asociaciones específicas entre los MRGs y ARGs [45].

Todos estos antecedentes sugieren la existencia de una relación directa entre estresores del tipo metales (como el Cu) y los antibióticos, tanto a nivel de la respuesta transcripcional global, donde la exposición a Cu o antibióticos activa mecanismos de resistencia para ambos estresores, como también a nivel de adquisición mecanismos de forma conjunta al exponer bacterias a la presencia de Cu. Por tanto, es posible suponer la existencia de posibles mecanismos regulatorios comunes capaces de responder a ambos estímulos. Al respecto, en términos de regulación transcripcional mediada por factores de transcripción, ninguno de los reguladores vinculados a resistencia a antibióticos o el regulador CopY son capaces de regular tanto mecanismos de resistencia a Cu y antibióticos, o incluso cambiar su abundancia a nivel transcripcional frente en ambos estímulos, abriendo un espacio importante en términos de la identificación de otros mecanismos transcripcionales involucrados en el control de la respuesta génica de *E. faecalis* tanto a Cu como antibióticos.

1.1.6. ARNs no codificantes

En estos últimos años dentro de los focos de mayor interés en el control de la expresión génica está el estudio de ARNs no codificantes (ARNnc) en bacterias. Estas moléculas de ARN actúan normalmente como reguladores postranscripcionales, siendo clasificadas como ribointerruptores, ARN mensajero de transferencia (ARNtm), ARN no codificantes largos (lARNnc), ARN no codificantes pequeños (sARN), entre otros [46]. En particular, los sARN corresponden a secuencias de longitud menor a 200 nucleótidos, cuya función está relacionada con el control de la expresión génica mediante la interacción con el ARN mensajero (ARNm).

Entre sus funciones, pueden operar como silenciador en la expresión génica a través de la unión por complementariedad imperfecta del sitio de unión al ribosoma (RBS) del ARNm objetivo, inhibiendo la asociación entre la subunidad ribosómica 30S ya sea con el inicio de la traducción [47], o directamente sobre el ARNm, como es el caso del ARNnc pequeño de interferencia o siARN [48]. También pueden inhibir la traducción por medio de mecanismos de alteración río abajo del RBS o bloqueo de otros mecanismos que permiten la unión del ribosoma [49], estimular la traducción, o prevenir la degradación del ARNm uniéndose por complementariedad de bases a la región 5'-UTR río arriba del RBS [50].

Los ARNnc que tienen como objetivo una secuencia de ARNm, se pueden clasificar en dos tipos: ARNs de acción en *cis* o en *trans*. En los ARNnc de acción en *cis*, el mecanismo de regulación consiste en la unión perfecta por complementariedad de bases al ARNm codificado en el gen al que regulan o gen objetivo, puesto que se transcriben en la hebra complementaria (antisentido), mientras que los de acción en *trans* pueden regular por complementariedad parcial, ya que son sintetizados en un lugar del genoma distinto al del gen objetivo [50]. Algunas de estas moléculas tienen roles importantes en la fisiología bacteriana, pues pueden afectar la homeostasis del hierro [51], metabolismo de carbohidratos [52], respuestas al estrés (como estrés oxidativo) y condiciones inducidas por daño en el ADN [53], composición de la superficie celular [54][55] y patogenicidad [56].

Los lARNnc son ARNs no codificantes de más de 200 nucleótidos, ampliamente estudiados en modelos eucariontes puesto que son capaces de regular la transcripción de genes al modificar histonas y ADN, principalmente mediante metilaciones y acetilaciones [57]. A pesar de que la evidencia no ha sido suficiente para determinar completamente la funcionalidad de todos los lARNnc, se ha demostrado que pueden alterar la estabilidad y traducción de los ARNm de genes cercanos o lejanos (modo de acción similar a los sARN), afectando así la expresión de genes en distintos niveles, ya sea interactuando no solo con el ARNm, sino que también con el ADN o proteínas [57]. En particular, los lARNnc que regulan en *cis* puede ser, por ejemplo, a través de un impedimento estérico que influya en la expresión de genes, y con ello posiblemente generar la pérdida o ganancia de funciones. Se ha descrito que al ADN se pueden unir actuando mayormente en *cis* que en *trans*, formando un bucle R o *R-loops* (estructura de tres hebras compuesta por un híbrido de ADN:ARN y el ADN monocatenario no molde asociado). La síntesis de los lARNnc al parecer es similar a la de los ARNm, dependiendo de la función de cada uno, pueden unirse a proteínas de unión a ARN (RBPs por sus siglas en inglés) y ribosomas [58].

En algunas bacterias se ha identificado y estudiado en profundidad a ARNnc capaces de regular a más de un gen. Tal es el caso de RyhB, un sARN que ha demostrado ser clave en la homeostasis de hierro cuando hay escasez de este micronutriente, pues controla a más de 50 genes relacionados con la utilización y captura de este metal, operando a nivel del control sobre la traducción, desestabilizando el ARNm, o promoviendo la expresión de alguno de

sus blancos [59]. Se ha observado la presencia de RyhB en otras especies, principalmente en enterobacterias, encontrando homólogos que poseen genes blanco conservados como aquellos asociados a la homeostasis de Cu, y que al mismo tiempo poseen roles en virulencia y formación de biopelículas [59], reafirmando la relación que existe entre los metales y patogénesis a nivel de regulación transcripcional.

En términos de la búsqueda de nuevos blancos o estrategias antimicrobianas, interesantemente, los ARNnc tienen potencial de ser usados como objetivos para el desarrollo de biomarcadores de enfermedades y de herramientas estratégicas para su uso terapéutico. Estudios realizados en ribointerruptores o *riboswitches* han mostrado resultados prometedores como blancos para nuevos antimicrobianos, como es el caso de los biointerruptores que modulan negativamente la transcripción tras la unión de guanina, cuya acción se vio reducida al afectar su afinidad con el sustrato en *Clostridioides difficile* [60]. También se han identificado ribointerruptores asociados a la resistencia antibiótica de aminoglucósidos en bacterias patógenas, quienes al unirse a moléculas del fármaco conducen a la inducción de la resistencia al antibiótico en la bacteria [61].

Así como los ribointerruptores mencionados, se han identificado en múltiples microorganismos una variedad de sARN asociados a genes de resistencia a antimicrobianos, bombas de eflujo, formación de biopelículas, modificación de la pared celular, entre otros [62], destacando la relación entre los sARN y los ARNm y potenciando la importancia de estas cortas secuencias en la respuesta antimicrobiana. [63]. Otro ejemplo de control de patógenos es la inyección de ARNnc mediante el uso de fagos, estrategia que ha permitido silenciar o eliminar fenotipos de *Escherichia coli* resistentes a antibióticos, logrando alterar de forma específica la traducción del ARNm que codifica para las proteínas de cloranfenicol acetiltransferasa y kanamicina fosfotransferasa, ambas involucradas en la degradación de antibióticos [64].

1.1.7. ARNs no codificantes en *E. faecalis*

Con respecto a ARNnc en *E. faecalis*, en el año 2014 el grupo de Michaux y colaboradores estudiaron estas secuencias como posibles elementos de respuesta a diferentes condiciones de estrés, como son la presencia ácido láctico, sales biliares, peróxido de hidrógeno, entre otras [65]. En términos operacionales, los ensayos y detección de estos ARNnc fueron realizados mediante microarreglos de mosaico, cuyos resultados fueron posteriormente comparados con los obtenidos por análisis *in silico*, concluyendo que la expresión obtenida de los ARNnc era similar en ambos métodos y que estos son complementarios para identificar este tipo de secuencias. Dentro del trabajo se detectó un total de 100 posibles candidatos a ARNnc, donde 6 de ellos fueron estudiados en profundidad comparando el comportamiento de la bacteria con y sin el ARNnc. Los resultados indicaron que la ausencia de alguno de estos elementos afecta la expresión de proteínas relacionadas con estrés, e incluso puede afectar la virulencia de *E. faecalis*, dando cuenta de la importancia de los ARNnc en estos procesos.

Otro trabajo recientemente publicado por el mismo grupo, mediante tecnología de secuenciación masiva por RNA-seq, describió la presencia de un total de 230 posibles ARNnc diferencialmente expresados bajo condiciones de estrés como las mencionadas anteriormente [66], complementando la información previa sumando un número importante de nuevos candidatos de respuesta. Si bien esta información da cuenta de la presencia de estas moléculas presentes en el genoma de *E. faecalis*, a la fecha no hay registros de la respuesta transcripcional de ARNnc frente a la adicción de antibióticos. Del mismo modo como se mencionó anteriormente, en cuanto a Cu, no se han encontrado registros que muestren ARNnc de respuesta diferencial en *E. faecalis* u otra bacteria, abriendo la posibilidad de identificar nuevos componentes no solo involucrados en el control de las concentraciones de este metal, sino también capaces de responder de igual modo a diferentes antibióticos.

Al respecto, poder contar con un conjunto de ARNnc de respuesta a Cu y antibióticos permite por un lado profundizar en los mecanismos de respuesta de *E. faecalis* frente a estos dos estresores, sumado a la posibilidad de abrir nuevas estrategias para el control de este patógeno.

Capítulo 2

Motivación

En consideración al continuo interés asociado al control de patógenos mediante antimicrobianos, surge la necesidad por investigar nuevos componentes involucrados en la respuesta transcripcional de *E. faecalis* ante diferentes factores de estrés, tales como antibióticos y metales pesados (como el Cu). En particular, identificar secuencias de ARNnc que cambian su abundancia frente a estos estresores, permitirá analizar y entender el control sobre la respuesta transcripcional en esta bacteria ante la presencia de antimicrobianos.

En este contexto, las investigaciones asociadas a ARNnc son acotadas, ya que se presentan limitantes dado lo reciente de su descubrimiento y escasez de técnicas relacionadas a su manipulación [67][68]. Adicionalmente, el mayor número de publicaciones se centran en la identificación y cuantificación de ARNm proveniente del transcriptoma de un microorganismo. Por otro lado, el avance actual y aparición de nuevas y mejores herramientas bioinformáticas para el análisis y tratamiento de datos transcripcionales abren un espacio para la investigación de ARNnc en bacterias.

Por todo esto, la principal motivación de este estudio considera tomar todos estos elementos con el fin de identificar secuencias no codificantes que cambien su respuesta transcripcional en la bacteria *E. faecalis* frente al tratamiento de Cu y antibióticos. Con los resultados obtenidos, se espera entregar un respaldo experimental para posteriores investigaciones, contribuyendo a la búsqueda de nuevos posibles blancos para la creación de fármacos o sistemas biotecnológicos con actividad terapéutica contra este patógeno.

Capítulo 3

Objetivos

3.1. Objetivo general

El objetivo principal de este trabajo es identificar ARNs no codificantes diferencialmente expresados bajo la exposición a cobre y antibióticos en *Enterococcus faecalis*.

3.2. Objetivos específicos

Como objetivos específicos, se plantea lo siguiente:

- Identificar ARNnc candidatos diferencialmente expresados frente a Cu y antibióticos en *E. faecalis*, a través del análisis de experimentos de RNA-seq provenientes de bases de datos públicos.
- Clasificar, jerarquizar y seleccionar ARNnc utilizando análisis *in silico* de conservación y blancos de regulación.
- Determinar posibles cambios de expresión de los ARNnc seleccionados mediante ensayos de qPCR en cultivos de *E. faecalis* expuestos a Cu y antibióticos.

Capítulo 4

Materiales y Metodología

4.1. Materiales

Los materiales que se utilizaron durante el trabajo de tesis fueron catalogados en 5 grupos distintos, como se muestra a continuación.

4.1.1. Reactivos

En la Tabla 4.1 se listan los reactivos y kits utilizados para llevar a cabo los experimentos.

Tabla 4.1: Reactivos y kits usados durante el trabajo en laboratorio.

Reactivo/Kit	Fabricante
Peptona (Bacto™ Peptone)	BD Difco
Extracto de Levadura (Bacto™ Yeast Extract)	BD Difco
Na ₂ HPO ₄	Merck
Glucosa	Merck
CuSO ₄	Merck
Clorhidrato de Vancomicina	Santa Cruz Biotechnology
Ceftriaxona, Sal Disódica, Hemiheptahidrato	Santa Cruz Biotechnology
NaCl	Merck
EDTA	Merck
Kit extracción RNA	Maccura
Agua libre de nucleasas	QIAGEN
Buffer RDD	QIAGEN
Buffer RLT	QIAGEN
Buffer RW1	QIAGEN
Buffer RPE	QIAGEN
Etanol	Winkler
Partidores	IDT
2× Brilliant II SYBR Green QPCR master mix	Agilent Technologies

4.1.2. Instrumentos y elementos

Los instrumentos y elementos utilizados en la realización de experimentos son las herramientas, utensilios o contenedores presentados en la Tabla 4.2.

Tabla 4.2: Instrumentos utilizados en el trabajo de laboratorio.

Instrumento	Función
Espátula	Tomar pequeñas cantidades de reactivos en polvo.
Frasco ISO	Almacenar reactivos y medios de cultivo.
Asa bacteriológica	Inocular microorganismos en medios de cultivo.
Tubos de microcentrífuga 1.5 mL	Almacenar muestras o reactivos.
Tubos de microcentrífuga 2 mL	Almacenar filtrado de muestras.
Tubos de centrífuga 15 mL	Crece cultivos o almacenar reactivos.
Tubos de centrífuga 50 mL	Crece cultivos o almacenar reactivos.
Placa de 96 pocillos	Crece cultivos bajo condiciones de estudio.
Tubos de ensayo	Depositar muestras para medir la densidad óptica en el espectrofotómetro.
Micropipeta	Tomar volúmenes de las muestras para ser transferidos.
Gradilla	Dar soporte a tubos con muestras o reactivos.
Columna RNeasy MinElute [®] Spin	Filtrar y purificar por elución ARN de una muestra.

4.1.3. Equipos

Los equipos necesarios usados durante el trabajo de investigación se muestran en la Tabla 4.3. Se indica el nombre del equipo y su respectiva función.

Tabla 4.3: Equipos utilizados durante el trabajo de tesis.

Equipo	Función
Microcentrífuga	Centrifugar las muestras para separar el pellet del sobrenadante.
Robot de extracción de ácidos nucleicos	Extraer ARN de muestras desde el kit de extracción.
Espectrofotómetro UV-VIS	Medir la concentración de ARN de una muestra.
Balanza electrónica	Medir en miligramos o gramos algunos reactivos.
Agitador tipo vórtex	Agitar tubos para disolver solutos en solventes.
Autoclave	Esterilizar instrumentos del laboratorio a alta presión y temperatura.
Incubadora con agitación	Mantener la temperatura y agitación constantes de los cultivos.
Termociclador	Realizar el qPCR de las muestras.
Campana de Flujo Laminar	Mantener condiciones estériles y extraer aire en caso de utilizar reactivos tóxicos.
Espectrofotómetro LED	Medir la densidad óptica de las muestras.
Lector multimodo	Agitar, mantener temperatura y medir densidad óptica en placas de múltiples pocillos.
Computador	Permite el acceso al servidor, hacer el análisis <i>in silico</i> , gráfico y estadístico.

4.1.4. Herramientas Computacionales

Los programas y herramientas bioinformáticas utilizados con sus funciones principales se encuentran listados en la Tabla 4.4.

Tabla 4.4: Herramientas computacionales utilizadas durante el trabajo de tesis.

Programa/software	Función
Terminal Linux	Acceso al servidor de la Universidad, donde se almacenan y procesan los datos.
Microsoft Excel	Construcción de gráficos y cálculo de valores estadísticos.
FastQC	Análisis de la calidad de secuenciaciones de genomas obtenidas desde bases de datos.
Trimmomatic	Filtrar secuencias en los datos de RNA-seq y adaptadores utilizados en la secuenciación.
Bowtie2	Mapeo de datos de RNA-seq en genoma de referencia.
BEDTools suite	Conversión de formatos de archivos, ensamble de transcritos, comparación de ensamble con referencia, filtrado y ordenamiento de datos.
BLASTN	Búsqueda de secuencias nucleotídicas coincidentes en <i>E. faecalis</i> en bases de datos.
IntaRNA	Determinación de posibles blancos regulatorios de ARNnc identificados.
RNAfold	Predicción de estructura secundaria de ARNnc.
GraphPad Prism	Creación de gráficos y análisis estadístico de resultados.
BioRender	Diseño de figuras que incluyen estructuras celulares.
LaTeX	Creación del documento principal.

4.1.5. Microorganismo

Todos los ensamblajes y mapeo de transcritos se realizaron utilizando el genoma de referencia de *E. faecalis* V583, cuyo genoma a la fecha se considera como el mejor anotado disponible (GenBank GCA_000007785.1). Los experimentos *in vitro* se realizaron con la cepa OG1RF de la bacteria *E. faecalis*, no patógena, caracterizada por su ausencia de plasmidios, sensible al uso de diferentes antibióticos, con genoma secuenciado y anotado. El microorganismo se encuentra autorizado para uso y disponible en el Laboratorio de Bioingeniería de la Universidad de O'Higgins (LB-UOH), Rancagua.

4.1.6. Composición del Medio de Cultivo

El medio de cultivo utilizado en el crecimiento bacteriano corresponde al Medio N (*Nutrient Broth* o *Medio Nutritivo*) [69], cuya composición se muestra en la Tabla 4.5.

Tabla 4.5: Composición del Medio N líquido [70].

Componente	Concentración [% p/v]
Glucosa	1
Peptona	1
Extracto de Levadura	0.5
Na ₂ HPO ₄	1

4.2. Metodología

Para facilitar la organización y comprensión del trabajo realizado, la metodología fue dividida en dos etapas. La primera, en bioinformática, enfocada en la recopilación de información desde bases de datos, tratamiento y análisis de los datos disponibles y obtención de resultados mediante el uso de herramientas bioinformáticas. La segunda parte, en microbiología y biología molecular, la cual constó del estudio *in vitro* de la abundancia transcripcional de los ARNnc de *E. faecalis* ante la exposición a Cu y antibióticos. A continuación, se encuentra la descripción detallada de cada etapa.

4.2.1. Etapa 1: Recopilación bibliográfica y uso de herramientas bioinformáticas

4.2.1.1. Búsqueda bibliográfica

- **Publicaciones científicas:** Para la recopilación de antecedentes relacionados con el trabajo de investigación, se recurrió a la base de datos de acceso público del “National Center for Biotechnology Information” (NCBI). Utilizando las palabras clave en inglés “*Enterococcus faecalis*”, “RNAseq”, “ncRNA”, “antibiotics resistance” y/o “copper resistance”, se rastrearon trabajos de expresión génica global en la bacteria expuesta a los estímulos en estudio. Una vez encontrados los artículos relacionados con las palabras clave mencionadas, se procedió a filtrar la información obtenida, descartando las publicaciones que no evidenciaban trabajo relacionado con la secuenciación de transcritos y análisis de transcriptomas.
- **Datos de secuencia de transcriptomas:** En paralelo a la búsqueda de artículos, se hizo una búsqueda de únicamente datos de RNA-seq liberados en la base de datos SRA (Sequence Read Archive) de NCBI [71] y repositorios públicos disponibles en el LB-UOH, utilizando las palabras en inglés “antibiotics”, “copper” y “*Enterococcus faecalis*”.

El criterio de selección utilizado fue de archivos que presentaron datos de secuenciación tanto *paired-end* como *single-end* con un mínimo de 3 réplicas biológicas, todos ellos obtenidos bajo secuenciación Illumina.

4.2.1.2. Descarga y análisis de datos de RNA-seq

- **Acceso al servidor y descarga de datos:** Teniendo la selección de experimentos asociados con secuenciaciones de transcriptomas, se procedió a acceder a un servidor de la Universidad de Chile, proporcionado por el Dr. Vinicius Maracaja Coutinho, el cual cuenta con el sistema operativo Ubuntu, 5 TB de disco duro y 1 TB de memoria RAM, capacidad suficiente para cualquier análisis transcriptómico en bacterias. En esta plataforma se realizó todo el análisis *in silico* dada la gran cantidad de espacio utilizado por los archivos generados durante el estudio. Se continuó con la descarga de datos de RNA-seq para *E. faecalis* expuesta a antibióticos. En paralelo, el LB-UOH, donde se realizó el presente trabajo, contaba con datos de RNA-seq de *E. faecalis* expuesta a Cu, datos que fueron incluidos en el análisis posterior. Los archivos descargados tenían formato FASTA, que al ser comprimidos se guardaron con la extensión *fasta.gz*.
- **Análisis de calidad y filtro de datos:** El tratamiento de datos de RNA-seq comenzó con el análisis de calidad de la secuenciación mediante el uso del programa FastQC [72], donde se generó un archivo HTML con el objetivo de evidenciar gráficamente la presencia o ausencia de adaptadores en las secuencias, además de errores de lectura que pudieron ocurrir durante la secuenciación. Como criterio de eliminación se consideró secuencias que se poseían un valor ‘Q’ menor a 28 y ausencia de adaptadores. Este paso se realizó a través del programa Trimmomatic [73], indicando al archivo *fasta.gz* como *input* en el código, además de cada parámetro deseado. Una vez aplicado el filtro, sólo se conservaron las lecturas mayores a 25 pares de bases, puesto que se disminuye el error de posición al momento de alinear la secuencia con el genoma de referencia. Para comprobar que la remoción de adaptadores y la mejora de la calidad fue exitosa, se analizó nuevamente las secuencias utilizando FastQC. En caso de que se identificaran algunas lecturas de baja calidad, se repitió el paso anterior con la herramienta Trimmomatic hasta que los datos cumplieran con las condiciones deseadas mencionadas para continuar con los pasos posteriores. Este criterio fue utilizado en base a protocolos estandarizados y publicados anteriormente [73].
- **Mapeo de secuencias en genoma de referencia:** Lo siguiente fue alinear las secuencias filtradas con el genoma de referencia, en este caso se utilizó el de la cepa de *E. faecalis* V583 (GenBank GCA_000007785.1). El archivo se obtuvo desde el SRA de NCBI y se descargó en formato FASTA. El mapeo se realizó utilizando la herramienta Bowtie2 [74], generando archivos con extensión SAM que posteriormente fueron transformados a BED, lo que permite asegurar el correcto funcionamiento de las herramientas siguientes en el análisis, estrategia utilizada previamente [74].

- **Ensamble de transcritos:** Los archivos BED generados fueron procesados para eliminar las columnas innecesarias para el ensamble. Estos datos editados sirvieron como entrada para la herramienta `sortBed` y `mergeBed` [74], que se encargaron de ordenar y transformar los datos a una extensión `.merge`. En esta etapa fue necesario indicar en el código si los datos son hebra específicos o no, dependiendo del tipo de secuenciación hecha. Luego, se concatenaron los archivos `.merge` de un mismo experimento, generando un archivo único concatenado por condición. Este archivo fue utilizado como entrada para realizar el ensamble final, el cual se generó a través de los programas `sortBed` y `mergeBed`, generando un archivo BED.
- **Comparación con referencia y filtrado:** Los transcritos ensamblados fueron comparados con el transcriptoma de ARNm de la cepa V583 de *E. faecalis*, para eliminar la existencia de similitudes entre las coordenadas identificadas entre ambos transcritos, o bien si había una sobreposición entre ellos. Para esto, la herramienta utilizada fue `intersectBed`, especificando si los datos eran o no hebra específicos. El filtrado posterior se realizó con el comando `awk` que se basa en el número de bases que tienen coincidencia con la anotación de la referencia; como se trata de la identificación de ARNs no codificantes y la anotación posee ARNs codificantes, se especificó que no haya coincidencia, es decir, cero bases. Finalmente se hizo un último filtrado considerando el número de bases de distancia a un gen o característica genómica, además del largo del transcrito ensamblado, utilizando `closestBed` y `awk`. En este caso, se deseaba que la distancia a un gen no fuera menor a 30 bases, pues podía ser parte de la región UTR de un gen, mientras que el largo máximo debía ser de 1500 bases. Esto generó una lista final de ARNnc diferencialmente expresados para cada una de las condiciones antibióticas y de exposición a Cu.

4.2.1.3. Comparación de la expresión de transcritos

Una vez detectadas las secuencias de ARNnc, se hizo un conteo de cada transcrito en cada condición de estudio, generando un archivo que permitiera observar específicamente el número de cuentas de cada ARNnc por secuenciación. Luego, se realizó una comparación entre cada condición de estrés estudiada con respecto a la condición basal, es decir, para el Cu o cada antibiótico con su respectivo control, obteniendo los ARNs no codificantes diferencialmente expresados. Este último paso se realizó en el programa RStudio [75] con lenguaje de programación en R, cuyo código se encuentra en el Anexo A.2.1. Posteriormente, de las listas obtenidas se compararon entre ellas con el objetivo de identificar ARNnc comunes y únicos entre todos los experimentos, seleccionando aquellos que se compartían entre Cu y otro antibiótico, conformando la lista final de posibles candidatos como ARNnc inducidos por los agentes estresores.

4.2.1.4. Clasificación y selección de los ARNnc candidatos

La selección de candidatos se realizó a través de herramientas bioinformáticas disponibles en plataformas de libre acceso. Primero, se observó que la expresión diferencial de los ARNnc tuviera valores estadísticamente significativos en las 3 condiciones de estudio con un *p-value*

menor a 0.05. Segundo, a través del programa IntaRNA [76], se determinaron los posibles blancos regulatorios de los ARNnc, es decir, la identificación de alguna posible función dentro del genoma, privilegiando resistencia a Cu o antibióticos. Tercero, se privilegió ARNnc sin repetición en el genoma y con un nivel de expresión diferencial (valor de \log_2FC) de al menos dos veces de cambio (aumento o disminución) comparado con el estado basal, en las 3 condiciones. Por último, se compararon las secuencias de los ARNnc con sus respectivos genes blanco mediante Clustal Omega [77], para confirmar que dichas secuencias eran complementarias y poseían un alto grado de identidad en caso de ser regulados en *cis*.

Para más detalles sobre los comandos utilizados directamente en el servidor e imágenes asociadas al análisis descrito en las etapas anteriores, ver Anexo A.

4.2.1.5. Predicción de estructuras de los ARNnc, conservación y codificación de genes blanco

Con el fin de conocer la estructura secundaria de los ARNnc, se utilizó el *software* RNAFold para predecir las interacciones entre las bases nitrogenadas, que permite entregar un modelo de plegamiento asociado a un valor de energía libre de la estructura secundaria óptima para cada ARNnc [78]. Luego, se estudió la conservación de los ARNnc en distintas especies, utilizando la herramienta BLASTN con un *word size* de 16 nucleótidos [79]. Finalmente, se estudió la estructura consenso entre todas las secuencias identificadas en diferentes cepas o especies para el mismo ARNnc en cuestión, mediante el uso de LocARNA [80].

4.2.2. Etapa 2: Validación *in vitro* y análisis de resultados

4.2.2.1. Crecimiento bacteriano bajo condiciones control

Para todos los experimentos posteriores, este fue el protocolo que se realizó para establecer las condiciones de crecimiento iniciales. Desde una placa de Petri sembrada con *E. faecalis* cepa OG1RF, se tomó una colonia aislada y se hizo crecer en medio N por toda una noche a 37°C y 150 RPM, que corresponde a una condición control o basar previamente validada en estudios anteriores [69]. Una vez crecido el cultivo se hizo un refresco, diluyendo en medio N con proporción 1:20 y se dejó creciendo por 2 horas a 37°C y 150 RPM para alcanzar una DO_{600nm} no superior a 2.0. Luego, se volvió a diluir la mezcla para alcanzar una DO_{600nm} de 0.1 como absorbancia inicial en el cultivo para todos los experimentos posteriores.

4.2.2.2. Determinación de concentraciones de agentes estresores

Se estudiaron 4 diferentes condiciones para analizar el comportamiento en el crecimiento de bacterias: control, Cu, ceftriaxona y vancomicina (antibióticos seleccionados en base a la búsqueda en literatura, ver la sección de Resultados 5.1.1). Primero, se debió determinar la concentración a utilizar de cada uno de los estresores; para ello se utilizó un ensayo de concentración mínima inhibitoria, MIC por su abreviatura en inglés (*minimum inhibitory concentration*). En una placa de 96 pocillos se agregaron 100 μ L de un cultivo de *E. faecalis*

OG1RF en medio N a una DO_{600nm} igual a 0.1 en cada pocillo. Posteriormente se prepararon diluciones seriadas de cada uno de los estresores tal como indica la Figura 4.1. Luego, la placa se dejó creciendo por un periodo de 8 horas con una agitación preestablecida por el equipo Synergy HTX Multimode Reader, de 282 RPM a $37^\circ C$, registrando la turbidez del cultivo (crecimiento) por absorbancia a DO_{600nm} cada 30 minutos. Obtenidos estos datos, se construyeron curvas de crecimiento, donde se comparó el crecimiento de la bacteria entre el medio con y sin estresores (control). Diferencias con el control se determinaron utilizando un análisis estadístico ANOVA, explicado en detalle más adelante.

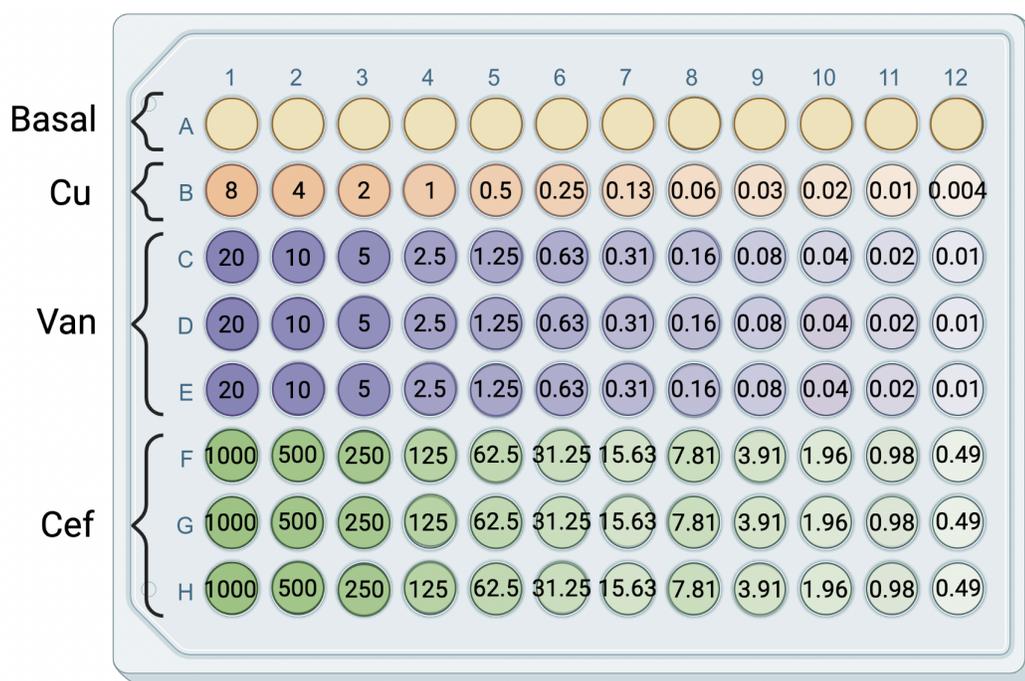


Figura 4.1: Placa de 96 pocillos con diluciones en serie. La ilustración muestra de manera gráfica cómo se realizaron las diluciones en serie, donde una mayor intensidad en los colores, en este trabajo, indican una mayor concentración de Cu, vancomicina o ceftriaxona.

Este experimento permitió determinar los valores máximos de concentraciones de estresores que no afecten el crecimiento de la bacteria (estado de homeostasis). Finalmente, puesto que los ensayos posteriores requieren un volumen mayor de trabajo, se confirmaron los valores máximos de los estresores en un cultivo de 15 mL. Para ello, escalando el mismo procedimiento, en un tubo de 50 mL se realizaron curvas de crecimiento en un volumen final de 15 mL de medio N, con y sin la adición de los estresores, registrando cada 30 minutos por 4 horas la turbidez a través de absorbancia a una DO_{600nm} . Diferencias entre las curvas se obtuvieron mediante un análisis estadístico ANOVA, explicado en detalle más adelante.

4.2.2.3. Extracción de ARN

La extracción de ARN se realizó en base a protocolos previamente validados y publicados [81][82]. Se utilizaron tubos de centrifuga de 50 mL con 15 mL de volumen total de cultivo en

triplicado. El procedimiento de crecimiento es el mismo descrito en la sección 4.2.2.1, tomando muestras de cultivo luego de aproximadamente 3 horas de exposición a los estresores. Para esto, pasado el tiempo se tomaron muestras de cada cultivo, se midió la DO_{600nm} en cada tubo y se ajustaron los valores de DO_{600nm} a la de menor valor, para normalizar el cultivo. Luego, de cada tubo se tomó un volumen a 5 mL, el cual se traspasó a nuevos tubos de centrífuga de 15 mL. Trabajando con guantes y hielo, se centrifugaron a 9000 RPM a 4°C por 15 minutos y se descartó el sobrenadante. Para eliminar el exceso de metal y antibióticos del pellet recuperado, se comenzó el proceso de lavado agregando 1 mL de NaCl 0.85 % a 4°C a cada tubo y se traspasó el resuspendido a tubos de microcentrífuga de 1.5 mL, centrifugando a 14,000 RPM por 3 minutos a 20°C. El segundo lavado, se realizó con 800 μ L de NaCl 0,85 % a 4°C, mientras que para el tercero se agregaron 800 μ L de EDTA 1 mM a 4°C. Para finalizar con los lavados, se añadieron 800 μ L de agua destilada estéril a 4°C a cada tubo. Cabe destacar que después de la adición de reactivos en cada lavado, se resuspendió con vórtex y se centrifugaron los tubos a 14,000 RPM por 3 minutos a 20°C, para posteriormente eliminar el sobrenadante y seguir con el siguiente paso.

Continuando con la extracción de ARN, los pellets resultantes se resuspendieron en 1 mL de lisozima con concentración final de 10 mg/mL para lisar las bacterias. Se incubaron las muestras a 37°C por 20 minutos y se procedió a extraer el ARN mediante el proceso automatizado en Auto-Pure32A Maccura Biotechnology, utilizando el kit RNA Mag-Bind RNA Extraction de Maccura Biotechnology y siguiendo las instrucciones del proveedor [81]. Terminada la extracción, se recuperaron las muestras en tubos de microcentrífuga de 1.5 mL y se realizó una digestión por DNasa utilizando el kit QIAGEN RNEasy, siguiendo el protocolo y recomendaciones del fabricante [82]. Una vez tratadas las muestras se agregaron 50 μ L de agua libre de nucleasas directamente a la membrana de las columnas, para posteriormente centrifugar a 8,000 g por 1 minuto para obtener el ARN total concentrado. Finalmente, se midió la concentración de ARN resultante de cada tubo en el espectrofotómetro NanoDrop™ One de Thermo Scientific™ para cuantificar la cantidad de ARN en concentración.

4.2.2.4. PCR en tiempo real (qPCR)

La síntesis de ADNc de hebra simple se realizó *in vitro* mediante el uso de partidores específicos usando como material de partida el ARN de las muestras tratamiento y control obtenidas en el punto anterior. La reacción de transcripción reversa se realizó a partir de 5 μ g de ARN total, 20 pmol del partidor correspondiente (2 μ L) y se aforó a un volumen de 12 μ L de agua libre de nucleasas. Se incubaron las muestras por 5 min a 65°C, y pasado este tiempo se agregaron 4 μ L de buffer de reacción 5x, 1 μ L de RiboLock RNase Inhibitor (20 U/ μ L), 2 μ L de dNTP Mix (10 mM) y 1 μ L de la enzima Transcriptasa Reversa RevertAid (200 U/ μ L). Para cada muestra el volumen final de reacción fue de 20 μ L. Se incubaron durante 1 hora a 42°C y calentadas posteriormente a 70 °C por 5 minutos para terminar la reacción. Las muestras se almacenaron a -20°C.

Para la cuantificación de la abundancia relativa de los ARNnc por qPCR, se utilizó un mix para cada ARNnc seleccionado y un mix para el gen normalizador 16S, en cada condición de estudio. A cada mix se agregaron 200 μL de 2x Master Mix, 24 μL del partidor correspondiente y 128 μL de agua libre de nucleasas. Las mezclas fueron distribuidas en una placa de PCR de 96 pocillos, agregando 22 μL del mix de cada partidor según correspondía. Cada ADNc obtenido de la síntesis anterior se diluyó con proporción 1:4; una parte de ADNc por 4 partes de agua libre de nucleasas. De estas diluciones, 3 μL fueron agregados a cada pocillo para lograr triplicados de cada condición, dejando blancos donde en vez de agregar ADNc se agregaron 3 μL de agua libre de nucleasas.

La reacción se llevó a cabo en el termociclador en tiempo real Bio-Rad CFX96, en un programa de 3 pasos según las instrucciones del manual Brilliant [83]: activación de la enzima por 10 minutos a 95°C (1 ciclo), amplificación de las secuencias (40 ciclos) con 30 segundos a 95°C, 1 minuto a 58°C y 30 segundos a 72°C, y finalmente se realiza la disociación siguiendo las instrucciones del fabricante. La obtención de los datos se realizó utilizando el programa Light Cycler Software v3.5 para posteriormente ser procesados y analizados.

4.2.2.5. Procesamiento de resultados de qPCR

Al obtener los resultados del qPCR, se necesitó cuantificar la cantidad de material genético detectada en cada muestra. Esto se expresó como la razón de cambio entre la abundancia normalizada del transcrito en la muestra tratada con respecto a la condición basal, y para ello se tomaron los valores de Ct entregados por el programa del termociclador. El valor de Ct (*cycle threshold* o umbral de ciclos), corresponde al número de ciclos en el cual la señal de fluorescencia detectada en el proceso de amplificación cruza un determinado umbral; mientras menor sea el valor del Ct, mayor es la cantidad de ARN presente en la muestra inicial. Así, la abundancia del transcrito normalizado se calculó utilizando como referencia un transcrito que no variara su abundancia en las distintas condiciones de tratamiento, que en este caso fue el ARN 16S [84] como *housekeeping*, utilizando la ecuación 4.1.

$$\Delta\Delta Ct = \Delta Ct_{\text{muestra tratada}} - \Delta Ct_{\text{muestra basal}} \quad (4.1)$$

donde cada ΔCt se calculó como:

$$\begin{aligned} \Delta Ct_{\text{muestra tratada}} &= \Delta Ct_{\text{muestra tratada}}^{\text{transcrito de interés}} - \Delta Ct_{\text{muestra tratada}}^{\text{transcrito housekeeping}} \\ \Delta Ct_{\text{muestra basal}} &= \Delta Ct_{\text{muestra basal}}^{\text{transcrito de interés}} - \Delta Ct_{\text{muestra basal}}^{\text{transcrito housekeeping}} \end{aligned}$$

Los valores de abundancia relativa fueron finalmente calculados como $2^{-\Delta\Delta Ct}$, los que fueron comparados y analizados para observar las diferencias en expresión de cada transcrito en las distintas condiciones de estudio. Cabe mencionar, además, que los valores de expresión diferencial expresados en $\log_2\text{FC}$ en el análisis de las secuenciaciones es una forma análoga de medir expresión al valor de $2^{-\Delta\Delta Ct}$ en los resultados de qPCR.

4.2.2.6. Análisis estadístico

Las réplicas técnicas y biológicas para cada experimento se detallan en el pie de las figuras o en la descripción de los protocolos, según corresponda, reportándose el promedio de las mediciones obtenidas con sus respectivas desviaciones estándar. Los datos se analizaron estadísticamente mediante un análisis de varianza ANOVA de una vía, con una prueba de Dunnet para establecer diferencias entre tratamientos y el control, utilizando el software GraphPad Prism 8.0. Ambos análisis se aplicaron utilizando los parámetros establecidos de base, considerando diferencias estadísticamente significativas si $p\text{-value} < 0.05$.

Capítulo 5

Resultados y Discusión

Como fue descrito en la metodología, el trabajo de tesis fue dividido en dos partes, por lo que los resultados y discusión serán presentados de la misma manera, paso a paso.

5.1. Identificación *in silico* de ARNnc

5.1.1. Análisis de datos de RNA-seq

Luego de buscar datos de secuencia de transcriptomas asociados a *E. faecalis* expuesto a cobre y antibióticos, desde la base de datos de NCBI se obtuvieron 4 conjuntos de datos de secuenciaciones de ARN total para *E. faecalis* en condiciones de estrés por distintos antibióticos: teixobactina [23], linezolid [85], vancomicina y ceftriaxona [86], correspondientes a 3 investigaciones previas distintas. Siguiendo los pasos especificados en la sección 4.2.1.2 de metodología, al filtrar los datos por coordenadas genómicas se obtuvo que para los primeros dos conjuntos de datos referentes a teixobactina y linezolid no fue posible identificar ARNnc, pues la cantidad de lecturas (*reads*) eran en su mayoría de regiones codificantes (CDS), siendo insuficientes para lograr ensamblar ARNnc.

Por el contrario, los datos de secuencias provenientes de *E. faecalis* V583 tratadas con vancomicina y ceftriaxona entregaron resultados favorables para la identificación de ARNnc. Al respecto, los tratamientos correspondieron a 256 $\mu\text{g}/\text{mL}$ de vancomicina por 45 minutos, y 2048 $\mu\text{g}/\text{mL}$ de ceftriaxona por 60 minutos. A este conjunto de datos, se incluyó el resultado de RNA-seq para el tratamiento de *E. faecalis* OG1RF con 0.5 mM de Cu y extracción de ARN a los 180 minutos, disponibles en el LB-UOH, lugar donde se ejecutó la presente tesis.

En relación con los resultados, de todos los datos analizados se encontró un total de 3207 transcritos que codifican para proteínas y que poseen función parcial o totalmente descrita, tanto en el análisis de RNA-seq de Cu como de antibióticos, mientras que los ARNnc detectados suman un total de 196 para antibióticos y 586 para Cu. En el caso de los genes codificantes de proteínas, el valor encontrado se acerca bastante al descrito para proteínas en bases de datos, pues para la cepa de *E. faecalis* V583 se ha encontrado un total de 3245

[87], mientras que el máximo número de ARNnc detectados en esta bacteria ha sido de 230 [65], todos correspondientes a sARN.

Los diagramas de Venn de la Figura 5.1 muestran la conservación del total de ARNnc en los tres RNA-seq analizados (5.1.A), cuya expresión positiva o negativa se comparte en los tratamientos de Cu, vancomicina o ceftriaxona (5.1.B). Al comparar los ARNnc identificados a partir de los datos de RNA-seq para cada condición, se observa que algunos de ellos se identificaron de forma total o parcial en las 3 condiciones, pero con distinto nivel de expresión. Las diferencias de tamaño observadas se explican puesto que algunos ARNnc no lograron ser identificados debido a la falta de cobertura en algunos datos. Como alternativa, se pueden modificar algunos parámetros utilizados en los filtros haciéndolos menos estrictos, sin embargo, esto no asegura que las secuencias identificadas sean de buena calidad o que presenten una expresión significativa, por lo que no se realizaron cambios a los protocolos publicados.

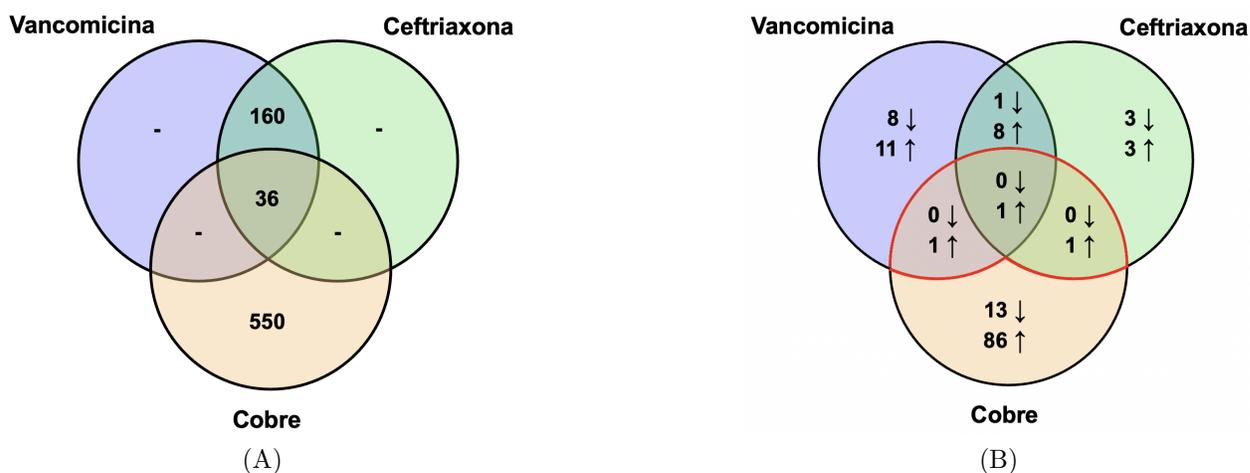


Figura 5.1: Diagrama de Venn para ARNs no codificantes identificados en los RNA-seq de Cu, vancomicina y ceftriaxona. (A) El diagrama muestra el total de ARNnc identificados en cada uno de los RNA-seq, además de cuáles de ellos se encontraron en más de un tratamiento. (B) Se observan aquellos ARNnc que con respecto a la condición basal aumentan su abundancia (flecha hacia arriba) o disminuyen su abundancia (flecha hacia abajo) en relación con el control. En rojo se marca el conjunto de ARNnc diferencialmente expresados en Cu y otro antibiótico.

5.1.2. Identificación y comparación de la expresión diferencial de ARNs no codificantes de respuesta a Cu y antibióticos

Obtenidos e identificados los transcritos de las secuenciaciones, el siguiente paso fue filtrar aquellos que no presentaran una respuesta diferencial significativa al tratamiento de Cu o antibióticos con respecto a la bacteria en crecimiento basal. Para ello se realizó una tabla de expresión diferencial a partir del código presentado en el Anexo A.2.1, obteniendo la cantidad

de veces de cambio (\log_2FC) o *Fold Change* con el *p-value* correspondiente de cada transcrito en condiciones de estrés con respecto a la condición basal. Al ser normalizado, este análisis estadístico permitió distinguir qué transcritos presentaron una expresión diferencial dentro del total de ARNnc. Como punto de corte, se seleccionaron aquellos ARNnc que presentaron un valor de \log_2FC sobre 1 (aumenta) o menor a -1 (disminuye), en ambos casos dos veces de cambio al menos en relación a la condición basal y con un *p-value* menor a 0.05.

La Tabla 5.1 muestra los resultados de las veces de cambio con respecto a la condición basal (\log_2FC) para vancomicina, ceftriaxona y Cu para los ARNnc cuyas secuencias fueron encontradas en los experimentos de RNA-seq (36 en total). Diferencias entre las cantidades de ARNnc totales encontradas en cada tratamiento se puede deber a razones como que provienen de cepas distintas (OG1RF para Cu y V583 para los antibióticos), al tipo de secuenciación utilizada, diferencias en el tiempo de tratamiento, o simplemente a la respuesta hacia el tratamiento, todos factores que combinados podrían explicar los resultados obtenidos. A pesar de estas diferencias, se identificó que 36 de los ARNnc se presentan en las 3 condiciones a distintos niveles de expresión, dando cuenta de un universo de elementos conservados entre cepas de *E. faecalis* que eventualmente responden a los tres estímulos, independientemente de la forma de cómo se aplicaron.

Uno de los objetivos del trabajo de tesis fue identificar secuencias que respondan de forma diferencial al menos a Cu y uno de los antibióticos, ya sea aumentando o disminuyendo Tomando los parámetros de corte para \log_2FC y *p-value*, aumentaron su abundancia de transcrito el ARNnc “assembly_3089” en Cu y ceftriaxona, el ARNnc con ID “assembly_3340” en Cu y vancomicina, mientras que el ARNnc “assembly_3683” presentó un aumento diferencial frente a los tres estresores. Estos resultados apoyan la hipótesis del trabajo, no obstante, se requiere una caracterización de su posible función, conservación y posterior validación del cambio de expresión para ratificar este punto.

Tabla 5.1: Expresión diferencial de los ARNnc en exposición a Cu, vancomicina y ceftriaxona. Destacados en celeste se muestran aquellos que presentaron una expresión diferencial en Cu y al menos un antibiótico.

ID ARNnc	Cu		Vancomicina		Ceftriaxona	
	\log_2FC	<i>p-value</i>	\log_2FC	<i>p-value</i>	\log_2FC	<i>p-value</i>
assembly_1152	-	NA	0.651	0.522	1.547	0.132
assembly_1299	-	NA	-0.918	0.000	-0.424	0.075
assembly_1539	-	NA	0.872	0.163	0.718	0.249
assembly_1772	0.885	0.213	0.503	0.474	0.294	0.644
assembly_1925	1.092	0.542	-0.066	0.916	0.226	0.723
assembly_2527	1.457	0.316	-0.637	0.311	-0.559	0.375
assembly_2765	2.014	0.000	0.050	0.928	-0.547	0.311
assembly_2839	-0.825	0.523	0.079	0.890	-0.117	0.836
assembly_3084	0.749	0.656	1.926	0.004	1.733	0.019
assembly_3089	2.701	0.037	1.194	0.060	1.212	0.042
assembly_3186	0.073	0.987	0.817	0.209	0.750	0.250
assembly_3340	2.517	0.029	3.454	0.001	2.372	0.065
assembly_366	-	NA	-0.868	0.065	-1.218	0.008
assembly_3683	1.647	0.028	1.806	0.020	1.998	0.044
assembly_3762	-1.438	0.007	0.471	0.042	0.508	0.029
assembly_3811	1.542	0.685	1.093	0.147	0.700	0.377
assembly_4314	-0.501	NA	1.362	0.021	1.785	0.002
assembly_4404	0.898	0.470	-0.912	0.138	-1.000	0.105
assembly_4427	0.567	0.679	1.373	0.022	1.681	0.008
assembly_4630	1.094	0.104	-0.479	0.435	-0.487	0.406
assembly_4822	-0.743	0.153	0.636	0.193	-0.151	0.760
assembly_6068	2.616	0.109	0.618	0.401	-0.006	0.993
assembly_6166	2.334	0.227	0.046	0.949	0.140	0.848
assembly_6176	5.374	0.028	-0.069	0.894	0.242	0.631
assembly_6736	1.688	0.125	0.636	0.394	1.003	0.157
assembly_6848	1.155	0.037	-0.199	0.600	0.102	0.790
assembly_7044	0.855	0.396	1.092	0.008	1.059	0.011
assembly_7518	-	NA	-0.697	0.004	-0.564	0.020
assembly_7580	0.399	0.649	1.070	0.121	0.912	0.263
assembly_7612	-0.281	0.674	1.539	0.029	-0.614	0.345
assembly_7865	2.340	0.252	1.263	0.043	1.576	0.017
assembly_8024	0.622	0.445	2.360	0.000	1.458	0.034
assembly_8062	1.492	0.267	1.056	0.137	1.184	0.081
assembly_8788	-2.701	0.046	-0.873	0.124	-0.631	0.243
assembly_8830	0.954	0.418	0.246	0.705	-0.289	0.678
assembly_9129	3.646	0.008	-1.850	0.065	-0.932	0.354

5.1.3. Selección final de ARNnc

Como se mencionó anteriormente, para poder seleccionar los ARNnc a estudiar en el análisis *in vitro*, primero se filtró la lista de ARNnc diferencialmente expresados en base a las veces de cambio en las 3 condiciones de estudi. De esta forma se obtuvo como posibles candidatos a los ARNnc “assembly_3089”, “assembly_3340” y “assembly_3683”. Luego de esto, se decidió incluir un segundo criterio en base a los genes blancos de ARNnc, predicción obtenida mediante el uso de la herramienta IntaRNA y RNA Predator, los cuales permiten predecir posibles interacciones entre ARNs, en este caso entre ARNnc y ARNm [76]. Ambas herramientas entregan valores de energía de interacción, en particular IntaRNA predice además interacciones con estructuras secundarias posiblemente formadas dependiendo de la secuencia de bases del ARNnc en cuestión, aumentando las posibilidades de interacción con genes blanco. Luego de utilizar ambos programas, como resultado se obtuvieron los blancos de regulación para cada ARNnc identificado, especificando además si la regulación era en *cis* o en *trans* (ver Tabla 5.2). En base a los resultados, en promedio se encontró un gen blanco regulado en *cis*, mientras que dos o tres se encontraban en *trans*. Cabe destacar que la complementariedad en *cis* es completa en la mayoría de los casos, mientras que en *trans* el ARNnc la complementariedad disminuye entre el ARNnc y su blanco dado las diferencias entre las secuencias.

En términos de la selección, se privilegió aquellos ARNnc que tuvieran blancos que codificaran para componentes involucrados en resistencia a Cu o antibióticos, dado esto, aunque algunos ARNnc podrían no tener una expresión diferencial bajo los tratamientos en los experimentos de RNA-seq, de igual forma se podrían incluir considerando su blanco de regulación y posteriormente ser analizados bajo un ensayo de qPCR.

No obstante, como se observa en la Tabla 5.2, para las proteínas codificadas por los blancos no existe una función directa con resistencia u homeostasis hacia los estresores. Considerando este punto, en la elección final se utilizó únicamente los resultados de cambios de los niveles de abundancia, seleccionando aquellos que presentaron las mayores diferencias entre el tratamiento y el estado basal. Así se seleccionaron los transcritos “assembly_3683” por presentar una expresión diferencial superior a 1.6 veces en las 3 condiciones de estudio, y “assembly_3340” por tener mayor expresión diferencial comparado entre el Cu y otro antibiótico, mostrando un cambio de al menos 2.3 veces más que la condición basal. En adelante, estas secuencias fueron denominadas “nc3683” y “nc3340”, respectivamente.

Tabla 5.2: Función de proteínas codificadas por genes blanco de los ARNnc identificados en Cu, vancomicina y ceftriaxona. Destacados en celeste se muestran los ARNnc seleccionados para el estudio por qPCR.

ARNnc	Blanco en cis	Función de la proteína codificada por el gen blanco en cis	Blanco en trans	Función de la proteína codificada por el gen blanco en trans
assembly_1152	-	-	EF1109 EF1115 EF1122	Proteína de unión al grupo de hierro y azufre Fenilalanil-tRNA sintetasa Ribonucleasa PH/Ham1
assembly_1299	EF0528 EF0529	Función desconocida Transposasa	EF0407 EF0857 EF1725	Regulador transcripcional, putativo Proteína hipotética Ligasa formiato-tetrahidrofolato
assembly_1539	EF0606	Proteína de la familia Dps	EF0941 EF0983 EF1262	Transportador ABC, permeasa Regulador transcripcional de la familia ArgR Proteína hipotética
assembly_1772	EF0696 EF0697	Tagatosa 1,6-difosfato aldolasa Proteína hipotética	EF0482 EF1327	Proteína hipotética Proteína de la familia ATPasa BadF/BadG/BcrA/BcrD
assembly_1925	EF0746 EF0747	Proteína de unión a penicilina, putativa Proteína hipotética	EF0817 EF0927 EF1033	Sistema PTS, componente IID Sensor histidina quinasa 6-aminohexanoato-dímero cíclico hidrolasa, putativa
assembly_2527	EF0949 EF0950 EF0951	Fosfotransacetilasa Proteína hipotética Acetiltransferasa de la familia GNAT	EF2307	Proteína hipotética
assembly_2765	EF1038	Lipoproteína, putativa	EF0704 EF1877	Lipoproteína, putativa Proteína de membrana, putativa
assembly_2839	EF1068	Aldosa 1-epimerasa	EF1598 EF2378	Desoxirribodipirimidina fotoliasa ADN polimerasa III, subunidad alfa
assembly_3084	EF1132 EF1133	Proteína con dominio CBS 2,3,4,5-tetrahidropiridina-2-carboxilato N-succiniltransferasa	EF0258 EF2435	Proteína de la familia de la biosíntesis de polisacáridos Sistema PTS, componentes IIBC
assembly_3089	EF1135	Proteína hipotética	EF0202 EF1394	Fosfatometilpirimidina quinasa, putativa Proteína hipotética
assembly_3186	EF1176	Proteína hipotética	EF0178 EF2515	Transportador ABC, proteína de unión a ATP Proteína hipotética
assembly_3340	EF1231	Proteína hipotética	EF0972 EF1269	Proteína de la familia de las exonucleasas reparadoras del ADN Proteína de la familia de anclaje a la superficie de la pared celular
assembly_366	EF0149	Sustancia de agregación, putativa	EF0836 EF1097 EF1619	Proteína hipotética Proteína hipotética Proteína de mecanismo de concentración de dióxido de carbono CcmL, putativa
assembly_3683	EF1362	Proteína hipotética	EF_RS02880 EF_RS02885 EF_RS11035	Función desconocida Función desconocida Función desconocida
assembly_3762	-	-	EF2070 EF2268 EF2320	ARNt (5-metilaminometil-2-tiouridilato)-metiltransferasa Proteína hipotética Proteína traE, putativa
assembly_3811	EF1406	Excinucleasa ABC, subunidad C	EF0062 EF2063	Proteína de la familia de las 5'-nucleotidasas Regulador transcripcional de la familia AraC
assembly_4314	-	-	EF0450 EF1134 EF1529	Mandelato racemasa/muconato, proteína de la familia de enzimas lactonizantes Peptidasas de la familia M20/M25/M40 Sistema PTS, componente IIC, putativo
assembly_4404	EF1624 EF1625 EF1626	Aldehído deshidrogenasa, putativa Familia de proteínas de microcompartimentos Proteína de utilización de etanolamina EutL	EF1623 EF_RS02065	Proteína de microcompartimento Función desconocida
assembly_4427	EF1632 EF1633	Sensor histidina quinasa Regulador de respuesta	EF1877	Proteína de membrana, putativa
assembly_4630	EF1693 EF1694	Proteína con dominio KH Proteína ribosomal S16	EF0950	Proteína hipotética
assembly_4822	EF1764	Proteína de interfaz de subunidad ribosómica	EF0573 EF1864	Proteína hipotética Regulador de la respuesta de unión al AND
assembly_6068	EF2164 EF2165 EF2166	Proteína de membrana, putativa Epimerasa dependiente de NAD Proteína de membrana, putativa	EF0646 EF0676	Proteína dependiente de NAD de la familia de la epimerasa/deshidratasa Represor de arginina
assembly_6166	EF2212	Proteína hipotética	EF1511 EF1718	Mandelato racemasa/muconato, proteína de la familia de enzimas lactonizantes Dihidroorotasa
assembly_6176	EF2217	Alfa-1,2-manosidasa, putativa	EF1201 EF2047	Proteína hipotética Proteína de la familia de las permeasas de aminoácidos

assembly_6736	EF2367 EF2368 EF2369	N-acetilmuramoiL-L-alanina amidasa Proteína hipotética Proteína hipotética	EF0967 EF1360 EF2513	Proteína hipotética Proteína de la familia de las dihidroxiacetona cinasas Lipoproteína, putativa
assembly_6848	EF2415 EF2416	Proteína hipotética Proteína ribosomal S21	EF0813	Glicosil hidrolasa, familia 35
assembly_7044	-	-	EF0901 EF1030 EF2199	Isopentenil difosfato delta isomerasa, putativa
assembly_7518	EF2632	Transposasa	EF0407 EF0857 EF1725	Regulador transcripcional, putativo Proteína hipotética Ligasa formiato-tetrahidrofolato
assembly_7580	EF2658	Proteína de la familia FemAB	EF0502 EF0801 EF1651	Proteína de membrana, putativa Leucil-tRNA sintetasa Proteína de infección abortiva
assembly_7612	EF2677	Regulador negativo de la competencia genética MecA, putativo	EF0728 EF1480 EF1977	ARN metiltransferasa de la familia TrmA Proteína hipotética Proteína hipotética
assembly_7865	EF2796	Proteína hipotética	EF0194 EF0328 EF0720	Butanol deshidrogenasa dependiente de NADH, putativa Proteína hipotética Proteína de la familia de los canales de cloruro dependientes de voltaje
assembly_8024	EF2860	Proteína de la familia ErfK/YbiS/YcfS/Ynh, putativa	EF0503 EF1093 EF2435	Proteína hipotética Proteína de la familia de anclaje a la superficie de la pared celular Sistema PTS, componentes IIBC
assembly_8062	EF2890 EF2891	Glicosil transferasa Glicosil transferasa	EF0573 EF2114 EF2343	Proteína hipotética Adenina metiltransferasa, putativa Proteína de la familia FtsK/SpoIIIE
assembly_8788	EF3160	Proteína hipotética	EF0967 EF1538 EF1784	Proteína hipotética Proteína hipotética Fosforribosilformilglicinamida sintasa
assembly_8830	EF3173	Proteína hipotética	EF0489 EF0880 EF1234	Proteína hipotética Defosfo-CoA quinasa, putativa Transportador ABC, proteína de unión a sustrato, putativa
assembly_9129	EF3286	Regulador transcripcional de unión a fosfosazúcar, putativo	EF0076 EF0369 EF0971	Oxidoreductasa de la familia de deshidrogenasas/reductasas de cadena corta Hidrolasa, familia similar a la deshalogenasa del haloácido Proteína hipotética

5.1.4. Posición en el genoma de los ARNnc seleccionados con sus blancos de regulación

La Figura 5.2, muestra a los dos candidatos con sus respectivos blancos de regulación. Dentro de las proteínas codificadas por los genes blanco, se destaca el alto número de proteínas hipotéticas o desconocidas sin función asignada, controladas preferentemente a nivel transcripcional por el ARNnc nc3683. Este resultado puede abrir un campo a futuro sobre la posible participación de estos componentes dentro de la respuesta a Cu y antibióticos, puesto que además, como se puede observar en la Figura 5.2, cambian la abundancia de su ARNm frente a los estímulos. En este contexto, con el objetivo de poder asignar una posible función a estos elementos y considerando que la anotación de estos genes data de la secuenciación del genoma de la cepa V583 en el año 2003, se realizó un BLASTP con la secuencia de estas proteínas frente a toda la base de datos del NCBI. Como era de esperarse, dada la disponibilidad con el paso de los años de un mayor número de proteínas con funciones conocidas, fue factible asignar a las proteínas hipotéticas codificadas por EF1362 y EF1231 una posible función.

El gen EF1362 regulado en *cis* por el ARN nc3683 codifica para una proteína hipotética del tipo peptidil prolil isomerasa, cuya función está dada por la interconversión de los isómeros *cis* y *trans* de los enlaces peptídicos con el aminoácido prolina. Por su parte, el ARN nc3340 regula en *cis* a EF1231, quien codifica para una proteína hipotética del tipo metalofosfoesterasa, la cual eventualmente podría estar implicada en la transferencia de grupos fosfato, metabolismo de nucleótidos o ácidos nucleicos ligados a diversos sistemas de resistencia a antibióticos. De hecho, en otras especies se ha detectado que proteínas de la familia de metalofosfoesterasas tienen funciones de modulación de la pared celular, sin embargo, aún se desconoce el mecanismo en el que ellas actúan [88]. Para este mismo ARNnc, sus blancos en *trans* codifican para proteínas vinculadas con exonucleasas reparadoras del ADN, de superficie y de anclaje de pared, y podrían estar participando en los efectos secundarios producidos por los estresores, tales como estrés oxidativo el cual podría dañar moléculas las cuales gatillan esta posible respuesta de reparación [87].

En paralelo, se analizó una posible correlación funcional entre la expresión diferencial de los ARNnc con sus respectivos genes objetivo. Para la secuencia nc3340 y el gen EF1231 (*cis*), se aprecia que para todas las condiciones en estudio existe un aumento en su abundancia. Esto podría significar que el ARN nc3340 estaría operando como un posible regulador positivo sobre el gen objetivo EF1231. Por el contrario, para el caso de la secuencia nc3683, ésta presenta un aumento de abundancia con respecto a la condición basal, mientras su gen blanco EF1362 (*cis*) presenta una expresión levemente menor frente a los estresores, lo que sugiere que el ARNnc cuestión podría regular negativamente a su gen objetivo en *cis*. Para el caso de los genes regulados en *trans*, se observa el gen EF0972 posee un aumento diferencial producto del Cu, vancomicina y ceftriaxona, situación compartida con el ARN nc3340 en el caso de Cu y vancomicina, lo que refuerza la posible acción como regulador positivo de sus genes blanco.

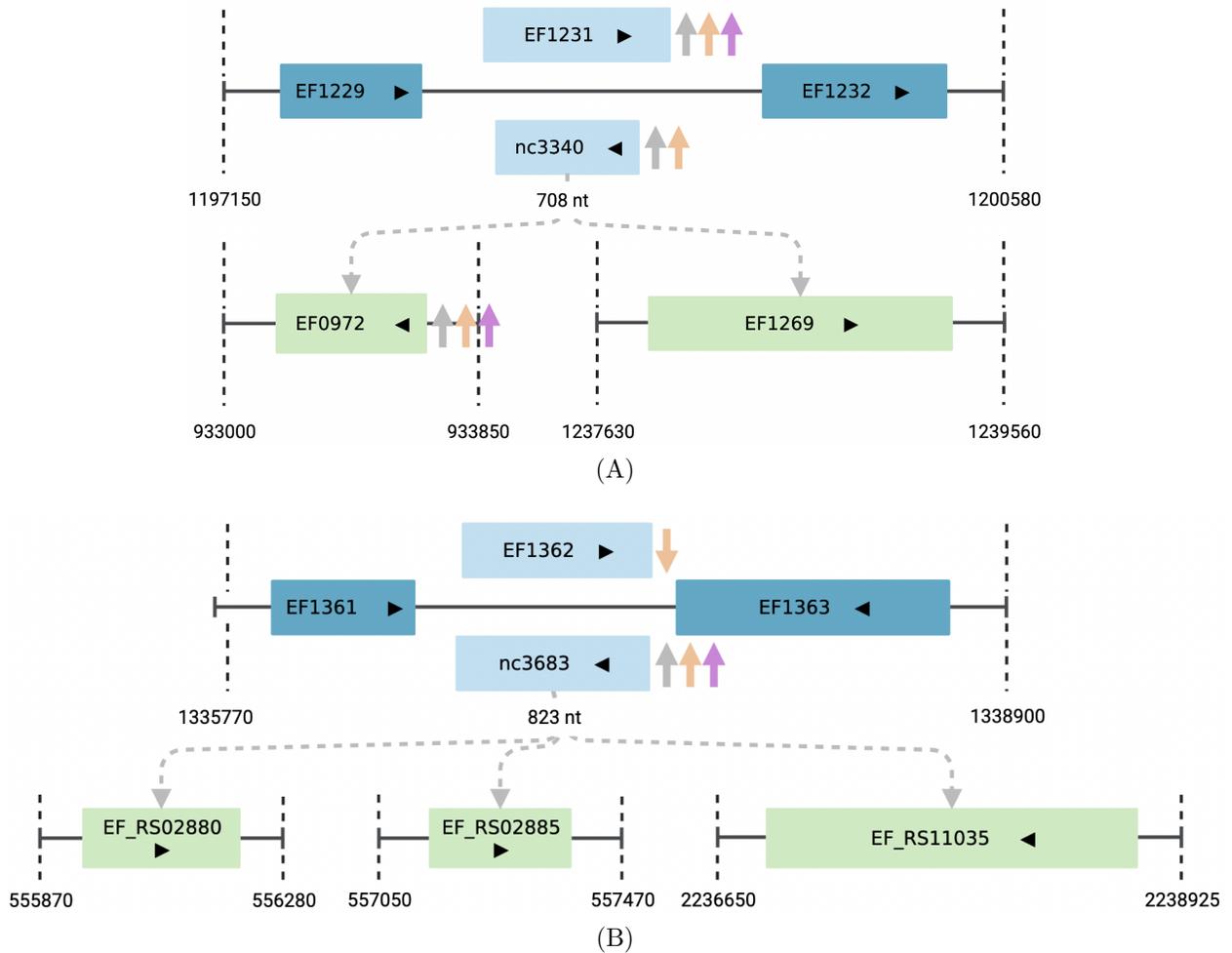


Figura 5.2: Contexto genómico de los ARNnc seleccionados. Los esquemas muestran los ARNnc y los genes locales entre ellos. En celeste se encuentran los genes blanco en *cis* con su respectivo ARNnc, en verde se destacan los genes blanco en *trans*, mientras que en azul se presentan los genes vecinos de los blancos en *cis*. Las flechas grises segmentadas indican los ARNnc regulados en *trans* y las flechas en grises, anaranjadas y rosas al lado derecho de las secuencias muestran la regulación positiva (flecha hacia arriba) o negativa (flecha hacia abajo) de los genes en los tratamientos con vancomicina, Cu y ceftriaxona, respectivamente. Los esquemas fueron creados en base al genoma publicado de *E. faecalis* V583 [87], indicando la posición del genoma por líneas segmentadas. (A) Contexto genómico para el ARN nc3340. (B) Contexto genómico para el ARN nc3683.

5.1.5. Análisis de la estructura y conservación de los ARNnc seleccionados

Con el objetivo de obtener una aproximación estructural de los ARNnc asociada a la conservación de estas moléculas en otras especies, primero se realizó una predicción de estructura secundaria de cada secuencia utilizando el servidor web RNAFold. Las estructuras fueron obtenidas en base al cálculo de la energía libre mínima óptima para cada configuración, lo que se asocia a la estabilidad de una molécula, obteniendo así un valor de -129.50 [kcal/mol]

para la estructura de nc3340, mientras que para la estructura secundaria óptima encontrada por predicción de ensamble termodinámico del ARN nc3683, se determinó que la energía libre mínima predicha corresponde a -162.90 [kcal/mol]. Las configuraciones predichas de los ARNnc se observan en las Figuras 5.3.A y 5.3.B.

Segundo, se realizó un análisis de la conservación de las secuencias de los ARNnc mediante el uso de la herramienta BLASTN, con el objetivo de estudiar si estas secuencias se encontraban conservadas en otras especies bacterianas pertenecientes al orden *Lactobacillales*. Según los resultados obtenidos (ver Tabla 5.3), los dos ARNs nc3340 y nc3683, se encontraron con una cobertura correspondiente al 100% e identidad sobre el 98.5% en especies de *Enterococcus*. En particular, se encontró que la secuencia nc3340 se conservaba parcialmente además en otros 2 organismos: en las bacterias *Lactococcus garvieae* y *Lactococcus petauri*.

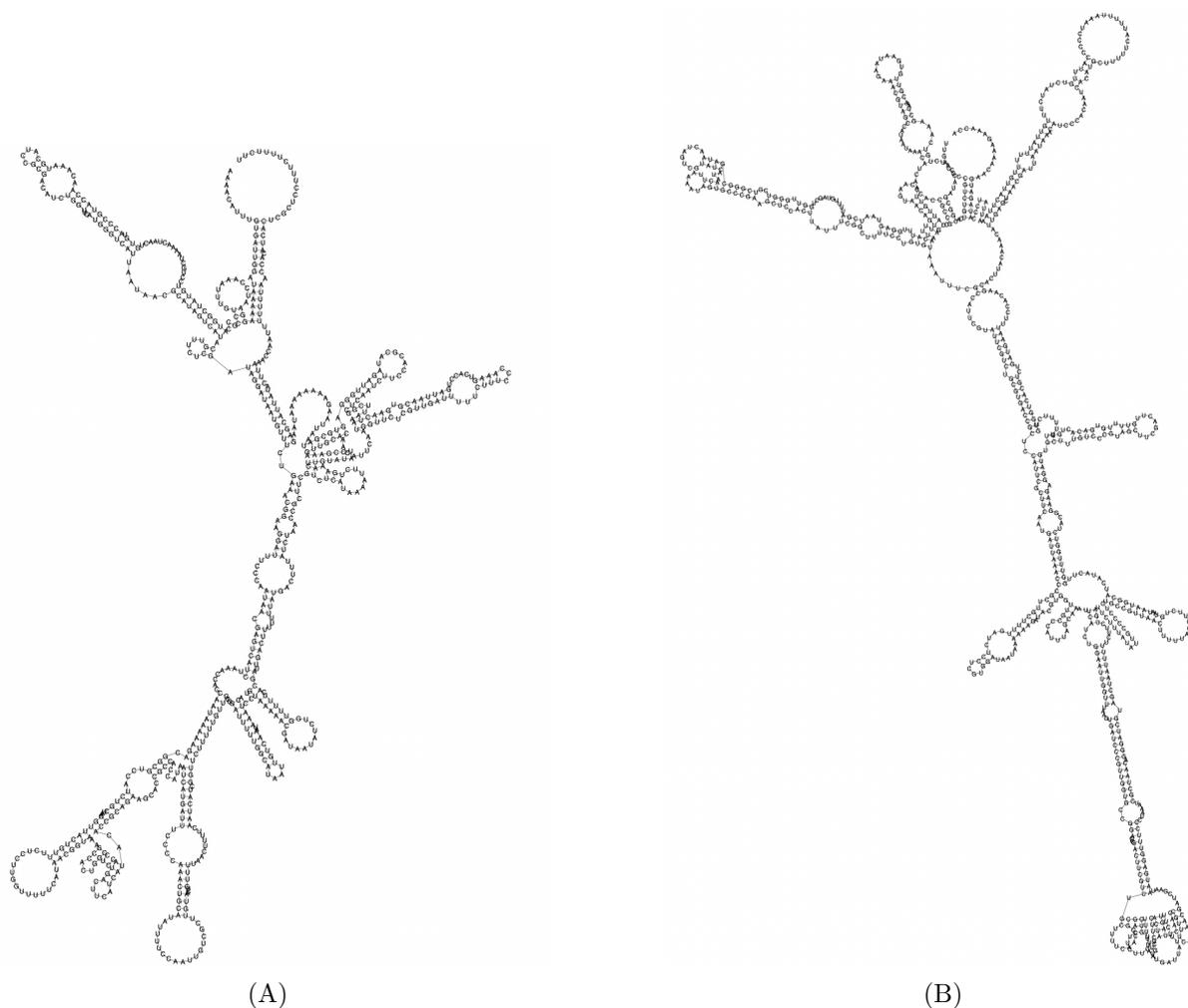


Figura 5.3: Predicción de la estructura secundaria de los ARNs no codificantes seleccionados. En la figura se observan las interacciones predichas por el programa RNAFold para cada base nitrogenada de los ARNnc seleccionados. (A) Estructura secundaria del ARN nc3340. (B) Estructura secundaria del ARN nc3683.

Tabla 5.3: Conservación de los transcritos según BLASTN.

ARNnc	Especie	Cobertura (%)	Identidad (%)	Error
nc3340	<i>Enterococcus</i> sp. PLM3	100	99.29	0.0
	<i>Enterococcus</i> sp. LX10	100	99.01	0.0
	<i>Lactococcus garvieae</i> strain JJJN1	87	74.68	1e-65
	<i>Lactococcus petauri</i> strain LG4	87	74.40	2e-62
nc3683	<i>Enterococcus</i> sp. LX10	100	98.54	0.0
	<i>Enterococcus</i> sp. PLM3	100	98.54	0.0

Tomando en consideración ambos parámetros, la Figura 5.4 muestra la predicción de la conservación del plegamiento obtenida mediante el uso del programa LocARNA [80]. Al comparar las estructuras conservadas obtenidas en ambos casos con sus respectivos plegamientos particulares (Figura 5.3), para los dos ARNnc se observan diferencias importantes en la configuración espacial de las horquillas y enlaces formados, pues básicamente presentan plegamientos distintos con sus respectivos homólogos por identidad encontrados en otras especies. Al respecto, a pesar de que las secuencias de *L. garvieae* y *L. petauri* tienen un grado de conservación a nivel de secuencia (ver Anexo A.3), la predicción del plegamiento conjunto es totalmente diferente al obtenido para las secuencias de los ARN nc3340 y nc3683, lo que sugiere que cada ARNnc en cada especie podría tener distintas funciones de regulación [89]. Si se observa en detalle la Tabla 5.3, las especies de *Enterococcus* que tienen alto porcentaje cobertura e identidad no son cepas taxonómicamente identificadas [90][91], por lo que podría tratarse de cepas de *E. faecalis* no clasificadas. De esta forma, se sostiene la premisa de que los ARNnc son poco conservados entre especies e incluso cepas de una misma especie, siendo particulares en la forma que adoptan y en la función relacionada a su estructura, a diferencia de otros reguladores del tipo protéicos tales como factores de transcripción, quienes se conservan dentro de todo el orden *Lactobacillales* y como es el caso del regulador CopY de homeostasis de Cu [89][92].

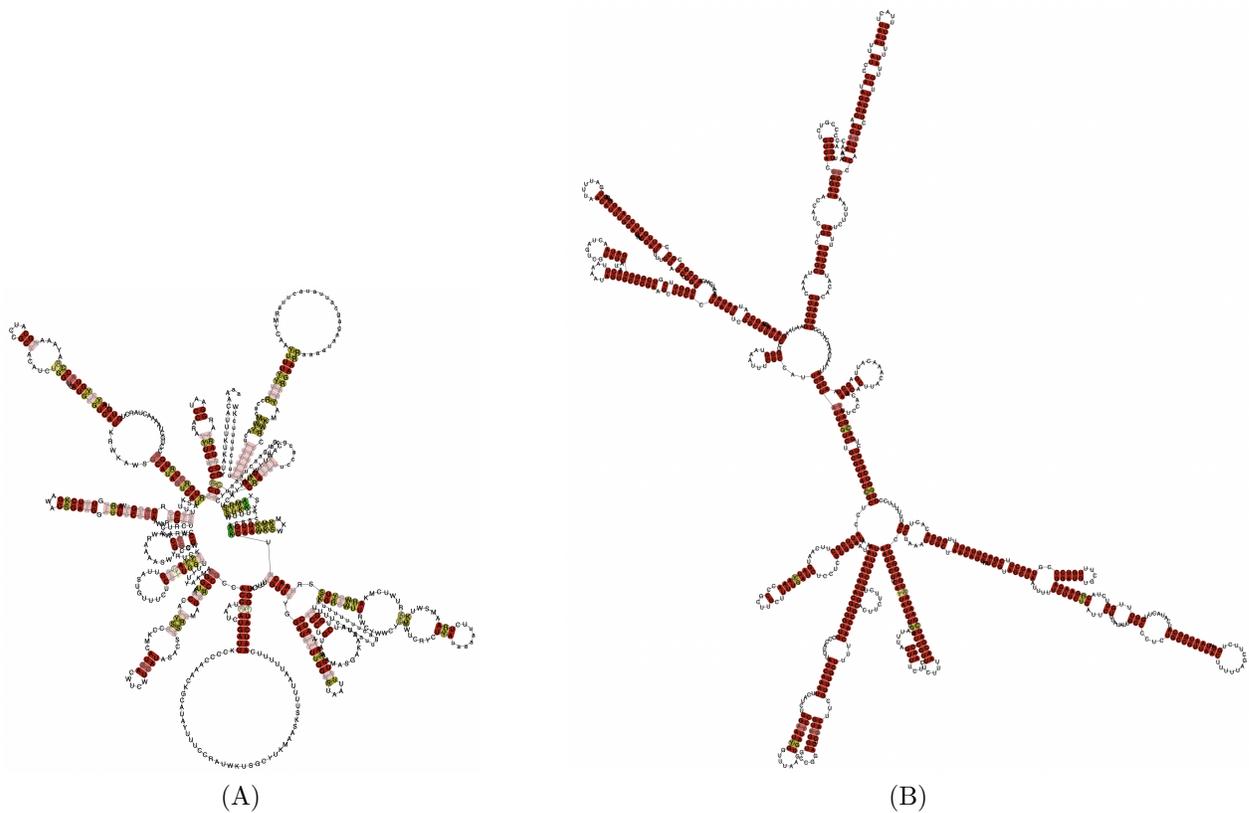


Figura 5.4: Predicción de la estructura secundaria de los ARNnc según su conservación en distintas especies. La Figura muestra las interacciones predichas por el programa LocARNA para cada base nitrogenada de los ARNnc seleccionados, considerando además la conservación de nucleótidos en las distintas especies, cuyos resultados fueron obtenidos previamente por BLASTN. (A) Estructura secundaria conservada del ARN nc3340. (B) Estructura secundaria conservada del ARN nc3683.

La determinación de las estructuras secundarias de los ARNs son de gran importancia para poder dilucidar los dominios de estas secuencias, facilitando la determinación de su función y unión a otros componentes celulares [93]. También se debe tener en consideración que la función de los ARNnc está fuertemente relacionada con su estructura secundaria más que con la secuencia primaria, y las predicciones permiten ayudar a descubrir sus funciones. Sin embargo, es importante recalcar que las predicciones exhibidas en la presente tesis fueron realizadas por métodos bioinformáticos y describen una estabilidad termodinámica calculada que no necesariamente refleja una estabilidad biológica, pues los ARNs en general pueden formar estructuras alternativas diferentes a las predichas que dependen además de la interacción con sus blancos [94][95]. Para poder determinar funciones e interacciones específicas, es necesario además realizar experimentos *in vitro* que involucren la coordinación de los transcritos con los genes objetivos identificados, lo cual complementaría y validaría las predicciones obtenidas en esta tesis.

5.2. Validación de abundancia de los ARNnc

En este punto se da comienzo a la presentación y discusión de resultados comprendidos en la segunda parte del presente trabajo, el cual apunta a la validación de los cambios de la abundancia de transcrito mediante ensayos de qPCR para los dos ARNnc en estudio durante la exposición a los diferentes estresores.

5.2.1. Determinación de las condiciones de tratamiento

Se estudió el crecimiento de *E. faecalis* OG1RF por 8 horas bajo 4 condiciones diferentes: basal, Cu, vancomicina y ceftriaxona. Cada curva se realizó por triplicado manteniendo los mismos parámetros para verificar que el crecimiento observado era similar para cada concentración de estresores y que su comportamiento se debía a la adición de Cu y antibióticos, y no a errores de manipulación o contaminación. Las condiciones de estrés poseían diferentes concentraciones de agentes estresores, como se mencionó en la sección 4.2.2.2 de Metodología. A pesar de ser una cepa diferente a la del genoma de referencia y la utilizada en los estudios de RNA-seq de vancomicina y ceftriaxona (*E. faecalis* V583), la cepa OG1RF también presenta una resistencia intrínseca a antibióticos, especialmente para la ceftriaxona [96]. Además, esta misma cepa fue la utilizada para la obtención de los datos de RNA-seq en exposición a Cu, por lo que es posible estudiar si comportamiento frente a este estresor y los antibióticos en estudio.

Los resultados de las curvas obtenidas del crecimiento en medio líquido en una placa de 96 pocillos, observados en la Figura 5.5, fueron analizados uno a uno. A partir del análisis experimental por medio de herramientas estadísticas se seleccionaron aquellas máximas concentraciones para cada condición que no presentaran diferencias significativas con respecto a la curva basal a las 3 horas, momento en el que se realiza la extracción de ARN. Para el crecimiento con tratamientos de Cu, en la Figura 5.5.A se observa que la condición basal es la curva que presenta mayor crecimiento a lo largo del tiempo, pues no presenta agentes estresores. Del mismo modo, el único caso con diferencia significativa a las 3 horas de crecimiento corresponde a la concentración de 2 mM de CuSO₄, disminuyendo el crecimiento bacteriano.

En el caso del crecimiento de *E. faecalis* en presencia de antibióticos ocurrió algo particular; el crecimiento observado en algunas de las curvas con tratamiento de las Figuras 5.5.B y 5.5.C indican que sus valores de absorbancia son mayores a las del cultivo basal. Estudios anteriores han revelado que bacterias como *E. faecalis* son capaces mostrar una resistencia temprana a estresores al estar expuestas a antibióticos, que se ve reflejada en un aumento en el crecimiento. Esta resistencia puede conducir a que nuevas generaciones puedan ser menos susceptibles al estresor al que se someten en comparación a generaciones anteriores, situación ocurre pasadas al menos 24 horas desde la inoculación del cultivo [97]. En el caso de este estudio, el ensayo de qPCR se realizó a las 3 horas, por lo que este fenómeno no interferiría

en la interpretación de los resultados.

Es así como en la Figura 5.5.B se muestra que concentraciones iguales o superiores a 2.5 $\mu\text{g}/\text{mL}$ de vancomicina disminuyen el crecimiento de la bacteria. Por su parte, concentraciones iguales o superiores a 175 $\mu\text{g}/\text{mL}$ de ceftriaxona generan el mismo efecto sobre el cultivo, tal como se observa en la Figura 5.5.C. Para este último caso, la concentración de 200 $\mu\text{g}/\text{mL}$ de ceftriaxona no generó una diferencia en el crecimiento a las 3 h de cultivo en relación a la condición basal, y si bien este resultado sugiere utilizar esta cantidad como el máximo tolerable por la bacteria, es posible que esta condición esté dentro de los márgenes muy cercanos a la toxicidad dado el resultado obtenido con 175 $\mu\text{g}/\text{mL}$, por lo cual se optó por establecer a esta última concentración como el límite inferior tóxico que afecta el crecimiento.

Luego de observar los resultados de las curvas y el análisis estadístico, se eligieron las concentraciones máximas que no afectan el crecimiento de la bacteria a las 3 horas de cultivo: 0.5 mM para Cu, 1.25 $\mu\text{g}/\text{mL}$ para vancomicina y 150 $\mu\text{g}/\text{mL}$ para ceftriaxona. En el caso del Cu (ver Figura 5.5.A), a pesar de que la concentración máxima que no presenta significancia corresponde a la de 1 mM, la concentración elegida corresponde a 0.5 mM debido a que este valor ha sido válido en investigaciones anteriores que además incluyen un análisis de disponibilidad de Cu intracelular [39].

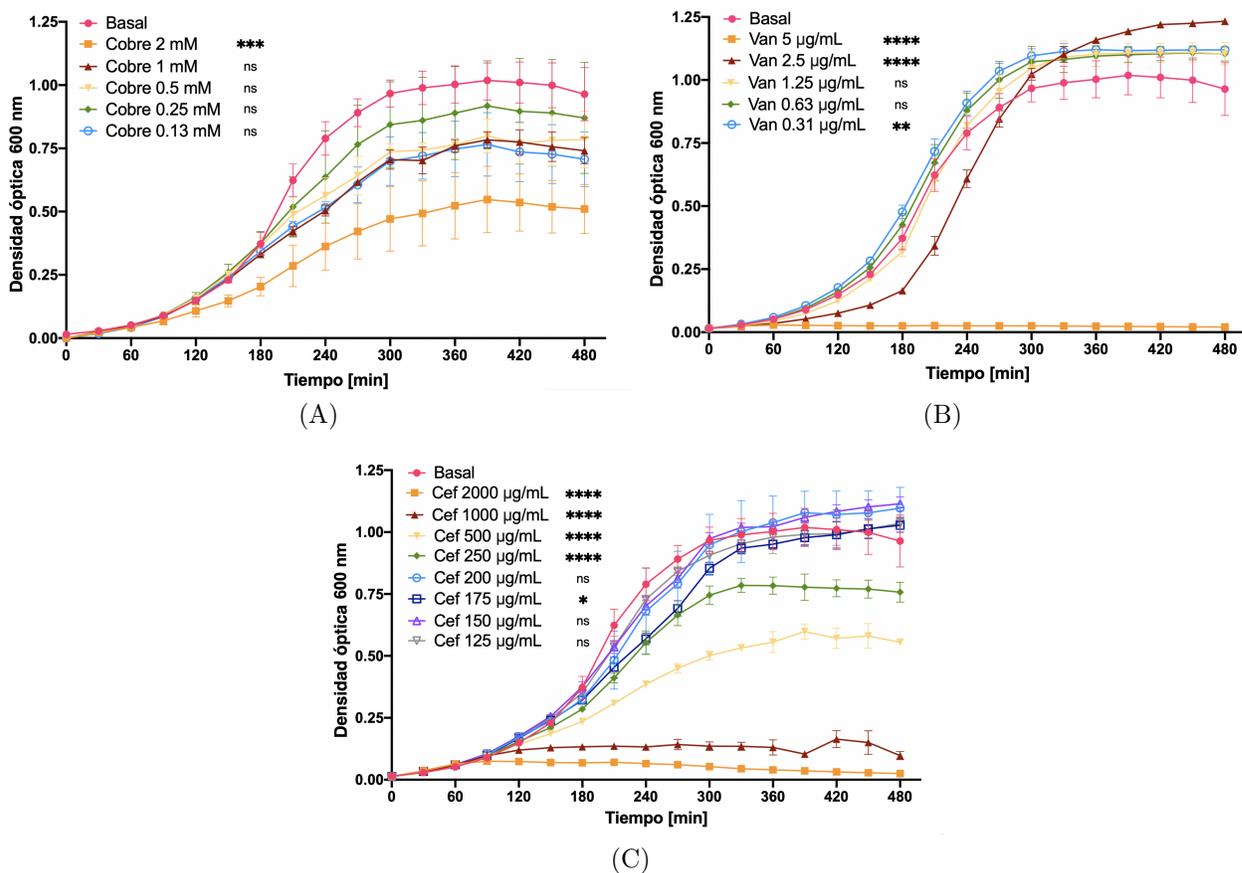


Figura 5.5: Curvas de crecimiento para cultivos de *E. faecalis* OG1RF expuestos a Cu, vancomicina y ceftriaxona. El gráfico representa las curvas de crecimiento de *E. faecalis* OG1RF crecidas en placas de 96 pocillos, expuestas a diferentes concentraciones de Cu, vancomicina y ceftriaxona. Cada punto representa el valor promedio obtenido por absorbancia (DO_{600nm}) de tres réplicas biológicas, las barras de error corresponden a la desviación estándar, mientras que los asteriscos (*) y letras 'ns' indican la significancia de cada curva con respecto a la basal. (A) Curvas con tratamientos de CuSO₄ (basal, 0.13 mM, 0.25 mM, 0.5 mM, 1 mM y 2 mM). (B) Curvas con tratamientos de vancomicina (basal, 0.31 µg/mL, 0.63 µg/mL, 1.25 µg/mL, 2.5 µg/mL y 5 µg/mL). (C) Curvas con tratamientos de ceftriaxona (basal, 125 µg/mL, 150 µg/mL, 175 µg/mL, 200 µg/mL, 250 µg/mL, 500 µg/mL, 1000 µg/mL, 2000 µg/mL).

Consolidados los resultados obtenidos en las placas de 96 pocillos, dado los experimentos posteriores, fue necesario corroborar que el crecimiento de *E. faecalis* en un volumen mayor se comportaba de manera similar, pues la oxigenación, mezclado, interacción entre bacterias, entre otros factores, es distinta [98]. Debido a que las extracciones de ARN total se debían realizar en cultivos de 15 mL tratados durante 3 horas por el estresor, se repitieron las curvas de crecimiento utilizando las mismas condiciones seleccionadas ajustadas a dicho volumen.

Si se comparan los resultados de las Figura 5.5.A, 5.5.B y 5.5.C con los de la Figura 5.6, lo primero que se puede notar es que la cantidad de bacterias multiplicadas para un

mismo tiempo es mucho mayor en los tubos de 50 mL. Sin embargo, el comportamiento de las curvas sigue el mismo patrón de crecimiento y a los 180 minutos en ambos casos el crecimiento pareciera estar en su fase exponencial. Como se muestra en la Figura 5.6, no existen diferencias significativas entre todos los tratamientos para las concentraciones estudiadas al ser comparados con el estado basal, ratificando que dichas concentraciones serían adecuadas para los experimentos posteriores.

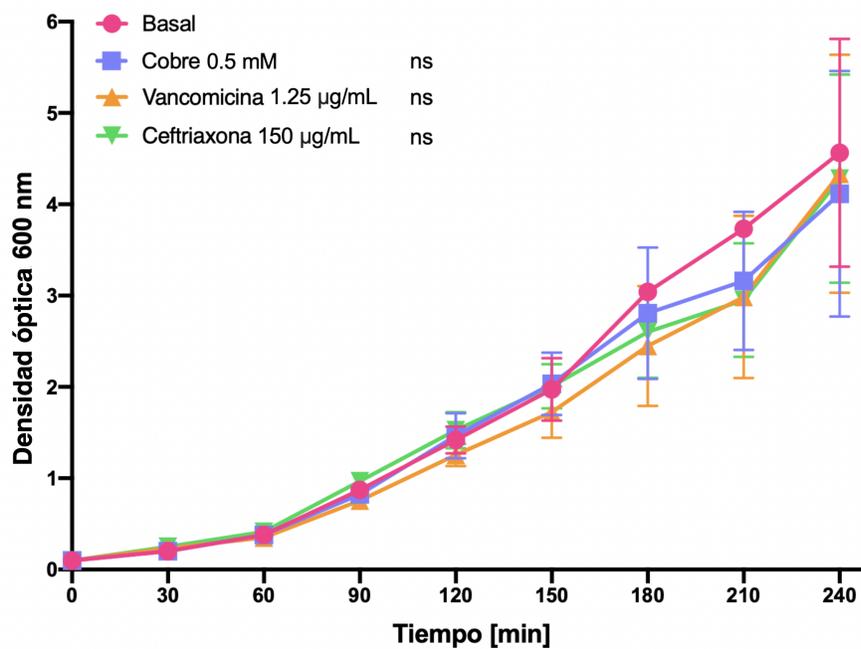


Figura 5.6: Curvas de crecimiento para cultivos de *E. faecalis* OG1RF de 15 mL expuestos a estresores. El gráfico representa la curva de crecimiento de *E. faecalis* OG1RF crecidas en tubos de centrífuga de 50 mL, expuestas a diferentes condiciones (basal, tratamientos con 0.5 mM de Cu, 1.25 µg/mL de vancomicina y 150 µg/mL de ceftriaxona). Cada punto representa el valor promedio obtenido por absorbancia (DO_{600nm}) de tres réplicas biológicas, mientras que las barras de error corresponden a la desviación estándar.

5.2.2. Extracción de ARN total y diseño de partidores

Seleccionadas las concentraciones de cada condición de crecimiento, se procedió con el crecimiento de los cultivos en tubos de centrífuga de 50 mL y a las 3 horas se realizó la extracción de ARN, según los pasos descritos en la sección de metodología. Para verificar la presencia de una concentración de ARN total en cada muestra después de la lisis celular, se tomaron alícuotas de cada muestra y se analizaron en el equipo NanoDrop. Los resultados de la concentración y calidad de estos ARNs se observan en las Tabla 5.4. En ella se muestran las razones de absorbancia obtenidas en las diferentes longitudes de onda en cada una de las muestras resultantes de la extracción de ARN de cultivos de *E. faecalis* cepa OG1RF en las condiciones de estudio (basal, tratamiento con 0.5 mM de Cu, 1.25 µg/mL de vancomicina y 150 µg/mL de ceftriaxona).

Tabla 5.4: Concentración y calidad de los ARN purificados.

Condición	Concentración [ng/ μ L]	A260/A280	A260/A230
Basal	1129.1	2.24	2.32
Cu	1067.3	2.22	1.90
Vancomicina	2168.7	2.24	2.44
Ceftriaxona	1785.3	2.24	2.31

Obtenidas las muestras con ARN concentrado, se procedió a la cuantificación de la abundancia de los ARNnc seleccionados. Para esto, los partidores utilizados en las reacciones de qPCR fueron diseñados utilizando como molde las secuencias de los ARNnc seleccionados, obtenidas a partir del genoma secuenciado de *E. faecalis* OG1RF (GenBank: CP002621.1) [99]. Teniendo dichas secuencias (detalladas en el Anexo B.1), primero se diseñaron los partidores específicos para la síntesis del ADN complementario de hebra simple (ADNc) para la retrotranscripción. No se utilizó un programa específico, pues se observó el fin de la secuencia de los ARNnc con dirección 5'-3', donde el partidor de ADNc correspondería al complementario a las últimas 20 bases, pero con dirección 3'-5' (reverso complementario). Las características de los partidores se detallan en la Tabla 5.5.

Tabla 5.5: Partidores diseñados para la síntesis del ADNc de hebra simple a partir de cada ARNnc. En la tabla se señala el nombre de cada partidor (ID), la secuencia respectiva y la temperatura de alineamiento o *melting* (Tm).

ID Partidor	Secuencia del partidor	Tm [°C]
cDNA3340	AAGAAAAGAAAGGAGCGAGT	61
cDNA3640	TGGTTTCTTTTGTTTCATCTA	56
16S_ADNc	ACCTTCCGATACGGCTA	62

Para la cuantificación de la abundancia de los transcritos, se diseñaron los partidores a partir de las secuencias de los ARNnc seleccionados, mediante el uso de la herramienta Primer3Plus. Dicha herramienta entregó múltiples opciones de partidores; la diferencia entre cada uno de ellos para un mismo ARNnc consistía en distintas temperaturas de *melting*, formación de estructuras secundarias, tamaño de producto y, por supuesto, secuencia. Se seleccionó la pareja con el puntaje mayor, basándose en las recomendaciones entregadas por el programa. Posteriormente, fue necesario verificar en el genoma completo de *E. faecalis* la especificidad de alineamiento, ya que se espera que sólo amplifiquen el transcrito de interés; este paso fue realizado utilizando BLASTN. Verificando este punto, los partidores seleccionados se detallan en la Tabla 5.6.

Tabla 5.6: Partidores diseñados para las reacciones de amplificación y cuantificación mediante qPCR. En la tabla se señala el nombre de cada partidor (ID), la secuencia respectiva y la temperatura de alineamiento o *melting* (Tm).

ID Partidor	Orientación	Secuencia	Tm [°C]	Tamaño de producto
nc3340F	Forward	GGCGTCCATCTGCTAAGGTT	60.1	107
nc3340R	Reverse	TGATTATGGTGGCGGTGCTT	60.0	
nc3683F	Forward	TTGTAGCTGGTTGGGTTCGTC	60.0	140
nc3683R	Reverse	CATTGAAGCGAATGGAGCGG	60.0	
16S_F	Forward	GTGAGGTAACGGCTCACCAA	60.0	143
16S_R	Reverse	TGCTCGGTCAGACTTTCGTC	60.0	

5.2.3. Cuantificación de los ARNnc en cultivos de *E. faecalis* expuestos a Cu, vancomicina y ceftriaxona

En la Figura 5.7 se presenta la cuantificación relativa normalizada de los 2 ARNnc en estudio (nc3340 y nc3683) en cada condición. En todos los tratamientos los datos evidencian un aumento en la abundancia de los ARNnc con respecto a la condición basal, llegando algunos a aumentar hasta 7 veces el producto con la adición del estresor.

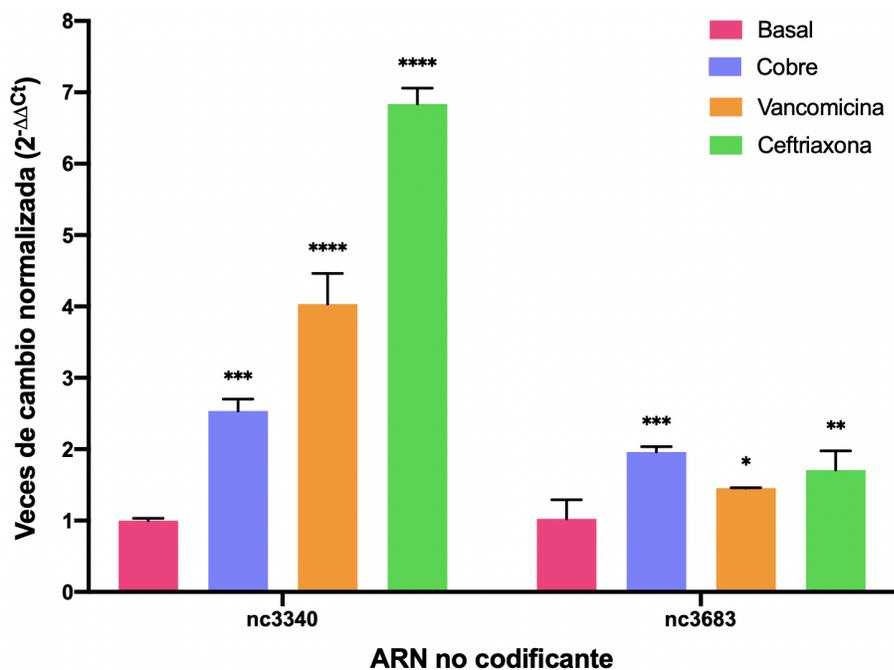


Figura 5.7: Abundancia transcripcional de los transcritos de ARNnc en estudio, obtenida por qPCR. El gráfico exhibe la razón de abundancia relativa de los ARNnc con respecto al gen normalizador 16S, para las condiciones basales y tratamiento con 0.5 mM de Cu, 1.25 $\mu\text{g}/\text{mL}$ de vancomicina y 150 $\mu\text{g}/\text{mL}$ de ceftriaxona. Los valores corresponden al promedio de tres mediciones dadas 3 réplicas técnicas; los asteriscos (*) indican el grado de significancia con respecto a condición basal.

En términos generales, dado las diferencias en los tratamientos, cepas y condiciones de crecimiento, era un esperado que la magnitud del cambio de la abundancia de los ARNnc no fueran iguales entre los experimentos de RNA-seq y los obtenidos por qPCR, no obstante, en ambos ensayos la dirección de los cambios se mantiene: en todos los casos existe un aumento de los transcritos en estudio. Atendiendo al hecho de que los experimentos de RNA-seq de antibióticos se realizaron en la cepa V583 de *E. faecalis* bajo otros tiempos y concentraciones, estos resultados suponen que al menos entre esta cepa y la OG1RF, la respuesta de los ARNnc es consistente, tal como ocurre con otros mecanismos reguladores como el regulador CopY de homeostasis de Cu, quien responde ante cualquier tiempo y concentración de exceso del metal, dando cuenta de la importancia en la respuesta por parte de este componente y cuya situación podría extrapolarse a los ARNnc seleccionados.

En particular, el ARNnc que presentó el mayor cambio de abundancia fue el nc3340, en especial al tratamiento con el antibiótico ceftriaxona, donde su expresión diferencial normalizada fue de casi 7 veces más con respecto al estado basal. Si se analiza la secuencia blanco en *cis*, esta codifica a una posible metalofosfoesterasa, con función hidrolasa para enlaces fosfodiéster. Se ha evidenciado que la resistencia a cefalosporinas como la ceftriaxona está mediada por una respuesta de sistemas reguladores de dos componentes (TCS por sus siglas en inglés) [96], donde se involucran histidinas kinasas (HK) y reguladores de respuesta (RR) que se asocian a transferencias y modificación de grupos fosfato. Tal es el caso del TCS CroRS, donde la HK fosforila al regulador de respuesta CroR que se cree que altera la transcripción a través de un dominio de unión al ADN [24]. Sin embargo, no existe evidencia que indique el ARN nc3340 o su respectivo gen objetivo estén ligados al sistema de respuesta CroRS, sugiriendo que las funciones del ARNnc y gen EF1231 pueden estar asociadas a otros procesos metabólicos. En el caso de la vancomicina ocurre algo similar, ya que existe un mecanismo de respuesta TCS llamado VanRS [100], donde el regulador de respuesta fosforilado, VanR, se une a una región promotora que activa la transcripción de los genes de resistencia [100]. De la misma manera, no hay certeza de que las secuencias en cuestión estén ligadas a dicho sistema de respuesta, no obstante abren un campo de exploración interesante en términos de poder vincular nuevas funciones a la resistencia a antibióticos.

En este sentido, un estudio realizado por Grillo-Puertas y colaboradores en 2014 sobre un análisis metagenómico de bacterias obtenidas desde el suelo, se observó en todas las muestras que el tipo de bacterias más abundantes eran bacterias solubilizadoras de fosfato (PSB), donde la mayoría de los aislados obtenidos eran además tolerantes a la sal, Cu y resistentes a la ampicilina [101]. Por otro lado, se ha demostrado una correlación en algunas bacterias entre la presencia de un alto nivel de fosfato en el medio y una mayor tolerancia al Cu [102], pues el flujo de fosfato inorgánico (Pi) estaría asociado a altos niveles de polifosfato (poliP) y, a la vez, su degradación a la presencia del metal. Teniendo esto en consideración, con respecto a la respuesta del ARN nc3340 en presencia de Cu, se puede observar en la Figura 5.7 que la abundancia relativa comparada con la condición basal fue de más de 2.5

veces. A nivel funcional, el gen blanco en *cis* codificante para una posible metalofosfoesterasa podría tener un vínculo directo con la homeostasis de Cu, ya que la presencia de este metal degradaría cadenas de poliP, y un alto nivel de Pi en el medio otorgaría una mayor tolerancia al Cu para la bacteria.

Dado que el gen objetivo en *cis* del ARN nc3683 no presenta una función específica, es complejo establecer una función de resistencia para este ARNnc. No obstante, como se presenta una expresión diferencial diferente al estado basal en la mayoría de las condiciones, no se descarta su participación en procesos metabólicos asociados a la resistencia de estos estresores. Adicionalmente, tal como se comentó con anterioridad, tanto la presencia del ARNnc como del conjunto de eventuales genes regulados que codifican para proteínas hipotéticas, abre un campo de exploración de nuevos componentes involucrados en la respuesta a Cu y antibióticos.

En resumen, en consideración a los resultados obtenidos, fue posible validar los cambios de expresión de ARNnc de respuesta a Cu y antibióticos.

Capítulo 6

Conclusiones

Los ARNnc seleccionados a partir del análisis de secuenciaciones de RNA-seq para el crecimiento de *E. faecalis* en presencia de estresores como vancomicina, ceftriaxona y/o Cu, presentan una respuesta transcripcional diferencial en comparación al estado basal, resultados obtenidos tanto en el análisis *in silico* como en los ensayos *in vitro*. Estas secuencias podrían tener genes objetivo con regulación en *cis* y *trans*, cuya complementariedad de interacción específica entre las moléculas de ARN requiere de ensayos complementarios que validen la regulación.

La conservación de los ARNnc en otras especies no presentó porcentajes de identidad importantes con otras especies, sugiriendo que las secuencias en estudio sólo estarían codificadas para cepas de *E. faecalis*, reafirmando resultados anteriores en otras bacterias que declaran la baja conservación de estos elementos, a diferencia de otros reguladores del tipo protéicos tales como factores de transcripción.

Si bien los genes blanco codifican para proteínas no directamente relacionadas con resistencia, los ensayos de expresión dan cuenta que tanto el ARNnc regulador como sus blancos responden transcripcionalmente a más de un estresor, evidenciando que la regulación transcripcional de *E. faecalis* es compleja, la cual requiere de una mayor profundidad en términos de interpretación de otros componentes. En esta línea, algunos de los blancos identificados codifican para proteínas hipotéticas débilmente caracterizadas, las que en algunos casos presentaron una expresión diferencial frente a estresores en los datos de RNA-seq, abriendo un espacio de exploración para el campo de resistencia a Cu y antibióticos.

La estrategia utilizada en esta tesis en base a herramientas bioinformáticas y posterior validación a través de ensayos de expresión por qPCR fue exitosa en la identificación de ARNnc de respuesta a Cu y antibióticos, método que es extrapolable para el estudio de otros estímulos.

6.1. Proyecciones

Una forma de estudiar el efecto que los ARNnc identificados y estudiados tienen sobre la resistencia de antimicrobianos en *E. faecalis*, es observando el comportamiento de la bacteria en ausencia de estos elementos, como por ejemplo diseñando cepas mutantes para cada uno de los ARNnc. Del mismo modo, podrían utilizarse vectores de expresión para aumentar la abundancia de estos ARNnc dentro de la célula, y con esto verificar si existen cambios tanto en la abundancia de sus posibles blancos como también sobre el fenotipo de resistencia en la bacteria expuesta a los diferentes estresores.

Por otro lado, conocer la función específica de cada una de las proteínas codificadas por los genes blanco también permitiría tener una aproximación de si los ARNnc seleccionados tienen influencia sobre los mecanismos de respuesta de la bacteria en presencia de antimicrobianos, lo que entregaría nuevos conocimientos en el campo de componente involucrados en la resistencia bacteriana.

Desde el punto de vista de ciencia aplicada, la caracterización de ARNnc podría permitir que estas secuencias sean blancos para fármacos específicos contra cepas de *E. faecalis* dada la baja conservación que tienen en otras especies, convirtiéndose en una estrategia para el desarrollo de tecnologías de control de patógenos con un alto nivel de especificidad. En este contexto, el estudio de las secuencias no codificantes sigue siendo un desafío dada la complejidad de identificación y caracterización, lo cual conduce al desarrollo de nuevas líneas de investigación vinculadas a la regulación transcripcional en especies bacterianas.

Abreviaturas

- **ABC**: Dependiente de ATP (*ATP-Binding Cassette*)
- **ADN**: Ácido desoxirribonucleico
- **ADNc**: Ácido desoxirribonucleico complementario
- **ARGs**: Genes de resistencia a antibióticos (*antibiotic resistant genes*)
- **ARN**: Ácido ribonucleico
- **ARNnc**: ARN no codificantes
- **BLAST**: *Basic Local Alignment Search Tool*
- **BLASTN**: BLAST de nucleótidos (*nucleotide BLAST*)
- **BLASTP**: BLAST de proteínas (*protein BLAST*)
- **Cef**: Ceftriaxona
- **CopA**: ATPasa transportadora de eflujo de Cu
- **copA**: Gen que codifica para la proteína CopA
- **CopY**: Regulador transcripcional - Represor operón *cop*
- **copY**: Gen que codifica para la proteína CopY
- **CopZ**: Chaperona de Cu
- **copZ**: Gen que codifica para la proteína CopZ
- **Ct**: Umbral de ciclos (*Cycle threshold*)
- **Cu**: Cobre
- **CuSO₄**: Sulfato de cobre
- **dNTP**: Desoxinucleótidos trifosfatos
- **DO**: Densidad óptica

- **EDTA**: Ácido etilendiamino tetra acético
- **HK**: Histidina kinasa (*histidine kinase*)
- **IntaRNA**: *Interacting RNAs*
- **LB-UOH**: Laboratorio de Bioingeniería de la Universidad de O'Higgins
- **Log₂**: Logaritmo en base 2
- **MRGs**: Genes de resistencia a metales (*metal resistant genes*)
- **NaCl**: Cloruro de sodio
- **Na₂HPO₄**: Fosfato sódico
- **NAG**: N-acetilglucosamina
- **NAM**: Ácido N-acetilmurámico
- **NCBI**: *National Center of Biotechnology Information*
- **PBP**: Proteína de unión a penicilina (*penicilin binding protein*)
- **qPCR**: PCR cuantitativa (*quantitative Polimerase Chain Reaction*)
- **RPM**: Revoluciones por minuto
- **RR**: Regulador de respuesta (*response regulator*)
- **RT**: Transcripción inversa o retrotranscripción
- **SRA**: *Sequence Read Archive*
- **TCS**: Sistema de dos componentes (*two-component system*)
- **TXRF**: Fluorescencia de rayos X por reflexión total (*Total X-Ray Reflection Fluorescence*)
- **UTR**: Region no traducida
- **Van**: Vancomicina
- **VRE**: Enterococos resistentes a vancomicina (*vancomycin-resistant enterococci*)

Bibliografía

- [1] Centers for Disease Control and Prevention, “CDC/NHSN Surveillance Definitions for Specific Types of Infections”, National Healthcare Safety Network, 2022, [En línea]. Disponible en: https://www.cdc.gov/nhsn/pdfs/pscmanual/17pscnosinfdef_current.pdf. [Acceso: 25 de diciembre de 2021].
- [2] Ministerio de Salud, República de Chile, “Manual de prevención y control de las infecciones intrahospitalarias (IIH) y normas del programa nacional de IIH”, División de Programas de Salud, Departamento de Epidemiología, 1993, [En línea]. Disponible en: <https://www.minsal.cl/portal/url/item/918f15de4a29106ae04001011e016773.pdf>. [Acceso: 25 de diciembre de 2021].
- [3] Ministerio de Salud, República de Chile, “Normas de aislamiento y manual de procedimientos”, 1988, [En línea]. Disponible en: <https://www.minsal.cl/portal/url/item/87957ff07b7cb1bee04001011f0135da.pdf>. [Acceso: 25 de diciembre de 2021].
- [4] Ministerio de Salud, Gobierno de Chile, “Modifica norma técnica N° 175 sobre Vigilancia nacional de resistencia a los antimicrobianos en agentes que pueden producir infecciones asociadas”, Gabinete Ministerial, División Jurídica, 2019, [En línea]. Disponible en: <https://www.minsal.cl/wp-content/uploads/2019/03/Exento-16-de-2019-Modifica-NT-175-Vigilancia-Nacional-de-Resistencia-a-los-Antimicrobianos-en-Agentes-que-pueden-producir-IAAS.pdf>. [Acceso: 25 de diciembre de 2021].
- [5] Unidad de IAAS del Ministerio de Salud, Gobierno de Chile, “Informe de Vigilancia de Infecciones Asociadas a la Atención en Salud 2019”, Departamento de Calidad y Seguridad de la Atención. Programa Control de IAAS, 2019, [En línea]. Disponible en: https://www.minsal.cl/wp-content/uploads/2021/08/Informe_2019_Ord.pdf. [Acceso: 25 de diciembre de 2021].
- [6] Centers for Disease Control and Prevention, “Antibiotic resistance threats in the United States”, U.S. Department of Health and Human Services, CDC, 2019, [En línea]. Disponible en: <https://www.cdc.gov/drugresistance/pdf/threats-report/2019-ar-threats-report-508.pdf>. [Acceso: 25 de diciembre de 2021].

- [7] J. Silva, L. Asserella, N. Bolados, N. Herrera, J. Leyton, “Resistencia a antimicrobianos en cepas de *Enterococcus* sp aisladas en hospitales del norte de Chile”, *Revista chilena de infectología : organo oficial de la Sociedad Chilena de Infectología*, vol. 23, n°3, pp. 226–231, 2006, [En línea]. Disponible en: <https://doi.org/10.4067/s0716-10182006000300005>. [Acceso: 07 de enero de 2023].
- [8] M. Castellano-González, A. Perozo-Mena, K. Gutierrez, J. Jimenez, M. Urdaneta., “Distribución de especies y susceptibilidad antimicrobiana en cepas de *Enterococcus* de origen clínico”, *Kasmera*, vol. 46, n°2, 2018, [En línea]. Disponible en: <https://produccioncientificaluz.org/index.php/kasmera/article/view/24663>. [Acceso: 27 de agosto de 2022].
- [9] A. Pourmand, M. Mazer-Amirshahi, G. Jasani, L. May, “Emerging trends in antibiotic resistance: implications for emergency medicine”, *American Journal of Emergency Medicine*, vol. 35, n°8, pp. 1172-1176, 2017, [En línea]. Disponible en: <https://doi.org/10.1016/j.ajem.2017.03.010>. [Acceso: 25 de diciembre de 2021].
- [10] I. Paulsen, *et al.*, “Role of Mobile DNA in the Evolution of Vancomycin-Resistant *Enterococcus faecalis*”, *Science*, vol. 299, n°5615, pp. 2071-2074, 2003, [En línea]. Disponible en: <https://www.science.org/doi/abs/10.1126/science.1080613>. [Acceso: 31 de octubre de 2022].
- [11] I. Levin-Reisman, I. Ronin, O. Gefen, I. Braniss, N. Shores, N. Balaban, “Antibiotic tolerance facilitates the evolution of resistance”, *Science*, vol. 355, n°6327, pp. 826–830, 2017, [En línea]. Disponible en: <https://doi.org/10.1126/science.aaj2191>. [Acceso: 31 de octubre de 2022].
- [12] M. Huycke, D. Sahn, M. Gilmore, “Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future”, *Emerging infectious diseases*, vol. 4, n°2, pp. 239–249, 1998, [En línea]. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/9621194/>. [Acceso: 31 de octubre de 2022].
- [13] J. Arredondo, A. Echeguren, P. Arzate, J. Medina, “Susceptibilidad antimicrobiana de *Enterococcus faecalis* y *faecium* en un hospital de tercer nivel”, *Revista Latinoamericana de Infectología Pediátrica*, vol. 31, n°2, pp. 56-61, 2018, [En línea]. Disponible en: <https://www.medigraphic.com/pdfs/infectologia/lip-2018/lip182d.pdf>. [Acceso: 02 de octubre de 2021].
- [14] S. Sood, M. Malhotra, BK. Das, A. Kapil, “Enterococcal infections & antimicrobial resistance”, *Indian Journal of Medical Research*, vol. 128, n°2, pp. 111-121, 2008, [En línea]. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/19001673>. [Acceso: 02 de octubre de 2021].
- [15] C. Arias, B. Murray, “The rise of the *Enterococcus*: beyond vancomycin resistance”, *Nature Reviews Microbiology*, vol. 10, pp. 266–278, 2012, [En línea]. Disponible en: <https://doi.org/10.1038/nrmicro2761>. [Acceso: 12 de octubre de 2022].

- [16] B. Murray, “The life and times of the *Enterococcus*”, *Clinical Microbiology Reviews*, vol. 3, n°1, pp. 46-65, 1990, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/CMR.3.1.46>. [Acceso: 12 de octubre de 2022].
- [17] B. Hollenbeck, L. Rice, “Intrinsic and acquired resistance mechanisms in *Enterococcus*”, *Virulence*, vol. 3, n°5, pp. 421-569, 2012, [En línea]. Disponible en: <https://doi.org/10.4161/viru.21282>. [Acceso: 12 de octubre de 2022].
- [18] M. Ahmed, K. Baptiste, “Vancomycin-Resistant Enterococci: A Review of Antimicrobial Resistance Mechanisms and Perspectives of Human and Animal Health”, *Microbial Drug Resistance*, vol. 24, n°5, pp. 590-606, 2017, [En línea]. Disponible en: <https://doi.org/10.1089/mdr.2017.0147>. [Acceso: 25 de diciembre de 2021].
- [19] E. Cercenado, “*Enterococcus*: resistencias fenotípicas y genotípicas y epidemiología en España”, *Enfermedades Infecciosas y Microbiología Clínica*, vol. 29, n°5, pp. 59-65, 2011, [En línea]. Disponible en: [https://doi.org/10.1016/S0213-005X\(11\)70045-3](https://doi.org/10.1016/S0213-005X(11)70045-3). [Acceso: 25 de diciembre de 2021].
- [20] M. Faron, N. Ledebøer, B. Buchan, C. Kraft, “Resistance Mechanisms, Epidemiology, and Approaches to Screening for Vancomycin-Resistant *Enterococcus* in the Health Care Setting”, *Journal of Clinical Microbiology*, vol. 54, n°10, pp. 2436-2447, 2016, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/JCM.00211-16>. [Acceso: 26 de diciembre de 2021].
- [21] J. Munita, C. Arias, “Mechanisms of Antibiotic Resistance”, *Microbiology Spectrum*, vol. 4, n°2, pp. 4.2.15, 2016, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/microbiolspec.VMBF-0016-2015>. [Acceso: 22 de diciembre de 2021].
- [22] T. Lemke, D. Williams, V. Roche, W. Zito, “Foye’s Principles of Medicinal Chemistry”, *Lippincott Williams Wilkins*, 7ma ed., pp. 1093–1094, 1099–1100, 2013.
- [23] R. Darnell, M. Knottenbelt, F. Todd Rose, I. Monk, T. Stinear, G. Cook, “Genomewide Profiling of the *Enterococcus faecalis* Transcriptional Response to Teixobactin Reveals CroRS as an Essential Regulator of Antimicrobial Tolerance”, *mSphere*, vol. 4, n°3, pp. e00228-19, 2019, [En línea]. Disponible en: <https://journals.asm.org/doi/10.1128/mSphere.00228-19>. [Acceso: 06 de septiembre de 2021].
- [24] Y. Comenge, R. Quintiliani Jr, L. Li, L. Dubost, J. Brouard, J. Hugonnet, M. Arthur, “The CroRS two-component regulatory system is required for intrinsic beta-lactam resistance in *Enterococcus faecalis*”, *Journal of bacteriology*, vol. 185, n°24, pp. 7184–7192, 2003, [En línea]. Disponible en: <https://doi.org/10.1128/JB.185.24.7184-7192.2003>. [Acceso: 31 de octubre de 2022].

- [25] C. Kristich, J. Little, C. Hall, J. Hoff, “Reciprocal Regulation of Cephalosporin Resistance in *Enterococcus faecalis*”, *mBio*, vol. 2, n°6, pp. e00199-11, 2011, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/mBio.00199-11>. [Acceso: 20 de noviembre de 2022].
- [26] S. Ono, T. Muratani, T. Matsumoto, “Mechanisms of Resistance to Imipenem and Ampicillin in *Enterococcus faecalis*”, *Antimicrobial Agents and Chemotherapy*, vol. 49, n°7, pp. 2954-2958, 2005, [En línea]. Disponible en: <https://doi.org/10.1128/AAC.49.7.2954-2958.2005>. [Acceso: 25 de diciembre de 2021].
- [27] W. Miller, J. Munita, C. Arias, “Mechanisms of antibiotic resistance in Enterococci”, *Expert Review of Anti-infective Therapy*, vol. 12, n°10, pp. 1221-1236, 2014, [En línea]. Disponible en: <https://doi.org/10.1586/14787210.2014.956092>. [Acceso: 25 de diciembre de 2021].
- [28] C. Seneviratne, T. Suriyanarayanan, S. Swarup, K. Burton, N. Nagarajan, C. Zhang, “Transcriptomics Analysis Reveals Putative Genes Involved in Biofilm Formation and Biofilm-associated Drug Resistance of *Enterococcus faecalis*”, *Journal of Endodontics*, vol. 46, n°3, pp. 949-955, 2017, [En línea]. Disponible en: [https://www.jendodon.com/article/S0099-2399\(17\)30069-9/fulltext](https://www.jendodon.com/article/S0099-2399(17)30069-9/fulltext). [Acceso: 06 de octubre d 2021].
- [29] C. Michaux, C. Martini, A. Hanin, Y. Auffray, A. Hartke, JC. Giard, “SlyA regulator is involved in bile salts stress response of *Enterococcus faecalis*”, *FEMS Microbiology Letters*, vol. 324, n°2, pp. 142–146, 2011, [En línea]. Disponible en: <https://doi.org/10.1111/j.1574-6968.2011.02390.x>. [Acceso: 14 de octubre de 2022].
- [30] C. Michaux, M. Sanguinetti, F. Reffuveille, Y. Auffray, B. Posteraro, M. Gilmore, A. Hartke, JC. Giard, “SlyA is a Transcriptional Regulator Involved in the Virulence of *Enterococcus faecalis*”, *Infection and Immunity*, vol. 79, n°7, pp. 2638-2645, 2011, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/IAI.01132-10>. [Acceso: 14 de octubre de 2022].
- [31] E. Riboulet-Bisson, A. Le Jeune, A. Benachour, Y. Auffray, A. Hartke, JC. Giard, “Ers a Crp/Fnr-like transcriptional regulator of *Enterococcus faecalis*”, *International Journal of Food Microbiology*, vol. 131, n°1, pp. 71-74, 2009, [En línea]. Disponible en: <https://doi.org/10.1016/j.ijfoodmicro.2008.06.022>. [Acceso: 14 de octubre de 2022].
- [32] N. Verneuil, A. Rincé, M. Sanguinetti, Y. Auffray, A. Hartke, JC. Giard, “Implication of *hypR* in the virulence and oxidative stress response of *Enterococcus faecalis*”, *FEMS Microbiology Letters*, vol. 252, n°1, pp. 137-141, 2005, [En línea]. Disponible en: <https://doi.org/10.1016/j.femsle.2005.08.043>. [Acceso: 14 de octubre de 2022].
- [33] N. Verneuil1, A. Rincé, M. Sanguinetti, B. Posteraro, G. Fadda, Y. Auffray, A. Hartke, JC. Giard, “Contribution of a PerR-like regulator to the oxidative-stress response and virulence of *Enterococcus faecalis*”, *Microbiology*, vol. 151, n°12, pp. 3997–4004, 2005,

[En línea]. Disponible en: <https://doi.org/10.1128/IAI.72.8.4424-4431.2004>. [Acceso: 14 de octubre de 2022].

- [34] N. Verneuil, M. Sanguinetti, Y. Le Breton, B. Posteraro, G. Fadda, Y. Auffray, A. Hartke, J.C. Giard, “Effects of the *Enterococcus faecalis* *hypR* Gene Encoding a New Transcriptional Regulator on Oxidative Stress Response and Intracellular Survival within Macrophages”,.
- [35] L. Stewart, D. Thaqi, B. Kobe, A. McEwan, K. Waldron, K. Djoko, “Handling of nutrient copper in the bacterial envelope”, *Metallomics*, vol. 11, n°1, pp. 50-63, 2019, [En línea]. Disponible en: <https://doi.org/10.1039/c8mt00218e>. [Acceso: 26 de diciembre de 2021].
- [36] S. Wanes, C. Keevil, “Mechanism of Copper Surface Toxicity in Vancomycin-Resistant Enterococci following Wet or Dry Surface Contact”, *Applied and Environmental Microbiology*, vol. 77, n°17, pp. 6049-6059, 2011, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/AEM.00597-11>. [Acceso: 03 de octubre de 2021].
- [37] International Copper Association, Codelco, “Antimicrobial Copper: new market opportunities”, [En línea]. Disponible en: https://www.codelco.com/prontus_codelco/site/artic/20111227/asocfile/201112-27161526/codelco_antimicrobial_copper_shanghai_28jul2010.pdf. [Acceso: 02 de octubre de 2021].
- [38] R. Portmann, K. Poulsen, R. Wimmer, *et al.*, “CopY-like Copper Inducible Repressors are Putative ‘Winged Helix’ Proteins”, *Biometals*, vol. 19, pp. 61-70, 2006, [En línea]. Disponible en: <https://doi.org/10.1007/s10534-005-5381-3>. [Acceso: 26 de octubre de 2021].
- [39] A. Reyes-Jara, M. Latorre, G. López, A. Bourgogne, B. Murray, V. Cambiazo, M. González, “Genome-wide transcriptome analysis of the adaptative response of *Enterococcus faecalis* to copper exposure”, *Biometals*, vol. 23, pp. 1105-1113, 2010, [En línea]. Disponible en: <https://doi.org/10.1007/s10534-010-9356-7>. [Acceso: 09 de septiembre de 2021].
- [40] E. Silveira, A. Freitas, P. Antunes, M. Barros, J. Campos, T. Coque, L. Peixe, C. Novais, “Co-transfer of resistance to high concentrations of copper and first-line antibiotics among *Enterococcus* from different origins (humans, animals, the environment and foods) and clonal lineages”, *Journal of Antimicrobial Chemotherapy*, vol. 69, n°4, pp. 899-906, 2014, [En línea]. Disponible en: <https://academic.oup.com/jac/article/69/4/899/705795>. [Acceso: 03 de octubre de 2021].
- [41] S. Zhang, D. Wang, Y. Wang, H. Hasman, F. Aarestrup, H. Alwathnani, Y. Zhu, C. Rensing, “Genome sequences of copper resistant and sensitive *Enterococcus faecalis*

- strains isolated from copper-fed pigs in Denmark”, *Stand in Genomic Science*, vol. 10, n°35, 2015, [En línea]. Disponible en: <https://doi.org/10.1186/s40793-015-0021-1>. [Acceso: 04 de octubre de 2021].
- [42] H. Hasman, I. Kempf, B. Chidaine, R. Cariolet, A. Ersbøll, H. Houe, H. Bruun Hansen, F. Aarestrup, “Copper resistance in *Enterococcus faecium*, mediated by the *tcrB* gene, is selected by supplementation of pig feed with copper sulfate”, *Applied and Environmental Microbiology*, vol. 72, n°9, pp. 5784-5789, 2006, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/AEM.02979-05>. [Acceso: 04 de octubre de 2021].
- [43] X. Zou, M. Weng, X. Ji, R. Guo, W. Zheng, W. Yao, “Comparison of antibiotic resistance and copper tolerance of *Enterococcus* spp. and *Lactobacillus* spp. isolated from piglets before and after weaning”, *Journal of Microbiology*, vol. 55, pp. 703–710, 2017, [En línea]. Disponible en: <https://doi.org/10.1007/s12275-017-6241-x>. [Acceso: 09 de enero de 2023].
- [44] J. Chen, J. Li, H. Zhang, W. Shi, Y. Liu, “Bacterial Heavy-Metal and Antibiotic Resistance Genes in a Copper Tailing Dam Area in Northern China, journal = *frontiers in microbiology*”, vol. 10, pp. 1916, 2019, [En línea]. Disponible en: <https://doi.org/10.3389/fmicb.2019.01916>. [Acceso: 20 de enero de 2023].
- [45] K. Poole, “At the Nexus of Antibiotics and Metals: The Impact of Cu and Zn on Antibiotic Activity and Resistance”, *Trends in Microbiology*, vol. 25, n°10, pp. 820-832, 2017, [En línea]. Disponible en: <https://doi.org/10.1016/j.tim.2017.04.010>. [Acceso: 20 de enero de 2023].
- [46] L. Waters, G. Storz, “Regulatory RNAs in bacteria”, *Cell*, vol. 136, n°4, pp. 615-628, 2009, [En línea]. Disponible en: <https://doi.org/10.1016/j.cell.2009.01.043>. [Acceso: 02 de octubre de 2022].
- [47] G. Storz, J Opdyke, A. Zhang, “Controlling mRNA stability and translation with small, noncoding RNAs”, *Current Opinion in Microbiology*, vol. 7, n°2, pp. 140-144, 2004, [En línea]. Disponible en: <https://doi.org/10.1016/j.mib.2004.02.015>. [Acceso: 02 de octubre de 2022].
- [48] G. Storz, “An Expanding Universe of Non-coding RNAs”, *Science*, vol. 296, n°5571, pp. 1260-1263, 2002, [En línea]. Disponible en: <https://www.science.org/lookup/doi/10.1126/science.1072249>. [Acceso: 01 de octubre de 2021].
- [49] S. Brantl, R. Brückner, “Small regulatory RNAs from low-GC Gram-positive bacteria”, *RNA Biology*, vol. 11, n°5, pp. 443-456, 2014, [En línea]. Disponible en: <https://doi.org/10.4161/rna.28036>. [Acceso: 28 de Agosto de 2022].
- [50] M. Jørgensen, J. Pettersen, B. Kallipolitis, “sRNA-mediated control in bacteria: An

- increasing diversity of regulatory mechanisms”, *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, vol. 1863, n°5, pp. 194504, 2020, [En línea]. Disponible en: <https://doi.org/10.1016/j.bbagr.2020.194504>. [Acceso: 28 de agosto de 2022].
- [51] E. Massé, H. Salvail, G. Desnoyers, M. Arguin, “Small RNAs controlling iron metabolism”, *Current Opinion in Microbiology*, vol. 10, n°2, pp. 140-145, 2007, [En línea]. Disponible en: <https://doi.org/10.1016/j.mib.2007.03.013>. [Acceso: 31 de octubre de 2022].
- [52] C. Vanderpool, “Physiological consequences of small RNA-mediated regulation of glucose-phosphate stress”, *Current Opinion in Microbiology*, vol. 10, n°2, pp. 146-151, 2007, [En línea]. Disponible en: <https://doi.org/10.1016/j.mib.2007.03.011>. [Acceso: 31 de octubre de 2022].
- [53] S. Altuvia, D. Weinstein-Fischer, A. Zhang, L. Postow, G. Storz, “A Small, Stable RNA Induced by Oxidative Stress: Role as a Pleiotropic Regulator and Antimutator”, *Cell*, vol. 90, n°1, pp. 43-53, 1997, [En línea]. Disponible en: [https://doi.org/10.1016/S0092-8674\(00\)80312-8](https://doi.org/10.1016/S0092-8674(00)80312-8). [Acceso: 31 de octubre de 2022].
- [54] J. Vogel, K. Papenfort, “Small non-coding RNAs and the bacterial outer membrane”, *Current Opinion in Microbiology*, vol. 9, n°6, pp. 605-611, 2006, [En línea]. Disponible en: <https://doi.org/10.1016/j.mib.2006.10.006>. [Acceso: 31 de octubre de 2022].
- [55] P. Valentin-Hansen, J. Johansen, A. Rasmussen, “Small RNAs controlling outer membrane porins”, *Current Opinion in Microbiology*, vol. 10, n°2, pp. 152-155, 2007, [En línea]. Disponible en: <https://doi.org/10.1016/j.mib.2007.03.001>. [Acceso: 31 de octubre de 2022].
- [56] A. Toledo-Arana, F. Repoila, P. Cossart, “Small noncoding RNAs controlling pathogenesis”, *Current Opinion in Microbiology*, vol. 10, n°2, pp. 182-188, 2007, [En línea]. Disponible en: <https://doi.org/10.1016/j.mib.2007.03.004>. [Acceso: 31 de octubre de 2022].
- [57] X. Zhang, W. Wang, W. Zhu, J. Dong, Y. Cheng, Z. Yin, F. Shen, “Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels”, *International Journal of Molecular Sciences*, vol. 20, n°22, pp. 5573, 2019, [En línea]. Disponible en: <http://dx.doi.org/10.3390/ijms20225573>. [Acceso: 21 de noviembre de 2022].
- [58] K. Harris, R. Breaker, “Large Noncoding RNAs in Bacteria”, *Microbiology spectrum*, vol. 6, n°4, PP. 6.4.01, 2018, [En línea]. Disponible en: <https://doi.org/10.1128/microbiolspec.RWR-0005-2017>. [Acceso: 15 de diciembre de 2022].
- [59] S. Chareyre, P. Mandin, “Bacterial Iron Homeostasis Regulation by sRNAs”, *Microbiology Spectrum*, vol. 6, n°2, pp. 6.2.06, 2018, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/microbiolspec.RWR-0010-2017>. [Acceso: 10

de enero de 2022].

- [60] L. Yan, A. Le Roux, K. Boyapelly, A. Lamontagne, M. Archambault, F. Picard-Jean, D. Lalonde-Seguin, E. St-Pierre, R. Najmanovich, L. Fortier, D. Lafontaine, É. Marsault, “Purine analogs targeting the guanine riboswitch as potential antibiotics against *Clostridioides difficile*”, *European journal of medicinal chemistry*, vol. 143, pp. 755–768, 2018, [En línea]. Disponible en: <https://doi.org/10.1016/j.ejmech.2017.11.079>. [Acceso: 10 de enero de 2023].
- [61] X. Jia, J. Zhang, W. Sun, W. He, H. Jiang, D. Chen, A. Murchie, “Riboswitch control of aminoglycoside antibiotic resistance”, *Cell*, vol. 152, n°1-2, pp. 68-81, 2013, [En línea]. Disponible en: <https://doi.org/10.1016/j.cell.2012.12.019>. [Acceso: 10 de enero de 2022].
- [62] G. Parmeciano Di Noto, M. Molina, C. Quiroga, “Insights Into Non-coding RNAs as Novel Antimicrobial Drugs”, *Frontiers in genetics*, vol. 10, n°57, 2019, [En línea]. Disponible en: <https://doi.org/10.3389/fgene.2019.00057>. [Acceso: 10 de enero de 2022].
- [63] S. Maddocks, “Novel Targets of Antimicrobial Therapies”, *ASM Journals, Microbiology Spectrum*, vol. 4, n°2, 2016, [En línea]. Disponible en: <https://journals.asm.org/doi/10.1128/microbiolspec.VMBF-0018-2015>. [Acceso: 01 de octubre de 2021].
- [64] V. Libis, A. Bernheim, C. Basier, et al, “Silencing of Antibiotic Resistance in *E. coli* with Engineered Phage Bearing Small Regulatory RNAs”, *ACS, Synthetic Biology*, vol. 3, n°12, pp. 1003-1006, 2014, [En línea]. Disponible en: <https://doi.org/10.1021/sb500033d>. [Acceso: 06 de octubre de 2021].
- [65] C. Michaux, A. Hartke, C. Martini, S. Reiss, D. Albrecht, A. Budin-Verneuil, M. Sanguinetti, S. Engelmann, T. Hain, N. Verneuil, J.C. Giard, “Involvement of *Enterococcus faecalis* small RNAs in Stress Response and Virulence”, *ASM Journals, Infection and Immunity*, vol. 82, n°9, 2014, [En línea]. Disponible en: <https://journals.asm.org/doi/10.1128/IAI.01900-14>. [Acceso: 04 de octubre de 2021].
- [66] C. Michaux, E. Hansen, L. Jenniches, M. Gerovac, L. Barquist, J. Vogel, “Single-Nucleotide RNA Maps for the Two Major Nosocomial Pathogens *Enterococcus faecalis* and *Enterococcus faecium*”, *Frontiers in Cellular and Infection Microbiology*, vol. 10, 2020, [En línea]. Disponible en: <https://www.frontiersin.org/articles/10.3389/fcimb.2020.600325>. [Acceso: 04 de octubre de 2021].
- [67] D. Xu, W. Yuan, C. Fan, B. Liu, M. Lu, J. Zhang, “Opportunities and Challenges of Predictive Approaches for the Non-coding RNA in Plants”, *Frontiers in Plant Science*, vol. 13, 2022, [En línea]. Disponible en: <https://www.frontiersin.org/articles/10.3389/fpls.2022.890663>. [Acceso: 20 de enero de 2023].

- [68] M. Hueso, A. Mallén, M. Suñé-Pou, J. Aran, J. Suñé-Negre, E. Navarro, “ncRNAs in Therapeutics: Challenges and Limitations in Nucleic Acid-Based Drug Delivery”, *International Journal of Molecular Sciences*, vol. 22, n°21, pp. 11596, 2021, [En línea]. Disponible en: <https://doi.org/10.3390/ijms222111596>. [Acceso: 20 de enero de 2023].
- [69] M. Latorre, F. Olivares, A. Reyes-Jara, G. López, M. González, “CutC is induced late during copper exposure and can modify intracellular copper content in *Enterococcus faecalis*”, *Biochemical and biophysical research communications*, vol. 406, n°4, pp. 633-637, 2011, [En línea]. Disponible en: <http://repositorio.uchile.cl/handle/2250/123986>. [Acceso: 10 de octubre de 2021].
- [70] A. Odermatt, M. Solioz, “Two trans-Acting Metalloregulatory Proteins Controlling Expression of the Copper-ATPases of *Enterococcus hirae*”, *Journal of Biological Chemistry*, vol. 270, n°9, pp. 4349-4354, 1995, [En línea]. Disponible en: <https://doi.org/10.1074/jbc.270.9.4349>. [Acceso: 28 de Agosto de 2022].
- [71] R. Leinonen, H. Sugawara, M. Shumway, “The sequence read archive”, *Nucleic acids research*, vol. 39, pp. 19-21, 2011, [En línea]. Disponible en: <https://doi.org/10.1093/nar/gkq1019>. [Acceso: 11 de octubre de 2021].
- [72] S. Wigett, S. Andrews, “FastQ Screen: A tool for multi-genome mapping and quality control”, *F1000 Research*, vol. 7, pp. 1338, 2018, [En línea]. Disponible en: <https://doi.org/10.12688/f1000research.15931.2>. [Acceso: 11 de octubre de 2021].
- [73] A. Bolger, M. Lohse, B. Usadel, “Trimmomatic: a flexible trimmer for Illumina sequence data”, *Bioinformatics*, vol. 30, n°15, pp. 2114-2120, 2014, [En línea]. Disponible en: <https://doi.org/10.1093/bioinformatics/btu170>. [Acceso: 06 de octubre de 2021].
- [74] B. Langmead, S. Salzberg, “Fast gapped-read alignment with Bowtie 2”, *Nature Methods*, vol. 9, pp. 357-359, 2012, [En línea]. Disponible en: <https://doi.org/10.1038/nmeth.1923>. [Acceso: 07 de octubre de 2021].
- [75] Y. Xie, J.J. Allaire, G. Grolemond, “R Markdown: The Definitive Guide”, *Chapman and Hall/CRC.*, Ed. 1, 2018, [En línea]. Disponible en: <https://doi.org/10.1201/9781138359444>. [Acceso: 07 de octubre de 2021].
- [76] P. Wright, J. Georg, M. Mann, D. Sorescu, A. Richter, S. Lott, R. Keinkauf, W. Hess, R. Backofen, “CoproRNA and IntaRNA: predicting small RNA targets, networks and interaction domains”, *Nucleic Acids Research*, vol. 42, n°1, pp. 119-123, 2014, [En línea]. Disponible en: <https://doi.org/10.1093/nar/gku359>. [Acceso: 10 de octubre de 2021].
- [77] F. Sievers, D.G. Higgins, “Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences”, *Methods in Molecular Biology*, vol. 1079, pp. 105–116, 2014, [En línea].

Disponible en: https://doi.org/10.1007/978-1-62703-646-7_6. [Acceso: 10 de octubre de 2022].

- [78] R. Lorenz, S. Bernhart, C. Höner zu Siederdisen, H. Tafer, C. Flamm, P. Stadler, I. Hofacker, “ViennaRNA Package 2.0”, *Algorithms for Molecular Biology*, vol. 6, n°1, pp. 26, 2011, [En línea]. Disponible en: <https://doi.org/10.1186/1748-7188-6-26>. [Acceso: 10 de octubre de 2022].
- [79] NCBI, “Introducing the BLAST homepage and form elements/functions of selected search pages”, 2019, [En línea]. Disponible en: <https://www.ncbi.nlm.nih.gov/blast/producttable.shtmlmega>. [Acceso: 09 de octubre de 2021].
- [80] R. Denman, “Using RNAFOLD to predict the activity of small catalytic RNAs”, *BioTechniques*, vol. 15, n°6, pp. 1090–1095, 1993, [En línea]. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/8292343/>. [Acceso: 20 de diciembre de 2022].
- [81] Maccura Biotechnology, “SARS-CoV-2 Fluorescent PCR Kit”, 2020, [En línea]. Disponible en: <https://www.fda.gov/media/137026/download>. [Acceso: 03 de agosto de 2022].
- [82] QIAGEN, “RNeasy[®] Mini Handbook”, 2019, [En línea]. Disponible en: <https://www.qiagen.com/us/resources/resourcedetail?id=14e7cf6e-521a-4cf7-8cbc-bf9f6fa33e24lang=en>. [Acceso: 03 de agosto de 2022].
- [83] Agilent Technologies, “Brilliant II QPCR Master Mix, Instruction Manual”, 2015, [En línea]. Disponible en: <https://www.agilent.com/cs/library/usermanuals/public/600804.pdf>. [Acceso: 03 de agosto de 2022].
- [84] S. Wu, Y. Liu, H. Zhang, L. Lei, “The Susceptibility to Calcium Hydroxide Modulated by the Essential walR Gene Reveals the Role for *Enterococcus faecalis* Biofilm Aggregation”, *Journal of Endodontics*, vol. 45, n°3, pp. 295-301, 2019, [En línea]. Disponible en: <https://doi.org/10.1016/j.joen.2018.11.011>. [Acceso: 03 de agosto de 2022].
- [85] R. Hua, Y. Xia, W. Wu, J. Yan, M. Yang, “Whole transcriptome analysis reveals potential novel mechanisms of low-level linezolid resistance in *Enterococcus faecalis*”, *Gene*, vol. 647, pp. 143-149, 2018, [En línea]. Disponible en: <https://doi.org/10.1016/j.gene.2018.01.008>. [Acceso: 06 de septiembre de 2021].
- [86] X. Jiang, U. Nguyen, C. Ziegler, D. Goel, K. Palmer, F. Morcos, “VanRS and CroRS cross-talk revealed by coevolutionary modeling regulates antibiotic resistance in VanA-type vancomycin-resistant *Enterococcus faecalis*”, *bioRxiv*, 2022, [En línea]. Disponible en: <https://www.biorxiv.org/content/10.1101/2022.05.05.490712v1>. [Acceso: 02 de septiembre de 2022].

- [87] I. Paulsen, et al., “Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*”, *Science*, vol. 299, n°5615), pp. 2071–2074, 2003, [En línea]. Disponible en: <https://doi.org/10.1126/science.1080613>. [Acceso: 02 de diciembre de 2022].
- [88] N. Matange, M. Podobnik, S. Visweswariah, “Metallophosphoesterases: structural fidelity with functional promiscuity”, *Biochemical Journal*, vol. 467, n°2), pp. 201–216, 2015, [En línea]. Disponible en: <https://doi.org/10.1042/BJ20150028>. [Acceso: 28 de diciembre de 2022].
- [89] P. Johnsson, L. Lipovich, D. Grandér, K. Morris, “Evolutionary conservation of long non-coding RNAs; sequence, structure, function”, *Biochimica et biophysica acta*, vol. 1840, n°3, pp. 1063–1071, 2014, [En línea]. Disponible en: <https://doi.org/10.1016/j.bbagen.2013.10.035>. [Acceso: 15 de diciembre de 2022].
- [90] S. Ishikawa, M. Huang, A. Tomita, Y. Kurihara, R. Watanabe, H. Iwai, K. Arakawa, “Complete genome sequences of four bacteria isolated from the gut of a spiny ant (*Polyrhachis lamellidens*)”, *Microbiology Resource Announcements*, vol. 11, n°7 pp. e00333-22, 2022, [En línea]. Disponible en: <https://doi.org/10.1128/mra.00333-22>. [Acceso: 20 de enero de 2023].
- [91] C. Schoch, S. Ciufu, M. Domrachev, C. Hutton, S. Kannan, R. Khovanskaya, D. Leipe, R. Mcveigh, K. O’Neill, B. Robbertse, S. Sharma, V. Soussov, J. Sullivan, L. Sun, S. Turner, I. Karsch-Mizrachi, “NCBI Taxonomy: a comprehensive update on curation, resources and tools”, *Database*, vol. 2020, pp. baaa062, 2020, [En línea]. Disponible en: <https://doi.org/10.1093/database/baaa062>. [Acceso: 20 de enero de 2023].
- [92] D. Matelska, M. Kurkowska, E. Purta, J. Bujnicki, S. Dunin-Horkawicz, “Loss of Conserved Noncoding RNAs in Genomes of Bacterial Endosymbionts”, *Genome biology and evolution*, vol. 8, n°2, pp. 426–438, 2016, [En línea]. Disponible en: <https://doi.org/10.1093/gbe/evw007>. [Acceso: 15 de diciembre de 2022].
- [93] J. Graf, M. Kretz, “From structure to function: Route to understanding lncRNA mechanism”, *BioEssays*, vol. 42, pp. e2000027, 2020, [En línea]. Disponible en: <https://doi.org/10.1002/bies.202000027>. [Acceso: 02 de diciembre de 2022].
- [94] K. Sato, M. Akiyama, Y. Sakakibara, “RNA secondary structure prediction using deep learning with thermodynamic integration”, *Nature Communications*, vol. 12, pp. 941, 2021, [En línea]. Disponible en: <https://doi.org/10.1038/s41467-021-21194-4>. [Acceso: 02 de diciembre de 2022].
- [95] K. Doshi, J. Cannone, C. Cobaugh, et al., “Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction”, *BMC Bioinformatics*, vol. 5, pp. 105, 2004, [En línea]. Disponible en: <https://doi.org/10.1186/1471-2105-5-105>. [Acceso: 02 de diciembre de 2022].

- [96] W. Miller, J. Munita, C. Arias, “Mechanisms of antibiotic resistance in enterococci”, *Expert Review of Anti-infective Therapy*, vol. 12, n°10, pp. 1221-1236, 2014, [En línea]. Disponible en: <https://doi.org/10.1586/14787210.2014.956092>. [Acceso: 20 de noviembre de 2022].
- [97] C. Reding-Roman, M. Hewlett, S. Duxbury, F. Gori, I. Gudelj, R. Beardmore, “The unconstrained evolution of fast and efficient antibiotic-resistant bacterial genomes”, *Nature Ecology Evolution*, vol. 1, pp. 0050, 2017, [En línea]. Disponible en: <https://doi.org/10.1038/s41559-016-0050>. [Acceso: 06 de septiembre de 2022].
- [98] L. Harris, J. Theriot, “Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis”, *Trends in Microbiology*, vol. 26, n°10, pp. 815-832, 2018, [En línea]. Disponible en: Trends in Microbiology. [Acceso: 20 de noviembre de 2022].
- [99] A. Bourgoigne, et al., “Large scale variation in *Enterococcus faecalis* illustrated by the genome analysis of strain OG1RF”, *Genome Biology*, vol. 9, 2008, [En línea]. Disponible en: <https://doi.org/10.1186/gb-2008-9-7-r110>. [Acceso: 02 de agosto de 2022].
- [100] F. Depardieu, I. Podglajen, R. Leclercq, E. Collatz, P. Courvalin, “Modes and modulations of antibiotic resistance gene expression”, *Clinical Microbiology Reviews*, vol. 20, n°1, pp. 79-114, 2007, [En línea]. Disponible en: <https://journals.asm.org/doi/abs/10.1128/CMR.00015-06>. [Acceso: 21 de noviembre de 2022].
- [101] P. Rathore, S. Joy, R. Yadav, W. Ramakrishna, “Co-occurrence and patterns of phosphate solubilizing, salt and metal tolerant and antibiotic-resistant bacteria in diverse soils”, *3 Biotech*, vol. 11, n°7, 2021, [En línea]. Disponible en: <https://doi.org/10.1007/s13205-021-02904-7>. [Acceso: 21 de enero de 2023].
- [102] M. Grillo-Puertas, L. Schurig-Briccio, L. Rodríguez-Montelongo, et al., “Copper tolerance mediated by polyphosphate degradation and low-affinity inorganic phosphate transport system in *Escherichia coli*”, *BMC Microbiology*, vol. 14, n°72, 2014, [En línea]. Disponible en: <https://doi.org/10.1186/1471-2180-14-72>. [Acceso: 21 de enero de 2023].
- [103] A. Quinlan, I. Hall, “BEDTools: a flexible suite of utilities for comparing genomic features”, *Bioinformatics*, vol. 26, n°6, pp. 841–842, 2010, [En línea]. Disponible en: <https://doi.org/10.1093/bioinformatics/btq033>. [Acceso: 02 de septiembre de 2021].

Anexos

Anexo A. Análisis *in silico*

En las siguientes subsecciones se detalla el uso de los programas relacionados con el análisis *in silico* de la primera etapa del trabajo, además de las líneas de código utilizadas en el servidor, paso a paso para obtener los resultados de los ARNnc target.

A.1. Programas utilizados

Los *softwares* utilizados para procesar los datos de RNA-seq obtenidos desde bases de datos fueron principalmente los siguientes:

- **FastQC:** Este programa permite realizar un análisis de la calidad de las secuencias a partir de los datos crudos de la secuenciación [72]. Entrega los análisis obtenidos de manera gráfica, donde en esta investigación se consideró principalmente la calidad de la secuenciación por base o “*Per base sequence quality*”
- **Trimomatic:** Esta herramienta permite remover los adaptadores provenientes de la secuenciación del material genético, bases con baja calidad en el análisis de la secuenciación o bases no determinadas (N), eliminar *reads* de largo menor a cierta cantidad de bases, entre otras. En el presente estudio no se necesitó remover adaptadores para las secuenciaciones de vancomicina, ceftriaxona ni cobe, pues los datos no poseían adaptadores y eran de buena calidad [73].
- **Bowtie2:** Es un paquete utilizado a lo largo del análisis de la secuenciación. Posee múltiples herramientas, donde se destaca su uso para el alineamiento de secuencias de *reads* con algunas de referencia [74]. En este caso, se utilizó como referencia el genoma de *E. faecalis* V583.
- **BedTools:** Corresponde a un *set* de herramientas útiles para el análisis de secuenciaciones. Permite intersectar, unir, contar, complementar y mezclar información genómica de distintos archivos en formatos variados [103].

A.2. Código

Luego del análisis de la secuenciación y remoción de adaptadores, el trabajo realizado desde el servidor consta de 13 pasos detallados a continuación.

1. Mapeo de los datos de RNAseq en el genoma de referencia de la cepa de *E. faecalis* OG1RF. Para ello se utiliza la herramienta Bowtie2. Se detallan los comandos utilizados para el caso de datos *paired-end* y *single-end*.
 - PAIRED-END: `bowtie2 -k 1 -p 20 -fr -mm -N 0 -x RUTA_INDICE -1 FASTQ_FORWARD -2 FASTQ_REVERSE -U`
 - SINGLE-END: `bowtie2 -k 1 -p 20 -fr -mm -N 0 -x RUTA_INDICE -1 FASTQ_FORWARD -2 FASTQ_REVERSE -U UNPAIRED_FORWARD, UNPAIRED_REVERSE -S OUT_NAME.sam`
2. Transformación de archivos BED a SAM. Se utiliza la herramienta *sam2bed*.
 - Para cada archivo SAM: `sam2bed -max-mem 70G < OUT_NAME.sam > OUT_NAME_RAW.bed`
3. Cortar las columnas que no son necesarias en los archivos BED generados previamente. Para ello se utiliza el comando *cut*.
 - Para cada archivo BED: `cut -f 1-6 OUT_NAME_RAW.bed > OUT_NAME_FINAL.bed`
4. Comenzar el ensamble de transcritos para cada archivo BED, mediante el uso de herramientas como *sortBed* y *mergeBed*.
 - DATOS HEBRA ESPECÍFICO: `sortBed -i OUT_NAME_FINAL.bed | mergeBed -s -c 6 -o distinct -delim ";" -i - > OUT_NAME.merge`
 - DATOS NO HEBRA ESPECÍFICO: `sortBed -i OUT_NAME_FINAL.bed | mergeBed -c 6 -o distinct -delim ";" -i - > OUT_NAME.merge`
5. Agrupar los archivos generados en el paso anterior para posteriormente realizar el ensamble final de transcritos que agrupa todas las muestras. El ensamble debe hacerse por separado según su tipo: para hebra específicos y para no hebra específicos. Se utiliza el comando *cat*.
 - `cat ARCHIVO_MERGE_1 ARCHIVO_MERGE_2 ARCHIVO_MERGE_3 ... ARCHIVO_MERGE_N > OUT_NAME.merge2`
6. Modificar el archivo MERGE2 para que sea compatible con las herramientas de ensamble. Se utiliza la herramienta *awk*.
 - `awk 'x="\t"; print $1 x $2 x $3 x "assembly_" NR x "0" x $4' OUT_NAME.merge2 > OUT_NAME_MODIFIED.merge2`
7. Realizar el ensamble final de transcritos según su tipo (hebra específico o no específico) y dejar los archivos en formato BED. Se utilizan las herramientas *sortBed* y *mergeBed* junto con *awk*.

- DATOS HEBRA ESPECÍFICO: `sortBed -i OUT_NAME_MODIFIED.merge2 | mergeBed -s -c 6 -o distinct -delim ";" -i - > OUT_NAME.merge3`
- DATOS NO HEBRA ESPECÍFICO: `sortBed -i OUT_NAME_MODIFIED.merge2 | mergeBed -c 6 -o distinct -delim ";" -i - > OUT_NAME.merge3`
- `awk 'x="\t"; print $1 x $2 x $3 x "assembly_" NR x "0" x $4' OUT_NAME.merge3 > FINAL_TRANSCRIPTS.bed`

8. Transformar el archivo del paso anterior de MERGE3 a GFF, pues es el formato requerido por el *HTSEQ-count* para determinar el número de *reads* de cada transcrito. Se realiza mediante un *script* en Perl.

- `perl 1_makeGFF_forHTseq.pl OUT_NAME.merge3 > OUT_NAME.gff`

El *script* utilizado se detalla a continuación:

```

1  if (@ARGV < 1){
2      die "Usage: perl $0 <MERGE3 FILE>\n";
3  }
4
5  my @lines_merge2 = `cat $ARGV[0]`;
6  chomp @lines_merge2;
7
8  my $index = 0;
9  foreach my $each_line (@lines_merge2){
10     $index++;
11
12     my @cols = split ("\t", $each_line);
13
14     print
15         "$cols[0]\tRNA-seq\tGene\t$cols[1]\t$cols[2]\t0\t$cols[3]\t0\tID=
↳ assembly_ $index\n";
16
17 }
18

```

9. Determinar el número de cuentas de cada transcrito desde cada muestra. Se utilizan los archivos SAM transformados en BAM de cada muestra utilizada para cada ensamble, por lo que este paso debe hacerse para cada BAM proveniente del mapeo de los archivos FASTQ. Los archivos deben quedar con extensión “.htseq” y la herramienta utilizada es *HTSEQ-count*.

- `htseq-count -q -f bam -s yes -t Gene -i ID -m intersection-nonempty RUTA_BAM OUT_NAME.gff > OUT_NAME.htseq`

- Unir todos los archivos resultantes del último paso para generar una matriz de cuentas. Se utiliza un *script* de Perl que toma todos los archivos generados por *HTSEQ-count* y forma una matriz de cuentas que posteriormente se usa para filtrar los datos.

```
• perl 1_joinHTSEQ_files.pl RUTA_CARPETA_CUENTAS >
  OUT_NAME.RawCounts
```

El *script* utilizado se detalla a continuación:

```
1  if (@ARGV < 1){
2      die "Usage: perl $0 <RUTE_FOLDER_HTSEQ_FILES>\n";
3  }
4
5
6  my @files_HTseq = `ls $ARGV[0]*.htseq`;
7  chomp @files_HTseq;
8
9  my %MATRIX;
10
11 my @name_samples = (); ## en array para que esten en orden
12 my @ID_genes = ();
13
14 my $index = 0;
15
16 foreach my $each_file (@files_HTseq){
17     my @lines_eachfile = `cat $each_file`;
18     chomp @lines_eachfile;
19
20     my $name_out = `ls $each_file | awk -F "/" '{print \$NF}'`;
21     chomp $name_out;
22
23     $name_out =~ s/.htseq//g;
24
25     $name_samples[$index] = $name_out;
26     $index++;
27
28     my $index_2 = 0;
29     foreach my $each_line (@lines_eachfile){
30         my @cols = split ("\t", $each_line);
31
32         my $ID = shift (@cols);
33         my $counts = shift (@cols);
34
35         $ID_genes[$index_2] = $ID;
36         $index_2++;
37
```

```

38     $MATRIX{$name_out}{$ID} = $counts;
39
40 }
41 }
42
43
44 print "ID";
45 foreach my $each_nameSample (@name_samples){
46     print "\t$each_nameSample";
47 }
48 print "\n";
49
50
51 foreach my $each_geneID (@ID_genes){
52     print "$each_geneID";
53
54     foreach my $each_nameSample (@name_samples){
55
56         if (defined $MATRIX{$each_nameSample}{$each_geneID}){
57             print "\t$MATRIX{$each_nameSample}{$each_geneID}";
58         } else {
59             print "\tNADA";
60         }
61     }
62
63     print "\n";
64
65 }
66

```

11. Filtrar los transcritos por un número determinado de cuentas. En este paso se filtran los transcritos basado en que en al menos en una condición analizada, el 50% de las muestras tenga más de 5 *reads*. Se necesita de un *script* en Perl.

- perl 2_filter_by_counts.pl SamplesCondition.tab OUT_NAME.RawCounts 5 > OUT_NAME.filtered

El archivo *SamplesCondition.tab* debe contener las condiciones a las que pertenece cada muestra; la columna 1 debe tener los mismos nombres que la primera fila del archivo *OUT_NAME.RawCounts* (archivo del paso 9). El *script* utilizado se detalla a continuación:

```

1  if (@ARGV < 3){
2      die "\tUsage: perl $0 <CONDITIONS_FILE> <COUNTS_RAW_FILE> <
    ↪ CUTOFF_COUNTS>\n";

```

```

3   }
4
5   ## ----- PART I ----- ##
6   ##      Saving the column number of each sample respect to ID of replicates
7
8   ## -- RECOVER ID REPLICATES-- ##
9   my @condition_file = `cat $ARGV[0]`;
10  chomp @condition_file;
11
12  my %SAMPLE_2_CONDITION;
13  foreach my $each_condition (@condition_file){
14      my @each_col = split ("\t", $each_condition);
15      my $sample = shift (@each_col);
16      my $condition = shift (@each_col);
17
18      $SAMPLE_2_CONDITION{$sample} = $condition;
19
20  }
21
22  my @id = `cut -f 2 $ARGV[0] | sed -e 's/[0-9]\\+//g' | sort -u`;
23  chomp @id;
24
25  my $head_counts = `head -1 $ARGV[1]`;
26  chomp $head_counts;
27
28  print "$head_counts\n";
29
30  my @div_head = split ("\t", $head_counts);
31
32  ## -- MAKE HASH OF HASH TO SAVE NUMBER OF COLUMN OF EACH
33  ##      ↪ REPLICATE-- ##
34  my %ID_1 = ();
35
36  foreach my $each_id (@id){
37
38      my $index_1 = 0;
39
40      foreach my $each_col_head (@div_head){
41
42          if ($SAMPLE_2_CONDITION{$each_col_head} =~ $each_id){
43              $ID_1{$each_id}{$each_col_head} = $index_1; ## Here save the column
44              ↪ number ex: "$ID_1{SAMPLE}{CRL1} = 7"
45
46          }
47
48          $index_1++;
49

```

```

46     }
47 }
48
49 ### ----- PART II ----- ###
50 ### Generating the parse based on a specific number of reads on each sample
51
52 ### -- OPEN COUNT FILE -- ###
53 open (IN,"<",$ARGV[1]);
54
55 <IN>;
56
57 while (<IN>){
58     chomp $_;
59
60     my $count_final = 0; # Count the number of condition with >= 50 % of samples
61     ↪ with at least "N" reads in their contig formed
62
63     my @cols_line = split ("\t", $_);
64
65     foreach my $each_id_first (@id){
66
67         my $number_samples_id = 0; # HERE the number of samples
68         my $count_of_replicates = 0; # HERE the number of replicates with more than
69         ↪ "N" reads
70
71         foreach my $values (values %{$ID_1{$each_id_first}}){
72             $number_samples_id++;
73             if ($cols_line[$values] >= $ARGV[2]){
74                 $count_of_replicates++;
75             }
76         }
77
78         my $percent = ($count_of_replicates * 100) / $number_samples_id; # HERE
79         ↪ percent of samples
80
81         if ($percent >= 50 ){
82             $count_final++;
83         }
84     }
85
86     if ($count_final >= 1){
87         my $recover_line = join ("\t", @cols_line);
88
89         if ($recover_line !~ /___/){

```

```

88     print "$recover_line\n";
89     }
90
91     }
92
93
94     }
95
96     close (IN);
97     close ($fh_1);
98

```

12. Para identificar los ARNnc se debe filtrar por coordenadas genómicas, pues se desea saber cuales de esos transcritos pertenecen a genes (CDS) y cuales no, siendo estos últimos potenciales ARNnc. Se deben comparar los transcritos ensamblados con anotaciones previamente hechas, por lo que se requiere recuperar las coordenadas de los transcritos que pasaron el filtro anterior y dejar todo en formato BED. Para realizar todo esto se usa un *script* de Perl.

- perl 1_filtered_2_BED.pl OUT_NAME.filtered OUT_NAME.gff
OUT_NAME.bed

El archivo *OUT_NAME.filtered*, corresponde al archivo del paso 11; *OUT_NAME.gff*, al archivo del paso 8; *OUT_NAME* es el nombre que se le da al archivo BED. El *script* utilizado es el que se detalla a continuación:

```

1  if (@ARGV < 2){
2      die "\tUsage: perl $0 <FILTERED_FILE> <GFF_ASSEMBLED(used in HTSEQ)>
    ↪ <NAME_OUT>\n";
3  }
4
5  my $filtered_file = $ARGV[0];
6  my $assembled_GFF = $ARGV[1];
7  my $name_out = $ARGV[2];
8
9  ### ----- GFF assembled in HASH
10 my @lines_assembledGFF = `cat $assembled_GFF`;
11 chomp @lines_assembledGFF;
12
13 my %ASSEMBLY_BED;
14
15 foreach my $each_line (@lines_assembledGFF){
16     my @cols = split ("\t", $each_line);
17
18     my $chr = $cols[0];

```

```

19     my $start = $cols[3];
20     my $end = $cols[4];
21     my $strand = $cols[6];
22     my $ID = $cols[-1];
23     $ID =~ s/ID=//g;
24
25     $ASSEMBLY_BED{$ID} = "$chr\t$start\t$end\t$ID\t0\t$strand";
26
27 }
28
29 ## ----- FILTERED ASSEMBLED TO BED
30 ↪ FILE
31
32 open (my $fh_1,">","$name_out");
33
34
35 my @lines_filtered = `cat $filtered_file`;
36
37 chomp @lines_filtered;
38
39
40 my $head = shift (@lines_filtered);
41
42
43 foreach my $each_line (@lines_filtered){
44     my @cols = split ("\t", $each_line);
45
46     my $ID = shift (@cols);
47     print $fh_1 "$ASSEMBLY_BED{$ID}\n";
48 }
49
50 close ($fh_1);
51
52

```

13. Realizar la comparación de los transcritos BED filtrados con la anotación disponible. El filtro en este paso se compone de 3 decisiones: el transcrito tiene o no *match* con alguna anotación, el transcrito no supera un largo de bases determinado (1500 bases) y el transcrito está alejado del gen más cercano por lo menos 30 bases, debido a que podría ser parte de la región UTR de algún gen. Para este paso se utiliza un *script* de Perl.

- perl 1_making_intersect.pl OUT_NAME.bed z_mergedFeatures_v583.bed 30 1500 yes 2

Con respecto a los archivos utilizados, OUT_NAME.bed, es el archivo generado en el paso anterior y z_mergedFeatures_v583.bed corresponde a un archivo con todas las anotaciones de la bacteria en formato BED. Los otros parámetros indicados son la distancia escogida al gen más cercano (30), el largo del transcrito (1500), información

de si los datos son hebra orientados (yes) y que la orientación de los datos es de hebra opuesta (2). El comando da como resultados varios archivos, cada uno de ellos teniendo el filtro realizado, donde los últimos archivos generados son los que se utilizarán más adelante, con nombre *OUT_script1_3**. Tienen en formato GFF y formato BED todos los transcritos que lograron pasar los filtros de este paso. El *script* utilizado en el comando se detalla a continuación:

```

1  if (@ARGV < 5){
2      die "\tUsage: perl $0 <FILTERED_BY_COUNTS(BED)> <
    ↳ RUTE_ANNOTATION(BED)> <DISTANCE_FILTER_CLOSESTBED> <
    ↳ SIZE_TRANSCRIPT> <STRAND[yes or not]> <TYPE_STRAND(1 or 2)>
3
4      NOTE: The script filter using INTERSECTBED, CLOSESTED and the Size of
    ↳ the transcript.
5      <DISTANCE_FILTER_CLOSESTBED> Based to make filter based on distance
    ↳ to some gene.
6      <STRAND[yes or not]> Write 'yes' for strand specific analysis.
7      <TYPE_STRAND> Write 1 for same strand orientation or 2 for opposite strand
    ↳ orientation. Ignore if <STRAND> is 'not'.\n";
8
9  }
10
11  my $rute_filtered = $ARGV[0];
12  my $rute_annotation = $ARGV[1];
13  my $distance = $ARGV[2];
14  my $size_filter = $ARGV[3];
15  my $strand = $ARGV[4];
16  my $type_strand = $ARGV[5];
17
18  my $name_out = `ls $rute_filtered | awk -F "/" '{print \$NF}'`;
19  chomp $name_out;
20
21  print "\n\t **** NOTE: Name file is '$name_out' -> out name only will have ";
22  $name_out =~ s/(OUT_script)*[0-9]*(_)*[0-9]*[a-z]+_|.bed//gi;
23  print "'$name_out' ****\n\n";
24
25
26  ## ----- MAKING INTERSECT BED
27  my @intersect_lines = ();
28
29  if ($strand eq "yes"){
30      if ($type_strand == 1){
31          @intersect_lines = `sortBed -i $rute_filtered | intersectBed -wao -s -a - -b
    ↳ $rute_annotation`; ## SAME STRAND
32          chomp @intersect_lines;

```

```

33
34     } else {
35         print "\n\tNOTE: You used type 2 reverse, i.e. Opposite strand for analysis.\n\n"
↪ n";
36         @intersect_lines = `sortBed -i $rute_filtered | intersectBed -wao -S -a - -b
↪ $rute_annotation`; ## OPPOSITE STRAND
37         chomp @intersect_lines;
38     }
39
40
41 } else {
42     @intersect_lines = `sortBed -i $rute_filtered | intersectBed -wao -a - -b
↪ $rute_annotation`;
43     chomp @intersect_lines;
44
45 }
46
47
48 ## ----- PARSING I - Based on INTERSECT
49 open (my $fh_1,">","OUT_script1_1_NotIntersected\_$_name_out.bed");
50
51 foreach my $each_line (@intersect_lines){
52     my @cols = split ("\t", $each_line);
53
54     if ($cols[-1] == "0"){
55         my $not_intersect = "$cols[0]\t$cols[1]\t$cols[2]\t$cols[3]\t$cols[4]\t$cols[5]";
56         print $fh_1 "$not_intersect\n";
57
58     }
59 }
60
61 close ($fh_1);
62
63
64 ## ----- MAKING CLOSEST BED
65 my @closest_lines = ();
66
67 if ($strand eq "yes"){
68     if ($type_strand == 1){
69         @closest_lines = `sortBed -i OUT_script1_1_NotIntersected\_$_name_out.bed
↪ | closestBed -d -s -a - -b $rute_annotation`;
70         chomp @closest_lines;
71
72     } else {
73         @closest_lines = `sortBed -i OUT_script1_1_NotIntersected\_$_name_out.bed

```

```

↪ | closestBed -d -S -a - -b $rute_annotation`;
74     chomp @closest_lines;
75 }
76
77
78 } else {
79     @closest_lines = `sortBed -i OUT_script1_1_NotIntersected\_$_name_out.bed |
↪ closestBed -d -a - -b $rute_annotation`;
80     chomp @closest_lines;
81
82 }
83
84
85 ## ----- PARSING II - Based on CLOSEST
86 open (my $fh_2,">","OUT_script1_2_NotClosest$distance\_$_name_out.bed");
87
88 foreach my $each_line (@closest_lines){
89
90     my @cols = split ("\t", $each_line);
91
92     if ($cols[-1] >= $distance){ ## HERE distance for filter
93         my $pass_the_distance = "$cols[0]\t$cols[1]\t$cols[2]\t$cols[3]\t$cols[4]\t$cols
↪ [5]";
94         print $fh_2 "$pass_the_distance\n";
95     }
96 }
97
98 }
99
100 close ($fh_2);
101
102
103 ## ----- PARSING III - Based on SIZE TRANSCRIPT
104 open (my $fh_3,">","OUT_script1_3_NotClosest$distance\
↪ _MinorThanNotClosest$size_filter\_$_name_out.bed");
105 open (my $fh_4,">","OUT_script1_3_NotClosest$distance\
↪ _MinorThanNotClosest$size_filter\_$_name_out.gff");
106
107 my @lines_transcripts = `cat OUT_script1_2_NotClosest$distance\_$_name_out.
↪ bed`;
108 chomp @lines_transcripts;
109
110 foreach my $each_line (@lines_transcripts){
111     my @cols = split ("\t", $each_line);
112

```

```

113     my $size_transcript = $cols[2] - $cols[1];
114
115     if ($size_transcript <= $size_filter){
116         print $fh_3 "$each_line\n";
117         print $fh_4 "$cols[0]\tRNA-seq\tGene\t$cols[1]\t$cols[2]\t0\t$cols[5]\t0\tID=
↪ $cols[3]\n";
118
119     }
120
121 }
122
123 close ($fh_3);
124 close ($fh_4);
125

```

Luego de los pasos anteriores se obtienen los transcritos que no tienen *match* con ninguna anotación, junto con aquellos que pasaron los filtros elegidos. Lo siguiente es recuperar las cuentas de cada transcrito desde el archivo generado en el paso 9 para así obtener los datos de expresión diferencial, mediante programación en R como se muestra en la siguiente sección.

A.2.1. Expresión diferencial

Haciendo uso de los archivos generados en los pasos descritos en la sección A.2, se obtienen los datos de expresión diferencial utilizando el código de programación en R que se muestra a continuación. El archivo final generado permitirá ver las veces de cambio de los ARNnc filtrados de una condición específica con respecto a la condición basal. En este caso y ejemplo de código se analizó la expresión diferencial de los ARNnc del transcriptoma de *E. faecalis* en presencia de ceftriaxona y vancomicina con respecto a la condición control o basal.

```

1 library("DESeq2")
2
3 cts <- as.matrix(read.csv("../04-RawCounts/OUT_script2b_FinalRawCounts.tab",sep="\t",
↪ row.names="ID"))
4 coldata <- read.csv("CONDITION.csv",sep=",", row.names=1)
5 coldata <- coldata[,c("condition", "type")]
6
7 dds <- DESeqDataSetFromMatrix(countData = cts,
8                               colData = coldata,
9                               design = ~ condition)
10
11 dds$condition <- factor(dds$condition, levels = c("Cef_OneHour",
12                                                  "Control_FourtyFiveMin",
13                                                  "Control_OneHour",
14                                                  "Van_FourtyFiveMin"))

```

```

15
16 dds <- DESeq(dds)
17
18 table_counts_normalized <- counts(dds, normalized=TRUE)
19 write.csv(as.data.frame(table_counts_normalized), file="DESEQ2_normalized.csv")
20
21 res <- results(dds, contrast=c("condition", "Cef_OneHour", "Control_OneHour"))
22 write.csv(as.data.frame(res),
23           file="1_Cef_OneHour_vs_Control_OneHour.csv")
24
25 res <- results(dds, contrast=c("condition", "Van_FourtyFiveMin", "Control_FourtyFiveMin"))
26 write.csv(as.data.frame(res),
27           file="2_Van_FourtyFiveMin_vs_Control_FourtyFiveMin.csv")

```

A.3. Conservación de los ARNnc

En las Figuras A.1, A.2, A.3 y A.5, se encuentra el alineamiento de los nucleótidos y conservación de la secuencia del transcrito no codificante nc3340 en distintas especies, utilizando BLASTN; en las Figuras A.6 y A.7, los del transcrito nc3683.

```

>Enterococcus sp. PLM3 DNA, complete genome
Sequence ID: AP025690.1 Length: 2859161
Range 1: 1087284 to 1087992

Score:1282 bits(694), Expect:0.0,
Identities:704/709(99%), Gaps:0/709(0%), Strand: Plus/Minus

Enterococcus faecalis V583 1      AAAACATTTGTGATTGGATAAAAAGGAATCCAATTTGTCCGCCATGGCTATGTCCTGCT 60
Enterococcus sp. PLM3      1087992 AAAACATTTGTGATTGGATAAAAAGGAATCCAATTTGTCCGCCATGGCTATGTCCTGCT 1087933
Enterococcus faecalis V583 61      AAACTAACTGTGTACCGTACCAACAATGCATCCGCGACATCTGGTTCATGGGTCATT 120
Enterococcus sp. PLM3      1087932 AAACTAACTGTGTACCGTACCAACAATGCATCCGCGACATCTGGTTCATGGGTCATT 1087873
Enterococcus faecalis V583 121     AATAACGCATAGTCATACGTTTCTCGATAGGATAATGTTTCTGAAACGGAAGGATTTCCC 180
Enterococcus sp. PLM3      1087872 AATAACGCATAGTCATACGTTTCTCGATAGGACAATGTTTCTGAAACGGAAGGATTTCC 1087813
Enterococcus faecalis V583 181     AATAACGAGTCATCTAAACCACCAATAAAAAGACGGCGTCCATCTGCTAAGGTTACTGTT 240
Enterococcus sp. PLM3      1087812 AATAACGAGTCATCTAAACCACCAATAAAAAGACGGCGTCCATCTGCTAAGGTTACTGTT 1087753
Enterococcus faecalis V583 241     TCTCCTTGTTTTTTCATAACGGTAAAGCCACTGGCCTCCATGACTTCATCATAACCCGCA 300
Enterococcus sp. PLM3      1087752 TCTCCTTGTTTTTTCATAACGGTAAAGCCACTGGCCTCCATGACTTCATCATAACCCGCA 1087693
Enterococcus faecalis V583 301     GAAGCACCGCCACCATAATCATGATTTCCCAAAGTGCATATTTTCCAATTGTCGCTTGT 360
Enterococcus sp. PLM3      1087692 GAAGCACCGCCACCATAATCATGATTTCCCAAAGTGCATATTTTCCAATTGTCGCTTGT 1087633
Enterococcus faecalis V583 361     AACTGTTTTAACTTTTCAATCATCGGTTCTTTTTGTTGCGGATTTTGGCATAATTGTCA 420
Enterococcus sp. PLM3      1087632 AACTGTTTTAACTTTTCAATCATCGGTTCTTTTTGTTGCGGATTTTGGCATAATTGTCA 1087573
Enterococcus faecalis V583 421     AATAAATCACCTGTA AAAACGATAAATATCTGGTTTTTTCACGATTGACTttttttATGACT 480
Enterococcus sp. PLM3      1087572 AATAAATCACCTGTA AAAACGATAAATATCTGGTTTTTTCACGATTGACTttttttATGACT 1087513
Enterococcus faecalis V583 481     TTATCTAACCGCTTCGTCTCATAAAAATTCTGAAACTTGTATATCCGAAAGTTGCACAACA 540
Enterococcus sp. PLM3      1087512 TTATCTAACCGCTTCGTCTCATAAAAATTCTGAAACTTGTATATCCGAAAGTTGCACAACA 1087453
Enterococcus faecalis V583 541     TTCAATGGTTCTCGTTGATTTTTCTTTCCCAAAGTCACTGATTAACGTGAACTAAATTC 600
Enterococcus sp. PLM3      1087452 TTCAATGGTTCTCGTTGATTTTTCTTTCCCAAAGTCACTGATTAACGTGAACTAAATTC 1087393
Enterococcus faecalis V583 601     GGCTCAATCTTCCAGCATAGATTGGGAGTGCATAATTAAGaaaaaaaTAAGAGCATT 660
Enterococcus sp. PLM3      1087392 GGCTCAATCTTCCAGCATAGATTGGGAGTGCATAATTAAGAAAAAAAAAATTAAGAGCATT 1087333
Enterococcus faecalis V583 661     ATACTTAAACCAAttttttAACACATCACTCGCTCCTTTCTTTCTT 709
Enterococcus sp. PLM3      1087332 ATACTTAAACCAATTTTTTAAACACATCACTCGCTCCTTTCTTTCTT 1087284

```

Figura A.1: Alineamiento de ARN nc3340 en *Enterococcus* sp. PLM3. La figura muestra el alineamiento de nucleótidos del transcrito nc3340 en el genoma de *Enterococcus* sp. PLM3. También se entrega información sobre el ID de la secuencia, el porcentaje de identidad del alineamiento y puntaje entregado por BLASTN.

```

>Enterococcus sp. LX10 chromosome, complete genome
Sequence ID: CP092784.1 Length: 2737931
Range 1: 1088250 to 1088958

Score:1271 bits(688), Expect:0.0,
Identities:702/709(99%), Gaps:0/709(0%), Strand: Plus/Minus

Enterococcus faecalis V583 1      AAAACATTTGTGATTGGATAAAAAGGAATCCAAATTTGTCGCGCATGGCTATGTCCTGCT 60
Enterococcus sp. LX10      1088958 AAAACATTTATGATTGGATAAAAAGGAATCCAAATTTGTCGCGCATGGCTATGTCCTGCT 1088899
Enterococcus faecalis V583 61      AAAACTAACTGTGTACCCGTACCAACAATGCATCCGCGACATCTGGTTCATGGGTCATT 120
Enterococcus sp. LX10      1088898 AAAACTAACTGTGTACCCGTACCAATAAATGCATCCGCGACATCTGGTTCATGGGTCATT 1088839
Enterococcus faecalis V583 121     AATAACGCATAGTCATACGTTTCTCGATAGGATAATGTTTCTGAAACGGAAGGATTTCCC 180
Enterococcus sp. LX10      1088838 AATAACGCATAGTCATACGTTTCTCGATAGGATAATGTTTCTGAAACGGAAGGATTTCCCT 1088779
Enterococcus faecalis V583 181     AATAACGAGTCATCTAAACCACCAATAAAAAGACGGCGTCCATCTGCTAAGGTTACTGTT 240
Enterococcus sp. LX10      1088778 AATAACGAGTCATCTAAACCACCAATAAAAAGACGGCGTCCATCTGCTAAGGTTACTGTT 1088719
Enterococcus faecalis V583 241     TCTCCTTGGTTTTTTCATAACGGTAAAGCCACTGGCCTCCATGACTTCATCATAACCCGCA 300
Enterococcus sp. LX10      1088718 TCTCCTTGGTTTTTTCATAACGGTAAAGCCACTGGCCTCCATGACTTCATCATAACCCCA 1088659
Enterococcus faecalis V583 301     GAAGCACCGCCACCATAATCATGATTTCCCAAACCTGCATATTTTCCAATTTGTCGCTTGT 360
Enterococcus sp. LX10      1088658 GAAGCACCGCCACCATAATCATGATTTCCCAAACCTGCATATTTTCCAATTTGTCGCTTGC 1088599
Enterococcus faecalis V583 361     AACTGTTTTAACTTTTCAATCATCGGTTCTTTTTGTTGCGGATTTTTGGCATAATTGTCA 420
Enterococcus sp. LX10      1088598 AACTGTTTTAACTTTTCAATCATCGGTTCTTTTTGTTGCGGATTTTTGGCATAATTGTCA 1088539
Enterococcus faecalis V583 421     AATAAATCACCTGTAAAAACGATAAATATCTGGTTTTTTCACGATTGACTtttttttATGACT 480
Enterococcus sp. LX10      1088538 AATAAATCACCTGTAAAAACGATAAATATCTGGTTTTTTCACGATTGACTTTTTTTATGACT 1088479
Enterococcus faecalis V583 481     TTATCTAACCGCTTCGTCTCATAAAATCTGAAACTTGTATATCCGAAAGTTGCACAACA 540
Enterococcus sp. LX10      1088478 TTATCTAACCGCTTCGTCTCATAAAATCTGAAACTTGTATATCCGAAAGTTGCACAACA 1088419
Enterococcus faecalis V583 541     TTCAATGGTTCTCGTTGATTTTTCTTTCCCAAAGTCACCTGATTAACGTGAACTAAATTC 600
Enterococcus sp. LX10      1088418 TTCAATGGTTCTCGTTGATTTTTCTTTCCCAAAGTCACCTGATTAACGTGAACTAAATTC 1088359
Enterococcus faecalis V583 601     GGCTCAATCTTCCACGCATAGATTGGGAGTGCATAATTAAGaaaaaaaTAAGAGCATT 660
Enterococcus sp. LX10      1088358 GGCTCAATCTTCCACGCATAGATTGGGAGTGCATAATTAAGAAAAAAAAAATAGAGCATT 1088299
Enterococcus faecalis V583 661     ATACTTAAACCAAttttttAACCACATCACTCGCTCCTTTCTTTCTT 709
Enterococcus sp. LX10      1088298 ATACTTAAACCAATTTTTTAAACCACATCACTCGCTCCTTTCTTTCTT 1088250

```

Figura A.2: Alineamiento de ARN nc3340 en *Enterococcus* sp. LX10. La figura muestra el alineamiento de nucleótidos del transcrito nc3340 en el genoma de *Enterococcus* sp. LX10. También se entrega información sobre el ID de la secuencia, el porcentaje de identidad del alineamiento y puntaje entregado por BLASTN.

```

>Lactococcus garvieae strain JJJN1 chromosome, complete genome
Sequence ID: CP026502.1 Length: 1965928
Range 1: 889572 to 890191

Score:265 bits(143), Expect:1e-65,
Identities:469/628(75%), Gaps:16/628(2%), Strand: Plus/Minus

Enterococcus faecalis V583 3 AACATTTGTGATTGGATAAAAAGGAATCCAAATTTGTCCGCATGGCTATGTCCTGCTAA 62
Lactococcus garvieae JJJN1 890191 AACATTTTTTATCGGGTAGAAGGGAATCCAGACCTGACCACCGTGACTATGTCCTGCTAA 890132
Enterococcus faecalis V583 63 AACTAACTGTGTACCCGTACCAACAAATGCATCCGCGACATCTGGTTCATGGGTCATTAA 122
Lactococcus garvieae JJJN1 890131 AACCAACTGTGTATCTGTACTAATAAAAAGCATCCGCAACATCAGGTTTCATGTGTCAT--GG 890073
Enterococcus faecalis V583 123 TAA--CGCATAGTCATACGTTTCTCGATAGGATAATGTTTCTGAAACGGAAGGATTTCCCA 181
Lactococcus garvieae JJJN1 890072 TAATGGCATAATCATAGTTTTGGCGTAAGCAAGCGTATCGGATACCGAAGGATTTCTTA 890013
Enterococcus faecalis V583 182 ATAACGAGTCATCTAAACCACCAATAAAAAGAC--GGCGTCCATCTGCT--AAGGTTACTG 238
Lactococcus garvieae JJJN1 890012 AAAGAGAGTCATCGAGTCCACC--TAAGAAAACCTGACGACCGTC--GCTCAATGTTAGTG 889956
Enterococcus faecalis V583 239 TTTCTCCTTGGTTTTTTCATAACGGTAAAGCCACTGGCCTCCATGACTTCATCATACACCG 298
Lactococcus garvieae JJJN1 889955 TTTCTCCTCGTTCCTTAGCAGCGTAAATCCACCCGCTGACATGACATCTTCATAGAC--G 889897
Enterococcus faecalis V583 299 CAGA--AGCACCGCCACCATAATCATGATTTCCCAAATGCATATTTTCC--A--ATTGTGCG 355
Lactococcus garvieae JJJN1 889896 CGGACTGCGCCACCACCATAATCATGATTGCCCCAAACGGCATACTTTCCGATATTGGC-- 889838
Enterococcus faecalis V583 356 CTTGTAAGTGTTTTAACTTTTCAATCATCGGTTCTTTTTGTTGGGATTTTGGCATAAT 415
Lactococcus garvieae JJJN1 889837 CTTA--AAGGCTTTTAAATTTTCAATCATTGGTGCCCTTTGTTCTGGATACTTGGAGTAAT 889779
Enterococcus faecalis V583 416 TGTCAAATAAATCACCTGTA AAAACGATAATATCTGGTTTTTTCACGATTGACttttttta 475
Lactococcus garvieae JJJN1 889778 TATCGAACAAATCCCCTGTA AACAGGAGAATATCTGGTTTTTGGGCATTGACCTTCTCAA 889719
Enterococcus faecalis V583 476 TGACTTTATCTAACCGCTTCGTCTCATAAAATTCTGAAACTTGATATCCGAAAGTTGCA 535
Lactococcus garvieae JJJN1 889718 TCACCTTATCTAAACGATTGTTTCATAAATACTCGGAAACCTGCAATCAGAGATTGTA 889659
Enterococcus faecalis V583 536 CAACATTCAATGGTTCTCGTTGATTTTTCTTTCCCAAAGTCACC--TGATTAACGTGAACT 594
Lactococcus garvieae JJJN1 889658 CAACATTTAGTGGTGTTCGCTCATTTTTCTTCCCA--GTTCAACATGATTGACATGCACA 889600
Enterococcus faecalis V583 595 AAATTCGGCTCAATCTTCCACGCATAGA 622
Lactococcus garvieae JJJN1 889599 AGGAAGGGCTCAACTTTCCAAGCATAGA 889572

```

Figura A.3: Alineamiento de ARN nc3340 en *Lactococcus garvieae* JJJN1. La figura muestra el alineamiento de nucleótidos del transcrito nc3340 en el genoma de *Lactococcus garvieae* JJJN1. También se entrega información sobre el ID de la secuencia, el porcentaje de identidad del alineamiento y puntaje entregado por BLASTN.


```

>Enterococcus sp. LX10 chromosome, complete genome
Sequence ID: CP092784.1 Length: 2737931
Range 1: 1088250 to 1088958

Score:1271 bits(688), Expect:0.0,
Identities:702/709(99%), Gaps:0/709(0%), Strand: Plus/Minus

Enterococcus faecalis V583 1      AAAACATTTGTGATTGGATAAAAAGGAATCCAATTTGTCGCGCATGGCTATGTCCTGCT 60
Enterococcus sp. LX10      1088958 AAAACATTTATGATTGGATAAAAAGGAATCCAATTTGTCGCGCATGGCTATGTCCTGCT 1088899
Enterococcus faecalis V583 61      AAAACTAACTGTGTACCCGTACCAACAATGCATCCGCGACATCTGGTTCATGGGTCATT 120
Enterococcus sp. LX10      1088898 AAAACTAACTGTGTACCCGTACCAATAAATGCATCCGCGACATCTGGTTCATGGGTCATT 1088839
Enterococcus faecalis V583 121     AATAACGCATAGTCATACGTTTCTCGATAGGATAATGTTTCTGAAACGGAAGGATTTCCC 180
Enterococcus sp. LX10      1088838 AATAACGCATAGTCATACGTTTCTCGATAGGATAATGTTTCTGAAACGGAAGGATTTCCCT 1088779
Enterococcus faecalis V583 181     AATAACGAGTCATCTAAACCACCAATAAAAAGACGGCGTCCATCTGCTAAGGTTACTGTT 240
Enterococcus sp. LX10      1088778 AATAACGAGTCATCTAAACCACCAATAAAAAGACGGCGTCCATCTGCTAAGGTTACTGTT 1088719
Enterococcus faecalis V583 241     TCTCCTTGGTTTTTTCATAACGGTAAAGCCACTGGCCTCCATGACTTCATCATAACCCGCA 300
Enterococcus sp. LX10      1088718 TCTCCTTGGTTTTTTCATAACGGTAAAGCCACTGGCCTCCATGACTTCATCATAACCCCA 1088659
Enterococcus faecalis V583 301     GAAGCACCGCCACCATAATCATGATTTCCCAAACCTGCATATTTTCCAATTTGTCGCTTGT 360
Enterococcus sp. LX10      1088658 GAAGCACCGCCACCATAATCATGATTTCCCAAACCTGCATATTTTCCAATTTGTCGCTTGC 1088599
Enterococcus faecalis V583 361     AACTGTTTTAACTTTTCAATCATCGGTTCTTTTTGTTGCGGATTTTTGGCATAATTGTCA 420
Enterococcus sp. LX10      1088598 AACTGTTTTAACTTTTCAATCATCGGTTCTTTTTGTTGCGGATTTTTGGCATAATTGTCA 1088539
Enterococcus faecalis V583 421     AATAAATCACCTGTAAAAACGATAAATATCTGGTTTTTTCACGATTGACTtttttttATGACT 480
Enterococcus sp. LX10      1088538 AATAAATCACCTGTAAAAACGATAAATATCTGGTTTTTTCACGATTGACTTTTTTTATGACT 1088479
Enterococcus faecalis V583 481     TTATCTAACCGCTTCGTCCTATAAAATCTGAAACTTGTATATCCGAAAGTTGCACAACA 540
Enterococcus sp. LX10      1088478 TTATCTAACCGCTTCGTCCTATAAAATCTGAAACTTGTATATCCGAAAGTTGCACAACA 1088419
Enterococcus faecalis V583 541     TTCAATGGTTCTCGTTGATTTTTCTTTCCCAAAGTCACCTGATTAACGTGAACTAAATTC 600
Enterococcus sp. LX10      1088418 TTCAATGGTTCTCGTTGATTTTTCTTTCCCAAAGTCACCTGATTAACGTGAACTAAATTC 1088359
Enterococcus faecalis V583 601     GGCTCAATCTTCCACGCATAGATTGGGAGTGCATAATTAAGaaaaaaaTAAGAGCATT 660
Enterococcus sp. LX10      1088358 GGCTCAATCTTCCACGCATAGATTGGGAGTGCATAATTAAGAAAAAAAAAATAGAGCATT 1088299
Enterococcus faecalis V583 661     ATACTTAAACCAAttttttAACCACATCACTCGCTCCTTTCTTTCTT 709
Enterococcus sp. LX10      1088298 ATACTTAAACCAATTTTTTAAACCACATCACTCGCTCCTTTCTTTCTT 1088250

```

Figura A.5: Alineamiento de ARN nc3340 en *Lactococcus garvieae* JJJN1. La figura muestra el alineamiento de nucleótidos del transcrito nc3340 en el genoma de *Lactococcus garvieae* JJJN1. También se entrega información sobre el ID de la secuencia, el porcentaje de identidad del alineamiento y puntaje entregado por BLASTN.

```

>Enterococcus sp. LX10 chromosome, complete genome
Sequence ID: CP092784.1 Length: 2737931
Range 1: 1241893 to 1242716

Score:1456 bits(788), Expect:0.0,
Identities:812/824(99%), Gaps:0/824(0%), Strand: Plus/Minus

Enterococcus faecalis V583 1 TTAGACCCCGTCTGTACGCGGATACCATCTGTCAAAGCTAAACGTTGTGAATAAGAAACG 60
Enterococcus sp. LX10 1242716 TTAGACCCCGTCTGTCACTGCGATACCATCTGTCAAAGCTAAACGTTGTGAATAAGAAACG 1242657

Enterococcus faecalis V583 61 TAGCCATAATAAACATCAACACCGCCGCCATTGTTCCAAGTAACAGTGATTTTAACATT 120
Enterococcus sp. LX10 1242656 TAGCCATAATAAACATCAACACCGCCGCCATTGTTCCAAGTAACAGTGATTTTAACATT 1242597

Enterococcus faecalis V583 121 GGAGTAATCGATTTTGTAGCTGGTTGGGTCGTCGGGCGATAACTAGTCGTATCATAACT 180
Enterococcus sp. LX10 1242596 GGAGTAATCGATTTTGTAGCTGGTTGGGTCGTCGGGCGATAATTAGTCGTATCATAACT 1242537

Enterococcus faecalis V583 181 AAAATAGTGCCCGAAGCTCCACTTATTTGGCTTTTCCTGTGTAATTTGCGCCATTCGTA 240
Enterococcus sp. LX10 1242536 AAAATAGTGCCCGAAGCTCCACTTATTTGGCTTTTCCTGTGTAATTTGCGCCATTCGTA 1242477

Enterococcus faecalis V583 241 TTCGCTGCGTGACCGCTCCATTGCTTCAATGATTAACCCGCTTGCTTTTGATCTCCT 300
Enterococcus sp. LX10 1242476 TTCGCTGCGTGACCGCTCCATTGCTTCAATGATTAACCCGCTTGCTTTTGATCTCCT 1242417

Enterococcus faecalis V583 301 CGTGGATAATAAAAAAGATACGTGGTTGCCATTAGCAAAATTCATCTGGAATTGGTTTA 360
Enterococcus sp. LX10 1242416 CGTGGATAATAAAAAAGATACGTGGTTGCCATTAGCAAAATTCATCTGGAATTGGTTTA 1242357

Enterococcus faecalis V583 361 AGTGATCCC GTTGGTGCCGACTCGACTTCGTTTGCTTTTCAACCGGACTTCTTCTGT 420
Enterococcus sp. LX10 1242356 AGTGATCCC GTTGGTGCCGACTCGACTTCGTTTGCTTTTCAACCGGACTTCTTCTGT 1242297

Enterococcus faecalis V583 421 CTTTTTAAGCCTTCACTATTTTGATTATTTAGTTTCTCAACCGTTTCTTAAACGATCGAA 480
Enterococcus sp. LX10 1242296 CTTTTTAAGCCTTCACTATTTTGATTATTTAGTTTCTCAACCGTTTCTTAAACGATTGAA 1242237

Enterococcus faecalis V583 481 TAATGAGGTTTCCCTAAATCGCTAACAGGGATCGTAGCTAATTTTTCTTCTGATTCTTTA 540
Enterococcus sp. LX10 1242236 TAATGAGGTTTCCCTAAATCGCTAACAGGAATCGTAGCTAATTTTTCTTCTGATTCTTTA 1242177

Enterococcus faecalis V583 541 TATTGCCCTCTGCCGTTAACTTTTTAGCTTCTGTAATAATGGCATCATACTTGGTTTGG 600
Enterococcus sp. LX10 1242176 TATTGCCCTCTGCCGTTAACTTTTTAGCTTCTGTAATAATGGCATCATACTTGGTTTGG 1242117

Enterococcus faecalis V583 601 TCTACGGAAGAGGATGTGCTTTGTCCCGTAGCTTCGACTTGTTTGTGACTGGTTTTT 660
Enterococcus sp. LX10 1242116 TCTACGGAAGAGGATGTGCTTTGTCCCGTAGCTTCGACTTGTTTGTGACTGGTTTTT 1242057

Enterococcus faecalis V583 661 TCTGTGGTCACGTCTGATGAATTTCCACAAGCACTTACAAACATTGAGCTAACGATTAAT 720
Enterococcus sp. LX10 1242056 TCCGTGGTCACGTCTGACGAACTTCCACAAGCACTTACAAACATTGAGCTAACGATTAAT 1241997

Enterococcus faecalis V583 721 AAACATCCCACAATCACATGCTTTTTCATTTTAAATCCCCCATTTGTCTATCTTTGTTA 780
Enterococcus sp. LX10 1241996 AAACATCCCACAATCACATGCTTTTTCATCTTTAAATCCCCCATTTGTTTATCTTTGTTA 1241937

Enterococcus faecalis V583 781 TTTTTGTACTTAAATTATCCCATAGATGAACAAAAGAAACCA 824
Enterococcus sp. LX10 1241936 TTTTTGTACTTAAATTATCCCATAGATGAACAAAAGAAACCA 1241893

```

Figura A.6: Alineamiento de ARN nc3683 en *Enterococcus* sp. LX10. La figura muestra el alineamiento de nucleótidos del transcrito nc3683 en el genoma de *Enterococcus* sp. LX10. También se entrega información sobre el ID de la secuencia, el porcentaje de identidad del alineamiento y puntaje entregado por BlastN.

```

>Enterococcus sp. PLM3 DNA, complete genome
Sequence ID: AP025690.1 Length: 2859161
Range 1: 1243168 to 1243991

Score:1456 bits(788), Expect:0.0,
Identities:812/824(99%), Gaps:0/824(0%), Strand: Plus/Minus

Enterococcus faecalis V583 1      TTGACCCCGTCTGTACGGCGATACCATCTGTCAAAGCTAAACGTTGTGAATAAGAAACG 60
Enterococcus sp. PLM3      1243991 TTGACCCCGTCTGTACGGCGATACCATCTGTCAAAGCTAAACGTTGTGAATAAGAAACG 1243932
Enterococcus faecalis V583 61      TAGCCATAATAAACATCAACACCGCCGCGCATTGTTCCAAGTAACAGTGATTTAACATTT 120
Enterococcus sp. PLM3      1243931 TAGCCATAATAAACATCAACACCGCCGCGCATTGTTCCAAGTAACAGTGATTTAACATTT 1243872
Enterococcus faecalis V583 121     GGAGTAATCGATTTTGTAGCTGGTGGGTCGTCGGGCGATAACTAGTCGTATCATAACTT 180
Enterococcus sp. PLM3      1243871 GGAGTAATCGATTTTGTAGCTGGTGGGTCGTCGGGCGATAACTAGTCGTATCATAACTT 1243812
Enterococcus faecalis V583 181     AAAATAGTGCCCGAAGCTCCACTTATTTCCGGCTTTTCCTGTGTAATTTGCCCATTCGTA 240
Enterococcus sp. PLM3      1243811 AAAATAGTGCCCGAAGCTCCACTTATTTCCGGCTTTTCCTGTGTAATTTGCCCATTCGTA 1243752
Enterococcus faecalis V583 241     TTCGTCTGCGTGACCGCTCCATTGCTTCAATGATTAACCCCGCTTGCTTTTGATCTCCT 300
Enterococcus sp. PLM3      1243751 TTCGTCTGCGTGACTGCTCCATTGCTTCAATGATTAACCCCGCTTGCTTTTGATCTCCT 1243692
Enterococcus faecalis V583 301     CGTGGATAATAAAAAAGATACGTGGTTGCCCATTTAGCAAATTCATCTGGAATGGTTTA 360
Enterococcus sp. PLM3      1243691 CGTGGATAATAAAAAAGATACGTGGTTGCCCATTTAGCAAATTCATCTGGAATGGTTTA 1243632
Enterococcus faecalis V583 361     AGTGATCCCGTTGGTGCCGGACTCGACTTCGTTTGCTTTTCAACCGCGACTTCTTCTTGT 420
Enterococcus sp. PLM3      1243631 AGTGATCCCGTTGGTGCCGGACTCGACTTCGTTTGCTTTTCAACCGCGACTTCTTCTTGT 1243572
Enterococcus faecalis V583 421     CTTTTAAGCCTTCACTATTTTGATTATTTAGTTTCTCAACCGTTTCTTAAACGATCGAA 480
Enterococcus sp. PLM3      1243571 CTTTTAAGCCTTCACTATTTTGATTATTTAGTTTCTCAACCGTTTCTTAAACGATCGAA 1243512
Enterococcus faecalis V583 481     TAATGAGGTTTCCCTAAATCGCTAACAGGGATCGTAGCTAATTTTTCTTCTGATCTTTA 540
Enterococcus sp. PLM3      1243511 TAGTGAGGTTTCCCTAAATCGCTAACAGGGATCGTAGCTAATTTTTCTTCTGATCTTTA 1243452
Enterococcus faecalis V583 541     TATTGCCCTCTGCCGTTAACTTTTTAGCTTCTGTAATAATGGCATCATACTTGGTTTGG 600
Enterococcus sp. PLM3      1243451 TATTGCCCTCTGCCGTTAACTTTTTAGCTTCTGTAATAATGGCATCATACTTGGTTTGG 1243392
Enterococcus faecalis V583 601     TCTACGGAAGAGGATGTGCTTTGTCGCCGAGCTTCGACTTGTTTTGTGACTGGTTTTT 660
Enterococcus sp. PLM3      1243391 TCTACGGAAGAGGATGTGCTTTGTCGCCGAGCTTCGACTTGTTTTGTGACTGGTTTTT 1243332
Enterococcus faecalis V583 661     TCTGTGGTACGCTGTATGAATTTCCACAAGCACTTACAAACATTGAGCTAACGATTAAT 720
Enterococcus sp. PLM3      1243331 TCCGTGGTACGCTGTAGCAATTTCCACAAGCACTTACAAACATTGAGCTAACGATTAAT 1243272
Enterococcus faecalis V583 721     AAACATCCACAATCACATGCTTTTTTCAATTTTAAATCCCCCATTGGTCTATCTTTGTTA 780
Enterococcus sp. PLM3      1243271 AAACATCCACAATCACATGCTTTTTTCAATTTTAAATCCCCCATTGGTCTATCTTTGTTA 1243212
Enterococcus faecalis V583 781     TTTTTTGTTACTTAAATTATCCCATAGATGAACAAAAGAAACCA 824
Enterococcus sp. PLM3      1243211 TTTTTTGTTACTTAAATTATCCCATAGATGAACAAAAGAAACCA 1243168

```

Figura A.7: Alineamiento de ARN nc3683 en *Enterococcus* sp. PLM3. La figura muestra el alineamiento de nucleótidos del transcrito nc3683 en el genoma de *Enterococcus* sp. PLM3. También se entrega información sobre el ID de la secuencia, el porcentaje de identidad del alineamiento y puntraje entregado por BlastN.

Además, mediante el uso del programa LocARNA se logró determinar la estructura secundaria conservada de los ARNs no codificantes según aquellos nucleótidos conservados. Dichas estructuras se muestran en las Figuras 5.4.A y 5.4.B en la sección de Resultados, mientras que las Figuras A.8 y A.9 exhiben la conservación determinada por LocARNA.

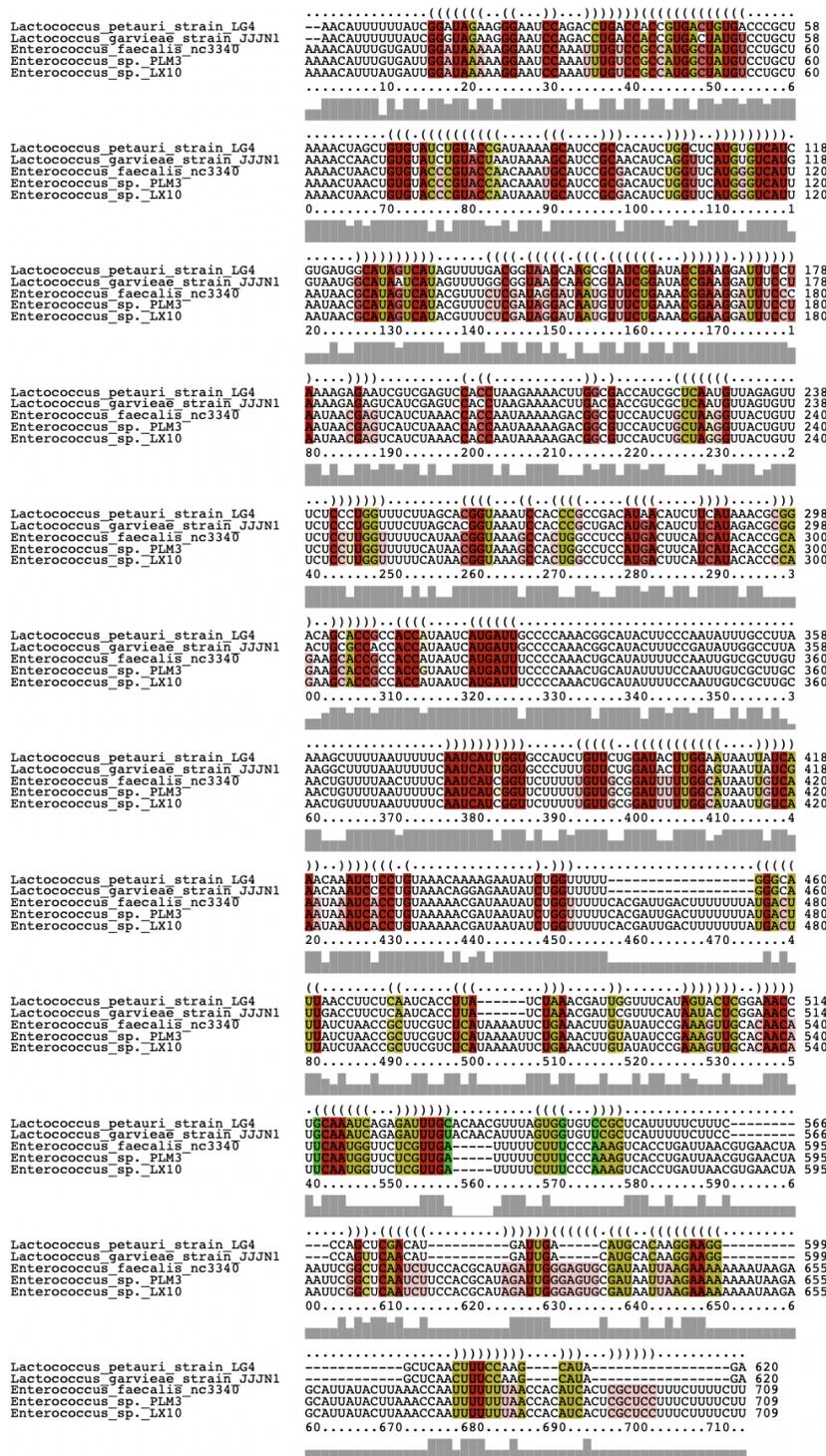


Figura A.8: Conservación de la secuencia del ARN nc3340 en varias especies. La figura enseña la conservación de nucleótidos en las especies *Lactococcus petauri* LG4, *Lactococcus garvieae* JJJN1, *Enterococcus* sp. PLM3 y *Enterococcus* sp. LX10. Los colores indican las bases conservadas y compatibles en la estructura secundaria. La saturación decrece con el número de bases incompatibles.

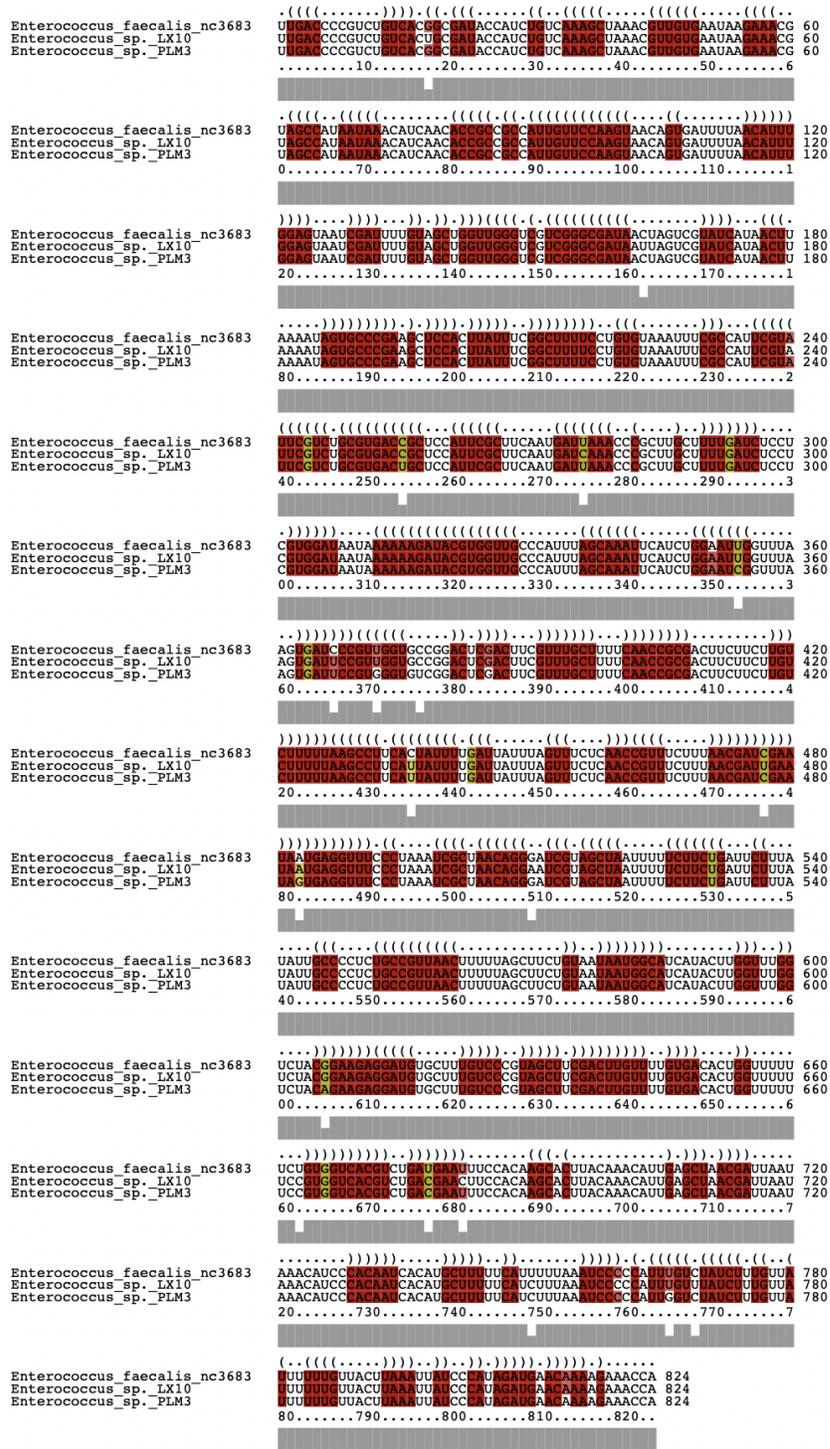


Figura A.9: Conservación de la secuencia del ARN nc3683 en varias especies. La figura enseña la conservación de nucleótidos en las especies *Enterococcus* sp. LX10 Y *Enterococcus* sp. PLM3. Los colores indican las bases conservadas y compatibles en la estructura secundaria. La saturación decrece con el número de bases incompatibles.

Anexo B. Validación de abundancia de los ARNnc

B.1. Diseño de partidores o *primers*

A partir de los resultados del análisis *in silico*, se obtuvo ARNs no codificantes cuya secuencia es conocida a partir de las coordenadas de inicio y fin, buscadas en el genoma de la cepa de *E. faecalis* V583 disponible en bases de datos. Utilizando dicha secuencia, es posible diseñar los partidores o *primers* para la síntesis del ADNc por retrotranscripción y qPCR.

Los partidores se pueden diseñar de forma manual o utilizando alguno de los softwares disponibles en línea. En este caso, se utilizó la página web [Primer3Plus](#), donde los pasos a seguir dependen de si el partidador es para ADNc o qPCR.

En el caso de qPCR, se incerta la secuencia del ARN no codificante en el recuadro disponible en la pestaña principal o *Main* y se cambia la configuración del servidor en *Load server settings* por qPCR. En la pestaña de configuraciones generales o *General settings* se configura el o los rangos del tamaño del producto deseado (*Product Size Ranges*), que en este caso fue entre 100 y 200 bp. Esto quiere decir que lo que se sintetice entre los partidores diseñados será de un largo entre 100 y 200 bp. Los demás valores se dejan sin cambios, por defecto, y finalmente se hace click en *Pick Primers* para que muestre los resultados de los partidores escogidos.

Es importante tener en cuenta los valores de ciertos parámetros entregados por los programas para verificar que sean adecuados para la bacteria con la que se está trabajando. Algunos de los filtros utilizados en la síntesis de los partidores fueron:

- Temperatura de melting (T_m) $\approx 60^\circ\text{C}$.
- *Product Size Ranges* ≈ 140 pb.
- Largo de partidores igual a 20 pb.
- Verificar que los partidores no formen estructuras secundarias que dificulten el alineamiento con el templado.
- Confirmar que la secuencia de los partidores hibrida únicamente en un segmento del genoma.

Los partidores del ADNc de hebra simple fueron diseñados manualmente, tomando las últimas 20 bases de la secuencia de los ARN no codificantes en estudio. Si los ARNs están en dirección 5'-3', entonces los *primers* serán de 3' a 5' y complementarios las últimas 20 bases de cada ARN.

Las secuencias de bases nitrogenadas de los 2 ARNnc seleccionados para estudiar su expresión diferencial son las que se muestran en la Tabla B.1.

Tabla B.1: Secuencia de los ARN no codificantes en estudio.

ARNnc	Secuencia
nc3340	AAAACAUUUGUGAUUGGAUAAAAAGGAAUCCAAAUUUGUCCGC CAUGGCUAUGUCCUGCUAAAACUAACUGUGUACCCGUACCAAC AAAUGCAUCCGCGACAUCUGGUUCAUGGGUCAUUAUAACGCA UAGUCAUACGUUUCUCGAUAGGAUAAUGUUUCUGAAACGGAAG GAUUUCCCAAUAACGAGUCAUCUAAACCACCAAUAAAAAGACG GCGUCCAUCUGCUAAGGUUACUGUUUCUCCUUGGUUUUUAUA ACGGUAAGCCACUGGCCUCCAUGACUUCAUCAUACACCGCAG AAGCACCGCCACCAUAAUCAUGAUUUUCCCAAACUGCAUAUUU UCCAAUUGUCGCUUGUAACUGUUUUAACUUUCAAUCAUCGGU UCUUUUUGUUGCGGAUUUUUGCAUAAUUGUCAAAUAAAUCAC CUGUAAAAACGAUAAUAUCUGGUUUUUCACGAUUGACUUUUU UAUGACUUUAUCUAACCGCUUCGUCUCAUAAAAUUCUGAAACU UGUUAUCCGAAAGUUGCACAACAUUCAUUGGUUCUCGUUGAU UUUUCUUUCCCAAAGUCACCUGAUUAAACGUGAACUAAAUCGG CUCAAUCUUCACGCAUAGAUUGGGAGUGCGAUAAUUAAGAAA AAAAAUAAGAGCAUUAUACUAAAACCAAUUUUUUUAACCACAU CACUCGCUCUUUCUUUUUCU
nc3683	UUGACCCCGUCUGUCACGGCGAUACCAUCUGUCAAAAGCUAAAC GUUGUGAAUAAGAAACGUAGCCAUAUAUAACAUCAACACCGCC GCCAUUGUCCAAGUAACAGUGAUUUUAACAUUUGGAGUAAUC GAUUUUGUAGCUGGUUGGGUCGUCGGGCGAUAAACUAGUCGUAU CAUAACUAAAAUAGUGCCCGAAGCUCACUUAUUUCGGCUUU UCCUGUGUAAAAUUCGCCAUUCGUAUUCGUCUGCGUGACCGCU CCAUUCGCUUCAUGAUUAAACCCGCUUGCUUUUGAUCUCCUC GUGGAUAAUAAAAAAGAUACGUGGUUGCCCAUUUAGCAAUUC AUCUGGAAUUGGUUUAAGUGAUCCCGUUGGUGCCGGACUCGAC UUCGUUUGCUUUUCAACCGCGACUUCUUCUUGUCUUUUUAAGC CUUCACUAUUUUGAUUAUUUAGUUUCUCAACCGUUUCUUUAAC GAUCGAAUAAUGAGGUUUCUCAAUAUCGCUAACAGGGAUCGUA GCUAAUUUUUCUUCUGAUUCUUUAUAUUGCCCCUCUGCCGUUA ACUUUUUAGCUUCUGUAAUAAUGGCAUCAUACUUGGUUUGGUC UACGGAAGAGGAUGUGCUUUGUCCCGUAGCUUCGACUUGUUUU GUGACACUGGUUUUUUCUGUGGUCACGUCUGAUGAAUUUCCAC AAGCACUACAAACAUUGAGCUAACGAUUAUAUAACAUCACAC AAUCACAUGCUUUUUCAUUUUUAAAUCCCCAAUUGUCUAUCU UUGUUAUUUUUGUUACUAAAUAUCCCAUAGAUGAACAAAA GAAACCA