



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Identificación *in silico* y caracterización molecular de la regulación genética basada en riboswitches que se transcriben en sentido opuesto al de su gen blanco (*antisense-acting* riboswitches)

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctora en Ciencias

PRESENTA:

M. en C. Mariela Serrano Gutiérrez

TUTOR PRINCIPAL

[Dr. Enrique Merino Pérez IBT UNAM](#)

MIEMBROS DEL COMITÉ TUTOR

[Dr. Mario Soberón Chávez IBT UNAM](#)

[Dra. Rosa Laura Camarena Mejía IIBO UNAM](#)

Ciudad de México. Enero, 2024.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

El presente trabajo de investigación se realizó en el departamento de Microbiología Molecular del Instituto de Biotecnología de la UNAM, bajo la dirección del Dr. Enrique Merino Pérez y como parte del Programa de Maestría y Doctorado en Ciencias Bioquímicas de la misma Universidad.

El Comité Tutorial estuvo integrado por:

Dr. Enrique Merino Pérez (Instituto de Biotecnología, UNAM)

Dra. Rosa Laura Camarena Mejía (Instituto de Investigaciones Biomédicas, UNAM)

Dr. Mario Soberón Chávez (Instituto de Biotecnología, UNAM)

El jurado designado estuvo integrado por:

Dra. Gloria Soberón Chávez (Instituto de Investigaciones Biomédicas, UNAM)

Dra. Katy Juárez López (Instituto de Biotecnología, UNAM)

Dr. Jose Arcadio Farías Rico (Centro de Ciencias Genómicas, UNAM)

Dr. Dimitris Georgellis (Instituto de Fisiología Celular, UNAM)

Dr. Daniel Genaro Segura González (Instituto de Biotecnología, UNAM)

Para la realización de esta investigación, el alumno recibió una beca CONACyT (No. CVU 702475).

Este estudio fue financiado con el donativo CONACYT no. 2040 y el apoyo DGAPA-UNAM PAPIIT no. IN202120.

A mi madre, por ser mi inspiración, y enseñarme que la disciplina y el trabajo duro siempre dan frutos.

A Luis, por ser mi apoyo y nunca soltar mi mano sin importar lo complicado que ha sido el camino.

AGRADECIMIENTOS

Al Instituto de Biotecnología y la Universidad Nacional Autónoma de México por permitirme formar parte de ustedes.

Agradezo profundamente al Dr. Enrique Merino Pérez por abrirme las puertas de su grupo de trabajo, por su gran apoyo y confianza incondicional, por ser una pieza clave en mi desarrollo como profesionista. Gracias por estar siempre dispuesto a ayudar y creer en mí.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías por la beca de doctorado que hizo posible la realización de este proyecto.

Al Programa de Apoyo a los Estudios de Posgrado por el apoyo económico otorgado para la estancia al XXXIII Congreso Nacional de Bioquímica de la Sociedad Mexicana de Bioquímica.

A la M.B. María Luisa Tabche Barrera por todas sus enseñanzas, por trabajar a mi lado y por su apoyo incondicional.

A todos mis compañeros de laboratorio, gracias por recibirme con los brazos abiertos y brindarme su amistad.

A mis amigos, por siempre creer en mí. Gracias por brindarme su amistad incondicional.

ÍNDICE

RESUMEN	1
ABSTRACT	3
INTRODUCCIÓN	5
RIBOSWITCHES.....	8
ESTRUCTURA DE LOS RIBOSWITCHES	10
ANTECEDENTES	19
INTERFERENCIA TRANSCRIPCIONAL	19
RIBOSWITCHES ANTISENTIDO (<i>ANTISENSE-ACTING</i> RIBOSWITCHES)	20
HIPÓTESIS	23
OBJETIVO	24
OBJETIVOS PARTICULARES	24
METODOLOGÍA	25
IDENTIFICACIÓN <i>IN SILICO</i> DE RIBOSWITCHES EN LOS GENOMAS DE BACTERIAS Y ARQUEAS DISPONIBLES	25
SELECCIÓN DE ORGANISMOS REPRESENTATIVOS	26
ANÁLISIS ESTADÍSTICO DE LOS RESULTADOS.....	26
ANÁLISIS DEL CONTEXTO GENÓMICO	27
ANÁLISIS DE LAS RUTAS METABÓLICAS	27
ANÁLISIS DE LA FUNCIÓN DE LOS GENES	28

ANÁLISIS DEL CONTEXTO SINTÉNICO DE LOS GENES RELACIONADOS CON RIBOSWITCHES	28
RESULTADOS.....	30
IDENTIFICACIÓN <i>IN SILICO</i> DE RIBOSWITCHES EN EL GENOMA DE BACTERIAS Y ARQUEAS	30
CLASIFICACIÓN DE TIPOS DE GENES REGULADOS POR RIBOSWITCHES ANTISENTIDO ...	37
APÉNDICE 1. LOCALIZACIÓN DE PROBABLES ERRORES DE ANOTACIÓN EN EL GENOMA DE BACTERIAS Y ARQUEAS BASADO EN LA IDENTIFICACIÓN DE RIBOSWITCHES EN REGIONES CODIFICANTES	61
DISCUSIÓN.....	72
CONCLUSIONES.....	78
PERSPECTIVAS.....	79
REFERENCIAS.....	80
ARTÍCULO PUBLICADO DURANTE EL DOCTORADO	89

ÍNDICE DE FIGURAS

FIGURA 1. MECANISMO DE REGULACIÓN MEDIADO POR THI-BOX PROPUESTO PARA EL OPERÓN THICOGE (TOMADO DE J. MIRANDA-RÍOS ET AL., 2001).....	8
FIGURA 2. MECANISMO DE REGULACIÓN BASADO EN RIBOSWITCHES (MODIFICADO DE SINUMVAYO ET AL., 2018).	9
FIGURA 3. LOCALIZACIÓN TÍPICA DEL RIBOSWITCH EN EL GENOMA (MODIFICADO DE MEHDIZADEH AGHDAM ET AL., 2016).....	10
FIGURA 4. ESTRUCTURA Y DOMINIOS QUE COMPONEN EL RIBOSWITCH (MODIFICADO DE EDWARDS & BATEY, 2010).....	11
FIGURA 5. REARREGLO ALOSTÉRICO DEL DOMINIO APTÁMERO EN RESPUESTA A LA UNIÓN DE SU LIGANDO (MODIFICADO DE W. C. WINKLER & BREAKER, 2005).....	12
FIGURA 6. PRINCIPALES MECANISMOS DE REGULACIÓN MEDIADOS POR RIBOSWITCH: (A) TERMINACIÓN DE LA TRANSCRIPCIÓN, (B) INICIO DE LA TRADUCCIÓN (MODIFICADO DE BREAKER, 2012).	15
FIGURA 7. MECANISMO DE ATENUACIÓN TRANSCRIPCIONAL MEDIADO POR RIBOSWITCH: (A) METABOLITO (M) UNIDO AL DOMINIO APTÁMERO, (B) SIN UNIÓN DE METABOLITO (MODIFICADO DE SHERWOOD & HENKIN, 2016).....	18
FIGURA 8. MODELOS PROPUESTOS DE LOS MECANISMOS REGULATORIOS DE LAS RNA POLIMERASAS QUE TRANSCRIBEN EN DIRECCIONES OPUESTAS: A. COLISIÓN DE RNA POLIMERASAS TRANSCRIBIENDO EN DIRECCIONES OPUESTAS, B. FORMACIÓN DE DÚPLEX (1)MRNA-(2)ASRNA, C. SUPERENRROLLAMIENTO (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	20
FIGURA 9. MECANISMO DE REGULACIÓN DEL RIBOSWITCH ANTISENTIDO S-BOX QUE REGULA EL OPERÓN UBIG/MCCAB EN C. ACETOBUTYLICUM (TOMADO DE ANDRÉ ET AL., 2008).....	21

FIGURA 10. MECANISMO DE REGULACIÓN DEL RIBOSWITCH ANTISENTIDO DE COBALAMINA QUE REGULA AL REGULADOR TRANSCRIPCIONAL POCHR EN <i>L. MONOCYTOGENES</i> (TOMADO DE MELLIN ET AL., 2013).	22
FIGURA 11. DIAGRAMA GENERAL DE TRABAJO.	29
FIGURA 12. DISTRIBUCIÓN POR PHYLUM Y FAMILIA DE LOS RIBOSWITCHES EN SENTIDO IDENTIFICADOS EN EL GENOMA DE BACTERIAS Y ARQUEAS	31
FIGURA 13. DISTRIBUCIÓN POR PHYLUM Y FAMILIA DE LOS RIBOSWITCHES ANTISENTIDO IDENTIFICADOS EN EL GENOMA DE BACTERIAS Y ARQUEAS.	36
FIGURA 14. GEN CODIFICANTE DE UNA ENZIMA INVOLUCRADA EN LA INTERCONVERSIÓN DE UN COMPUESTO RELACIONADO AL METABOLITO SENSADO POR EL RIBOSWITCH (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	39
FIGURA 15. GEN QUE CODIFICA PARA UNA ENZIMA QUE REQUIERE UN COFACTOR SENSADO POR EL RIBOSWITCH PARA ACTIVARSE O QUE PARTICIPA EN UNA RUTA METABÓLICA EN LA CUAL OTRA ENZIMA REQUIERE DE DICHO COFACTOR (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	42
FIGURA 16. GEN CODIFICANTE DE UNA PROTEÍNA TRANSPORTADORA DE UN COMPUESTO RELACIONADO AL METABOLITO SENSADO POR EL RIBOSWITCH (TOMADO DE SERRANO- GUTIÉRREZ & MERINO, 2023).	45
FIGURA 17. GEN QUE CODIFICA PARA UNA ENZIMA INVOLUCRADA EN UN SISTEMA DE SEÑALIZACIÓN CELULAR RELACIONADO CON EL METABOLITO SENSADO POR EL RIBOSWITCH (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	47
FIGURA 18. GEN CODIFICANTE DE UN REGULADOR TRANSCRIPCIONAL RELACIONADO CON EL METABOLITO SENSADO POR EL RIBOSWITCH (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	50
FIGURA 19. GEN CODIFICANTE PARA UNA TRANSPOSASA ADYACENTE A UN RIBOSWITCH ANTISENTIDO (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	54

FIGURA 20. PSEUDOGEN ADYACENTE A UN RIBOSWITCH ANTISENTIDO (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	56
FIGURA 21. GEN CODIFICANTE PARA UN ORF NO CARACTERIZADO ADYACENTE A UN RIBOSWITCH ANTISENTIDO (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	57
FIGURA 22. GEN CODIFICANTE PARA UNA PROTEÍNA HIPOTÉTICA ADYACENTE A UN RIBOSWITCH ANTISENTIDO (TOMADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	59
FIGURA 23. ANÁLISIS DE RIBOSWITCHES CANÓNICOS. (A1) RIBOSWITCH EN SENTIDO, (A2) RIBOSWITCH ANTISENTIDO, (B) PROPORCIÓN DE ASIGNACIÓN COG DE GENES REGULADOS POR RIBOSWITCHES CANÓNICOS, (C) DISTRIBUCIÓN DE TAMAÑO DE LOS GENES REGULADOS POR RIBOSWITCHES CANÓNICOS.	65
FIGURA 24. ANÁLISIS DE RIBOSWITCHES SEMI-INTERNOS. (A1) RIBOSWITCH EN SENTIDO SOBRE EL INICIO DE SU GEN REGULADO, (A2) RIBOSWITCH EN SENTIDO EN EL GEN ADYACENTE A SU GEN REGULADO, (A3) RIBOSWITCH ANTISENTIDO EN EL GEN ADYACENTE A SU GEN REGULADO, (A4) RIBOSWITCH ANTISENTIDO SOBRE EL FIN DE SU GEN REGULADO, (B) PROPORCIÓN DE ASIGNACIÓN COG DE GENES REGULADOS POR RIBOSWITCHES SEMI-INTERNOS, (C) DISTRIBUCIÓN DE TAMAÑO DE LOS GENES REGULADOS POR RIBOSWITCHES SEMI-INTERNOS.	67
FIGURA 25. ANÁLISIS DE RIBOSWITCHES INTERNOS. (A1) RIBOSWITCH EN SENTIDO SOBRE EL INICIO DE SU GEN REGULADO, (A2) RIBOSWITCH EN SENTIDO EN EL GEN ADYACENTE A SU GEN REGULADO, (A3) RIBOSWITCH ANTISENTIDO EN EL GEN ADYACENTE A SU GEN REGULADO, (A4) RIBOSWITCH ANTISENTIDO SOBRE EL FIN DE SU GEN REGULADO, (B) PROPORCIÓN DE ASIGNACIÓN COG DE GENES REGULADOS POR RIBOSWITCHES INTERNOS, (C) DISTRIBUCIÓN DE TAMAÑO DE LOS GENES REGULADOS POR RIBOSWITCHES INTERNOS.....	68
FIGURA 26. ANÁLISIS DE RIBOSWITCH CONTENIENDO UN ORF. (A1) RIBOSWITCH EN SENTIDO EN EL GEN ADYACENTE A SU GEN REGULADO, (A2) RIBOSWITCH ANTISENTIDO EN EL GEN	

ADYACENTE A SU GEN REGULADO, (B) PROPORCIÓN DE ASIGNACIÓN COG DE GENES REGULADOS POR RIBOSWITCHES CONTENIENDO UN ORF, (C) DISTRIBUCIÓN DE TAMAÑO DE LOS GENES REGULADOS POR RIBOSWITCHES CONTENIENDO UN ORF.	69
FIGURA 27. LOCALIZACIÓN DE UN PROBABLE ERROR DE ANOTACIÓN EN EL GENOMA BASADO EN LA IDENTIFICACIÓN DE UN RIBOSWITCH SEMI-INTERNO EN EL INICIO DE SU GEN REGULADO.	70
FIGURA 28. LOCALIZACIÓN DE UN PROBABLE ERROR DE ANOTACIÓN EN EL GENOMA BASADO EN LA IDENTIFICACIÓN DE UN RIBOSWITCH SEMI-INTERNO EN EL FIN DEL GEN ADYACENTE A SU GEN REGULADO.	71
FIGURA 29. LOCALIZACIÓN DE UN PROBABLE ERROR DE ANOTACIÓN EN EL GENOMA BASADO EN LA IDENTIFICACIÓN DE UN RIBOSWITCH SEMI-INTERNO EN EL INICIO DE SU GEN REGULADO Y EL INICIO DEL GEN ADYACENTE AL MISMO.	71

ÍNDICE DE TABLAS

TABLA 1. ORGANIZACIÓN DE LAS CLASES DE RIBOSWITCHES EN FAMILIAS (MODIFICADA DE (SERRANO-GUTIÉRREZ & MERINO, 2023).....	13
TABLA 2. RIBOSWITCHES ANTISENTIDO IDENTIFICADOS EN EL GENOMA DE BACTERIAS Y ARQUEAS (EXTRACTO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	32
TABLA 3. TENDENCIA DE LAS FAMILIAS DE RIBOSWITCHES PARA ACTUAR EN ANTISENTIDO (MODIFICADO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	34
TABLA 4. CLASIFICACIÓN DE LOS GENES REGULADOS POR RIBOSWITCHES ANTISENTIDO (EXTRACTO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	38
TABLA 5. GENES CODIFICANTES PARA TRANSPOSASAS ADYACENTES A RIBOSWITCHES ANTISENTIDO (EXTRACTO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	53
TABLA 6. PSEUDOGENES ADYACENTES A RIBOSWITCHES ANTISENTIDO (EXTRACTO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	55
TABLA 7. GENES CODIFICANTES PARA PROTEÍNAS HIPOTÉTICAS ADYACENTES A RIBOSWITCHES ANTISENTIDO (EXTRACTO DE SERRANO-GUTIÉRREZ & MERINO, 2023).	58

RESUMEN

Los riboswitches son elementos de RNA involucrados en la regulación de genes que participan en la biosíntesis o el transporte de metabolitos esenciales. Estos elementos se caracterizan por su habilidad para reconocer a su molécula blanco con gran afinidad y especificidad. Los riboswitches son comúnmente cotranscritos con su gen blanco y se localizan en el extremo 5' de su unidad transcripcional. A la fecha, únicamente han sido descritos dos casos excepcionales de riboswitches situados en el extremo 3' y transcribiendo en dirección antisentido a su gen regulado. El primer caso involucra un riboswitch de S-Adenosil Metionina (SAM) localizado en el extremo 3' del operón *ubiG-mccB-mccA* en *Clostridium acetobutylicum* involucrado en la interconversión de metionina y cisteína. El segundo caso involucra un riboswitch de cobalamina en *Listeria monocytogenes* que regula el factor transcripcional *PocR* relacionado al proceso patogénico de este organismo. En casi una década a partir de las primeras descripciones de riboswitches antisentido, no se han descrito nuevos ejemplos. En este proyecto, llevamos a cabo un análisis computacional para identificar nuevos ejemplos de riboswitches antisentido que tendrán un efecto negativo en la expresión de los genes río arriba. Se encontraron 292 casos en los que, de acuerdo a la información disponible, se dedujo que la regulación esperada para el riboswitch es consistente con la molécula señal sensada y la función metabólica del gen regulado. Además, las implicaciones metabólicas de este novedoso tipo de regulación fueron minuciosamente discutidas. Por otra parte, durante el análisis de una fracción de los riboswitches antisentido identificados, se encontró evidencia acerca de la existencia

de probables errores de anotación, basada en la observación de regulación de pequeños ORF anotados como codificantes para proteínas hipotéticas. Dado lo cual, y gracias a la búsqueda de riboswitches en regiones codificantes, se propone este método para la identificación tanto de errores en la anotación de genes, como de una mayor cantidad de riboswitches.

ABSTRACT

Riboswitches are RNA elements involved in regulating genes that participate in the biosynthesis or transport of essential metabolites. They are characterized by their ability to recognize their target molecules with high affinity and specificity. Riboswitches are commonly cotranscribed with their target genes and are located at the 5' end of their transcriptional units. To date, only two exceptional cases of riboswitches being situated at the 3' end and transcribing in the antisense direction of their regulated genes have been described. The first case involves a SAM riboswitch located at the 3' end of the *ubiG-mccB-mccA* operon in *Clostridium acetobutylicum* involved in converting methionine to cysteine. The second case concerns a cobalamin riboswitch in *Listeria monocytogenes* that regulates the transcription factor *PocR* related to this organism's pathogenic process. In almost a decade since the first descriptions of antisense-acting riboswitches, no new examples have been described. In this work, we performed a computational analysis to identify new examples of antisense-acting riboswitches that will have a negative effect in the expression of the upstream genes. We found 292 cases in which, according to the available information, we infer that the expected regulation of the riboswitch is consistent with the signaling molecule it senses and the metabolic function of the regulated gene. Besides, the metabolic implications of this novel type of regulation were thoroughly discussed. Moreover, during the analysis of a part of the antisense-acting identified riboswitches, we found evidence on the existence of probable annotation errors, based on the observation of the regulation of small ORFs annotated as codifying to a hypothetical protein. Given so, and thanks to the

riboswitch search in codifying regions, we propose this method for the identification of gene annotation errors, as well as for a greater number of riboswitches.

INTRODUCCIÓN

Todos los organismos poseen una enorme diversidad génica que requiere ser expresada de manera regulada con el fin de mantener un metabolismo exitoso. Debido a esto, todos los organismos hacen uso de una gran variedad de mecanismos de control genético que operan a todos los niveles de la expresión genética, desde la regulación de la transcripción del gen hasta la estabilidad de la proteína traducida. La gran diversidad de mecanismos de regulación permite que los organismos mantengan un estado metabólico controlado en el cual cada célula lleva a cabo un estricto monitoreo de su condición, así como del ambiente que lo rodea. El reconocimiento de señales bioquímicas ocurre a través de receptores proteicos y enzimas que modulan sus funciones en respuesta al cambio en las condiciones intra y extracelulares (Nahvi et al., 2002; Sinumvayo et al., 2018).

En los últimos años, ha habido un gran avance en el estudio de la regulación de la expresión genética mediada por elementos de RNA. Gracias a dichos análisis, se sabe que el RNA puede formar estructuras de plegamiento complejas que actúan como receptores altamente selectivos para diversas clases de compuestos. Además, recientemente se ha establecido que la capacidad de reconocimiento molecular que posee el RNA puede ser aprovechada para el establecimiento de sitios de unión alostérica en interruptores moleculares y, aunado a esto, se conoce que el RNA puede exhibir propiedades biocatalíticas y biosensoras, similares a aquellas propias de las proteínas. Considerando estas características funcionales del RNA, se propuso que el reconocimiento de moléculas de RNA por sus moléculas blanco podría servir como base para un mecanismo de control genético.

Específicamente, se comprobó que la unión entre un metabolito y un RNA mensajero podría provocar cambios estructurales que influyeran en la expresión del RNA sin la necesidad de factores proteicos adicionales (Mironov et al., 2002; W. Winkler et al., 2002).

Resulta de especial importancia destacar que no todo el RNA que es transcrito es traducido a proteína. Aquellos fragmentos de RNA que no codifican proteínas han sido denominadas non-coding RNA (ncRNA) y se sabe que aquellos fragmentos que pertenecen a una misma clase, comparten ciertas características tales como su secuencia y estructura, las cuales han sido conservadas a través de diversos procesos evolutivos. Uno de los ejemplos más significativos de moléculas de RNA funcionales conservadas son los riboswitches (Erdem et al., 2023; L. Waters & Storz, 2009).

A finales de los años 90, varios elementos genéticos conservados fueron identificados como motivos reguladores de la expresión genética en las regiones 5' no codificantes de diversas secuencias bacterianas, estando entre los más conocidos: el elemento S-box identificado por Tina Henkin, el elemento *btuB* descubierto por Robert Kadner y el elemento thi-box descrito por Juan Miranda-Ríos (Grundy & Henkin, 1998; Juan Miranda-Ríos et al., 1997; Nou & Kadner, 2000). Todos ellos presentaban la característica de llevar a cabo el proceso de regulación mediante un mecanismo hasta ese momento desconocido. Adicionalmente, el incremento en el número de genomas bacterianos analizados computacionalmente, ha contribuido al descubrimiento de muchos otros, como por ejemplo el elemento riboflavina (RFN) en la región 5' no codificante del gen involucrado en el proceso de

biosíntesis de riboflavina por Mikhail Gelfand (Vitreschak et al., 2002). A partir de esto, se propuso un método de estudio mediante el cual pudiera llevarse a cabo la evaluación de aquellos elementos reguladores cuyo mecanismo no estaba relacionado con la acción de factores proteicos accesorios. De tal manera, se comenzó con el análisis de aptámeros de RNA, mismos que fueron descubiertos gracias al uso de métodos de selección *in vitro*. Estas moléculas de RNA tienen la capacidad de interactuar directamente con moléculas pequeñas. El grupo de Ronald Breaker comprobó esta hipótesis en el año 2002 (Batey, 2015; Breaker, 2011; Vitreschak et al., 2003; W. Winkler et al., 2002).

Anteriormente, se consideraba que la regulación de los genes involucrados en la biosíntesis de tiamina estaba controlada por una proteína reguladora dependiente de tiamina pirofosfato (TPP). Tiempo después, se demostró que la TPP regulaba directamente la traducción del operón *thiCOGE* en *Rhizobium etli* a través de una secuencia conservada, denominada thi-box y una estructura de tallo-asa de mRNA sobrelapante al sitio de unión a ribosoma (RBS, por sus siglas en inglés) (Figura 1). Posteriormente, mediante análisis comparativo se predijo la existencia de una estructura más extensa que involucrara la formación de horquillas alternativas (THI-element), mismo que se comprobó y se demostró experimentalmente que su estructura regula la traducción del gen *thiM* en *E. coli*, así como también la traducción y transcripción de *thiC* en *E. coli* y la transcripción del operón de tiamina en *B. subtilis* (J. Miranda-Ríos et al., 2001; W. Winkler et al., 2002).

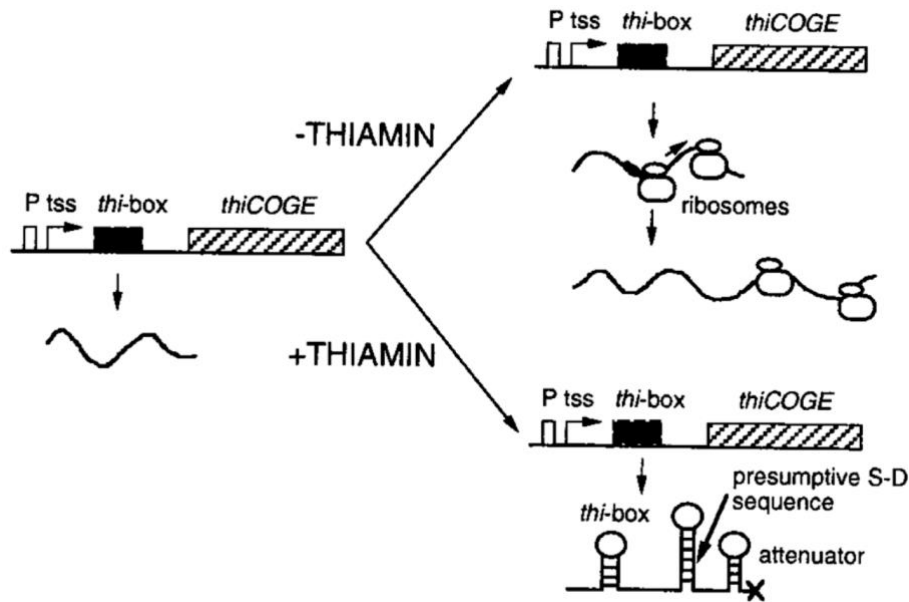


Figura 1. Mecanismo de regulación mediado por *thi*-box propuesto para el operón *thi*COGE (Tomado de J. Miranda-Ríos et al., 2001).

Riboswitches

El término riboswitch fue acuñado por Ronald Breaker en el año 2002 y los primeros estudios para su validación experimental fueron reportados en el mismo año (Nahvi et al., 2002; W. Winkler et al., 2002; W. C. Winkler et al., 2002).

Los riboswitches son elementos de RNA reguladores no codificantes que se localizan en la región intergénica 5' de ciertos genes y su principal función es monitorear el ambiente celular y controlar la expresión de su gen blanco (Garst et al., 2011; Kavita & Breaker, 2023).

La regulación mediada por riboswitches no requiere de factores macromoleculares accesorios, además de la maquinaria de expresión propia de cada organismo, por lo que su función ocurre de manera directa (Figura 2). Los

riboswitches se encuentran regulando genes involucrados en múltiples rutas metabólicas fundamentales, principalmente en procesos de biosíntesis, catabolismo, señalización o transporte de su metabolito de unión con el fin de mantener los niveles adecuados de dicha molécula (Barrick & Breaker, 2007; Sherwood & Henkin, 2016).

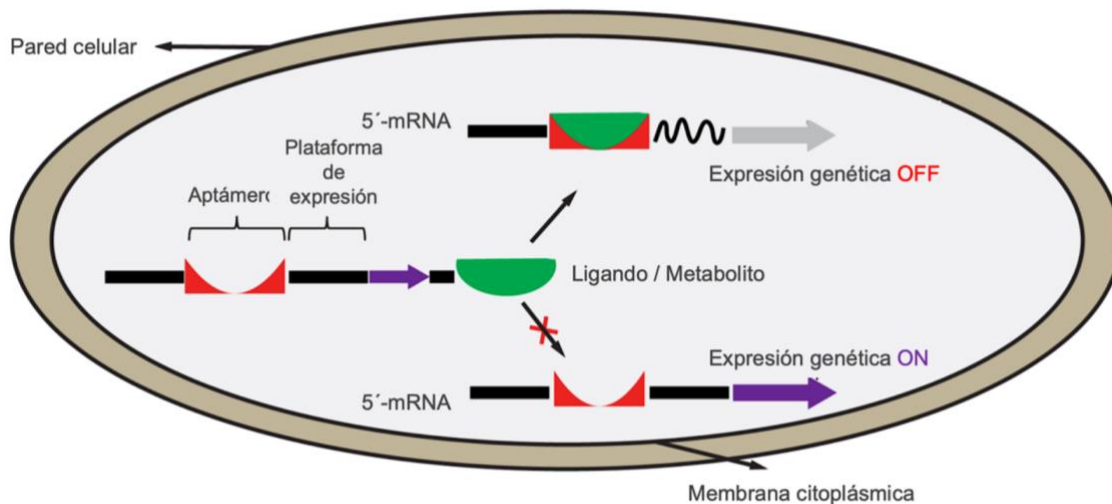


Figura 2. Mecanismo de regulación basado en riboswitches (Modificado de Sinumvayo et al., 2018).

Los riboswitches están presentes en una gran variedad de especies que pertenecen a los tres dominios de la vida, estando mayormente presentes en bacterias y arqueas. Muchas bacterias usan riboswitches para monitorear la concentración de iones y metabolitos fundamentales para, posteriormente, controlar la expresión de genes cuyos productos proteicos están directamente relacionados con la conservación de la homeostasis de dichos metabolitos (Bervoets & Charlier, 2019; Pavlova et al., 2019).

Como ya se mencionó anteriormente, los riboswitches están típicamente localizados en la región 5' no codificante de los mRNA de los genes cuya expresión controlan (Figura 3). Se considera que la razón de tal disposición podría ser que, de esta manera es posible que el riboswitch sea sintetizado primero y, posteriormente, exista la oportunidad de que este elemento responda a la unión del metabolito antes de que el gen blanco sea procesado. Resulta probable que las bacterias hayan favorecido tal disposición, debido a que esta aporta un efecto beneficioso al generar un ahorro de tiempo y evita el gasto de energía en la síntesis de un transcrito innecesario en ese preciso momento (Breaker, 2011).

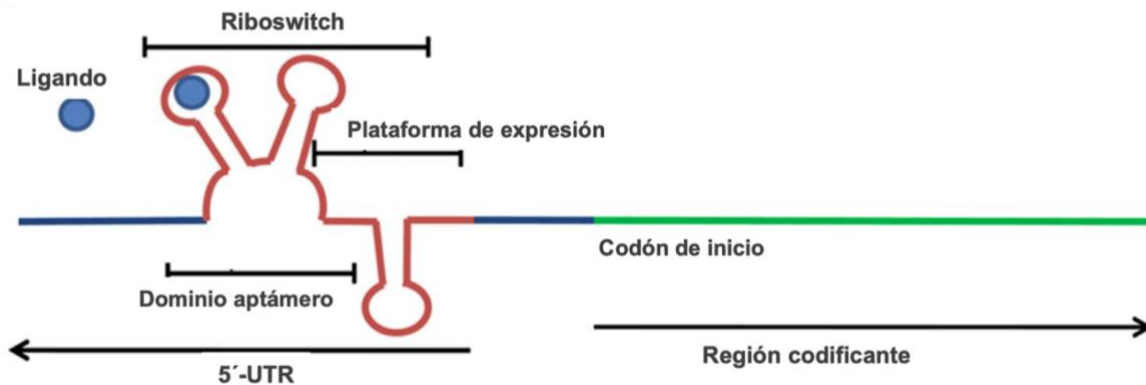


Figura 3. Localización típica del riboswitch en el genoma (Modificado de Mehdizadeh Aghdam et al., 2016).

Estructura de los riboswitches

Estructuralmente, los riboswitches están compuestos por 2 dominios: el aptámero y la plataforma de expresión (Figura 4).

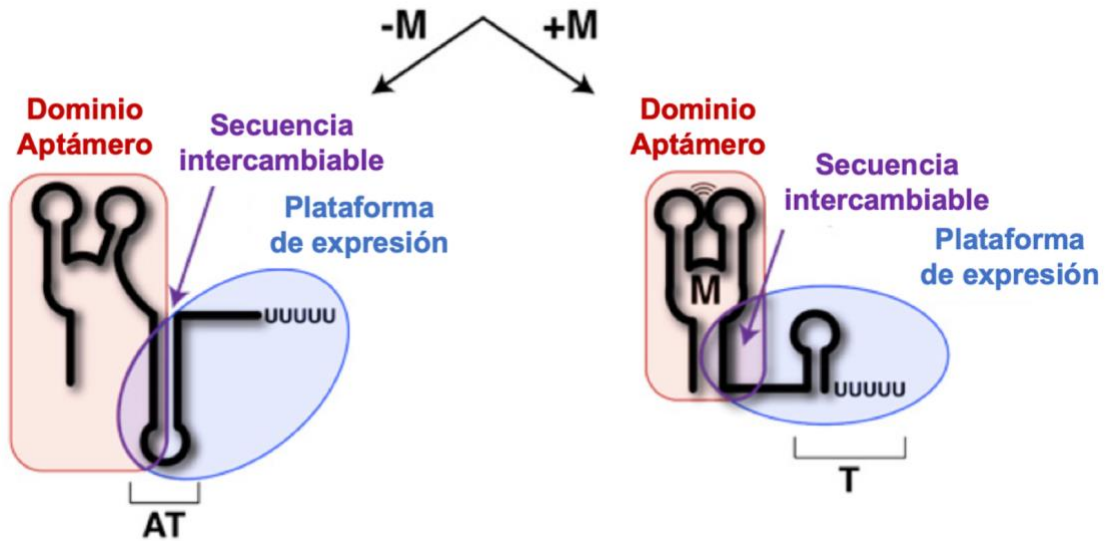


Figura 4. Estructura y dominios que componen el riboswitch (Modificado de Edwards & Batey, 2010).

Dominio aptámero

El dominio aptámero o plataforma de reconocimiento es una secuencia altamente conservada, y es la responsable del monitoreo del estado metabólico celular mediante su interacción directa con un metabolito determinado, para lo cual es necesaria la formación de una estructura tridimensional compleja (Figura 5). Típicamente, su extensión oscila entre 70 y 170 nucleótidos (Garst et al., 2011; Pavlova et al., 2019; Sherwood & Henkin, 2016).

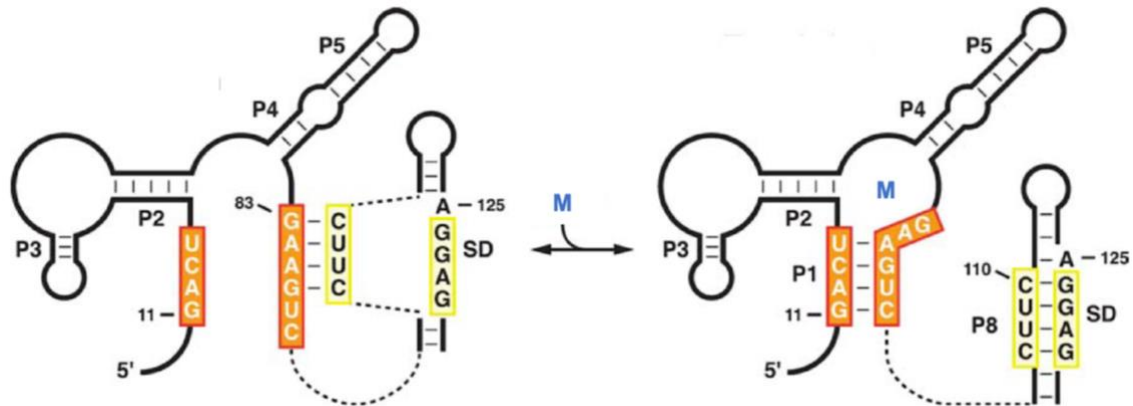


Figura 5. Rearreglo alostérico del dominio aptámero en respuesta a la unión de su ligando (Modificado de W. C. Winkler & Breaker, 2005).

Este dominio es uno de los elementos celulares más altamente conservados que se han estudiado, lo cual es debido a su necesidad de formar compartimientos de unión altamente selectivos para un metabolito blanco usando únicamente los 4 nucleótidos comunes (adenina, citosina, guanina y timina). Tal es la selectividad de estos elementos, que poseen la habilidad de distinguir entre análogos cercanamente relacionados y reconocer a sus efectores con alta especificidad, por lo que son usados para denominar las diferentes clases de riboswitches (Breaker, 2011; Garst et al., 2011; Mccown et al., 2017).

Los riboswitches poseen la capacidad de unirse a una enorme variedad de moléculas, tales como:

- Aminoácidos : Lisina, glutamina y glicina.
- Coenzimas : Cobalamina, S-adenosilmetionina (SAM), Tiamin Pirofosfato (TPP), Tetrahidrofolato (THF), etc.

- Segundos mensajeros: di-AMP cíclico (c-di-AMP), di-GMP cíclico (c-di-GMP), ZMP/ZTP, etc.
- Otros metabolitos: Manganeseo (Mn^{2+}), Aminoglicósidos (AAC-AAD), Glucosamina 6-P (GlcN6P), etc.

Considerando la molécula reconocida por la plataforma de reconocimiento o dominio aptámero del riboswitch, este toma su nombre o bien la clase a la cual pertenece. Además, también es posible organizar las clases de riboswitches en familias considerando la similitud de las moléculas que son reconocidas por el aptámero, aunque esto no tenga efecto sobre la especificidad tridimensional del reconocimiento de las mismas (Tabla 1) (Antunes et al., 2018; Breaker, 2011; Serrano-Gutiérrez & Merino, 2023).

Tabla 1. Organización de las clases de riboswitches en familias (Modificada de (Serrano-Gutiérrez & Merino, 2023)).

Familia	Clase	Metabolito o condición sensada
T-box	T-box	tRNA no aminoacilado
Cobalamin	Cobalamin	Cobalamina
	AdoCbl_riboswitch	Adenosil cobalamina
	AdoCbl_variant	Variante de Adenosil cobalamin
TPP	TPP	Tiamin Pirofosfato
SAM	SAM	S-Adenosil Metionina
	SAM-IV	S-Adenosil Metionina-IV
	SAH_riboswitch	S-Adenosil Homocisteína
	SAM_alpha	S-Adenosil Metionina (Alfaproteobacteria)
	SAM-I-IV-variant	Variante de S-Adenosil Metionina-I/IV
	SMK_box_riboswitch	S-Adenosil Metionina-III
	SAM-SAH	S-Adenosil Metionina - S-Adenosil Homocisteína
	SAM_V	S-Adenosil Metionina-V
SAM_VI	S-Adenosil Metionina-VI	
FMN	FMN	Flavin Mononucleótido
Glycine	Glycine	Glicina
	glycine-GGAnGA	Glicina-GGAnGA
ydaO-yuaA	ydaO-yuaA	di-AMP cíclico
c-di-GMP	c-di-GMP-I	di-GMP-I cíclico

	c-di-GMP-II	di-GMP-II cíclico
	c-di-GMP-I-GGC	di-GMP-I-GGC cíclico
	c-di-GMP-II-GAG	di-GMP-II-GAG cíclico
	c-di-GMP-I-UAU	di-GMP-I-UAU cíclico
	c-di-GMP-II-GCG	di-GMP-II-GCG cíclico
Purine	Purine	Purina
	2dG-II	Deoxiguanosina
	MFR	Purina (<i>Mesoplasma florum</i>)
Lysine	Lysine	Lisina
yybP-ykoY	yybP-ykoY	Manganeso
crcB / Fluoride	crcB	Fluoruro
ykkC / Guanidine	ykkC-yxkD	Guanidina-I
	mini-ykkC	Guanidina-II
	ykkC-III	Guanidina-III
MOCO_RNA_motif	MOCO_RNA_motif	Cofactor Molibdeno / Tungsteno
ykoK / M-box	ykoK	Ion Magnesio
pfl / ZMP-ZTP	pfl	ZMP/ZTP
glmS	glmS	Glucosamina-6-fosfato
PreQ1	PreQ1	Pre-queuosina1
	preQ1-II	Pre-queuosina1-II
	preQ1-III	Pre-queuosina1-III
THF	THF	Tetrahidrofolato
nhaA-I	nhaA-I	Ions / Protones Sodio
sul1	sul1	Sulfato
DUF1646	DUF1646	Sodio
raiA	raiA	Respuesta estrés
glnA	glnA	Glutamina
AAC_AAD_leader	AAC_AAD_leader	Aminoglicósidos
Mg_sensor	Mg_sensor	Magnesio
NiCo	NiCo	Ions Niquel / Cobalto
	NiCo-AnGGAG	Ions Niquel / Cobalto - AnGGAG

Dominio de expresión

Por su parte, la plataforma de expresión está encargada de controlar la expresión genética mediante un mecanismo determinado, como resultado de una modificación estructural o rearrreglo alostérico, lo que genera una transición entre dos estructuras mutuamente excluyentes (Pavlova et al., 2019; W. C. Winkler & Breaker, 2005).

La transcripción a partir del riboswitch implica un plegamiento secuencial, para lo cual ocurre una compleja interacción de diversos factores, tales como la velocidad de transcripción, la cinética y la termodinámica de la formación de la estructura de RNA; además de la unión a ligando. El mecanismo de respuesta del riboswitch puede ser descrito como un proceso donde el reconocimiento del ligando define el plegamiento del RNA naciente durante su expresión y la estructura resultante determina la interacción con el aparato transcripcional o traduccional celular (Figura 6) (Breaker, 2012; Pavlova et al., 2019).

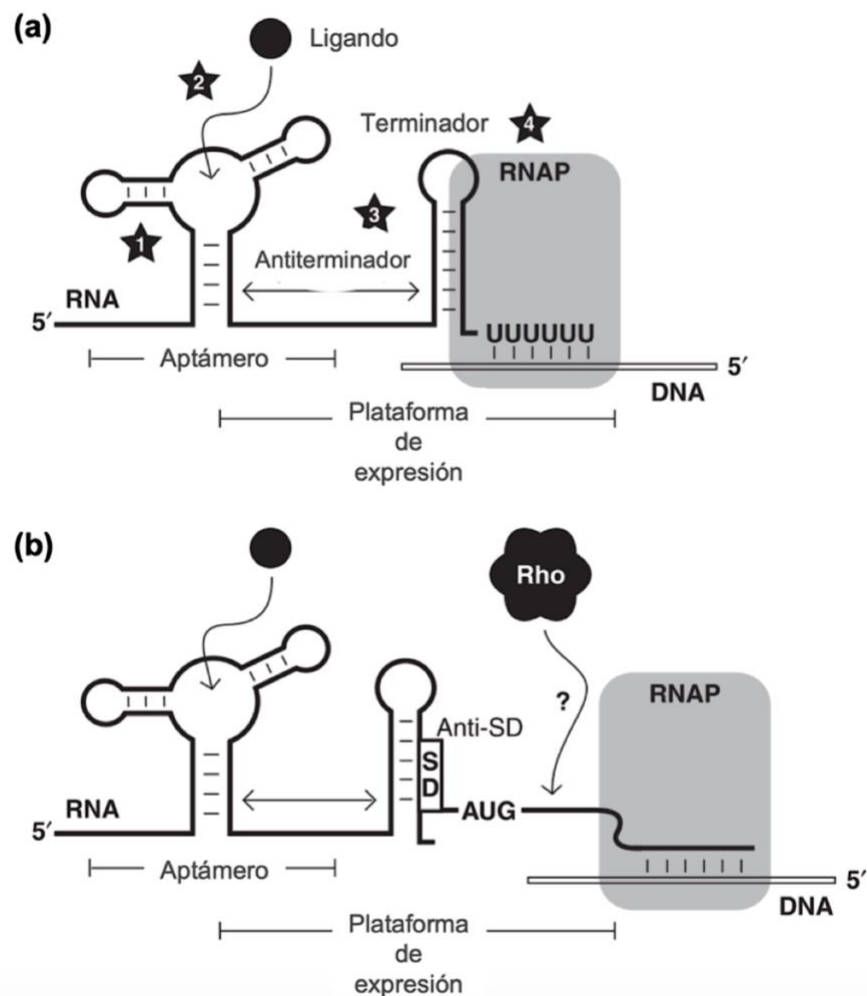


Figura 6. Principales mecanismos de regulación mediados por riboswitch: (a) Terminación de la transcripción, (b) Inicio de la traducción (Modificado de Breaker, 2012).

Para entender el mecanismo de acción del riboswitch, resulta importante hacer notar que los estados de ligando-libre y ligando-unido no se encuentran en un equilibrio dinámico. Para ello, la célula debe haber evolucionado un mecanismo mediante el cual se permita suficiente flexibilidad en la cadena de RNA naciente con el fin de mantener la competencia para el reconocimiento del ligando y garantizar el inicio del plegamiento en ausencia del mismo, generando una estructura secundaria diferente. Por lo tanto, la estructura de reconocimiento, aunque sea similar a la conformación de unión, no permite el cambio en un estado termodinámicamente estable, ya que bloquea la vía de plegamiento alternativo y el mecanismo de regulación resultante (Haller et al., 2011; Sherwood & Henkin, 2016).

Los riboswitches son entidades cinéticas, esto significa que su respuesta regulatoria está dada por las tasas de asociación y disociación de unión a ligando y no por el equilibrio de la afinidad de su unión. La conformación de una de estas estructuras dependerá de la concentración de su metabolito blanco (Batey, 2015).

Principalmente, son dos los mecanismos mediante los que los riboswitches llevan a cabo su proceso de regulación: control del inicio traduccional y control del final transcripcional. Ambos procesos utilizan cambios en el plegamiento del RNA que involucran apareado de bases de tipo Watson-Crick (Breaker, 2012; Vitreschak et al., 2003; W. C. Winkler & Breaker, 2005).

El enorme progreso en la caracterización de los estados de unión de los riboswitches a su respectivo ligando, a nivel molecular, se refleja en un impresionante número de estructuras de complejos aptámero-ligando que han sido determinadas en alta resolución, las cuales han contribuido significativamente al

entendimiento de conceptos acerca de tal interacción (Garst et al., 2011; Haller et al., 2011).

Regulación traduccional

En el caso de la regulación a nivel traduccional, formando parte de la plataforma de expresión de algunos riboswitches, existe un elemento helicoidal de extensión e identidad variable, que se localiza cercano a la región de la secuencia Shine Dalgarno (SD). Esta secuencia se identifica como una hélice de secuestro de la secuencia SD debido a que es susceptible al apareado de bases con la misma. Alternativamente, una porción de la hélice es capaz de complementarse con una secuencia en el dominio aptámero, formando una hélice antisequestro. Esta formación estructural sugiere que estas moléculas de RNA regulan la eficiencia del inicio de la traducción mediante el bloqueo del RBS bajo determinadas condiciones (Barrick & Breaker, 2007; Pavlova et al., 2019; Sinumvayo et al., 2018).

Regulación transcripcional

A nivel transcripcional, cuando un riboswitch se encuentra en presencia de un ligando específico que se une a su dominio aptámero, usualmente se forma una estructura en la región 5' no codificante que detiene la transcripción del mRNA del gen blanco. En ausencia del ligando, una estructura de antiterminador es formada en lugar del terminador, por lo tanto, el mRNA completo es transcrito y traducido a una proteína (Breaker, 2011; Sinumvayo et al., 2018).

La terminación transcripcional es uno de los mecanismos de regulación más comúnmente usados por las bacterias. Los terminadores transcripcionales

intrínsecos (Rho-independientes) son regiones fuertes, ricas en GC, seguidas por una serie de residuos de uridina. Estos elementos estructurales desestabilizan el complejo de RNA polimerasa, deteniendo la transcripción y eventualmente, liberando el templado de DNA y el producto de RNA naciente, en un proceso denominado atenuación (Figura 7) (Garst et al., 2011; W. C. Winkler & Breaker, 2005).

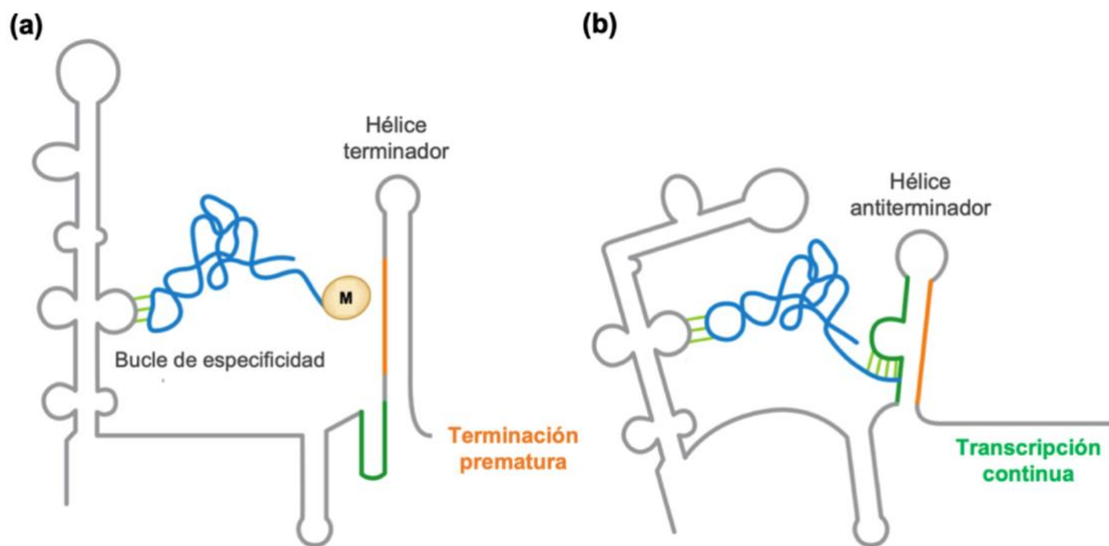


Figura 7. Mecanismo de atenuación transcripcional mediado por riboswitch: (a) Metabolito (M) unido al dominio aptámero, (b) Sin unión de metabolito (Modificado de Sherwood & Henkin, 2016).

El mecanismo típico de atenuación involucra el control de la formación de una horquilla de terminación, resultado de la unión del ligando al dominio aptámero del riboswitch, lo cual dirige la formación del tallo terminador mediante la configuración opcional de un antiterminador mutuamente excluyente (Garst et al., 2011; Pavlova et al., 2019).

ANTECEDENTES

Interferencia transcripcional

Tomando en cuenta los dos reportes mencionados anteriormente, resulta muy probable que la regulación basada en riboswitches antisentido ocurra mediante un fenómeno de interferencia transcripcional. Este mecanismo se define como la influencia supresora de un proceso transcripcional sobre otro que ocurre de manera adyacente debido al tráfico de RNA polimerasas. Durante la transcripción antisentido, la colisión entre RNA polimerasas convergentes resulta en la terminación prematura de la transcripción de uno o ambos eventos de transcripción. El choque de RNA polimerasas parece ser una interacción que ocurre a mayor distancia o bien, un efecto del superenrollamiento de DNA al frente de la maquinaria de transcripción (Figura 8) (Brantl, 2007; Georg & Hess, 2011; Sesto, 2013).

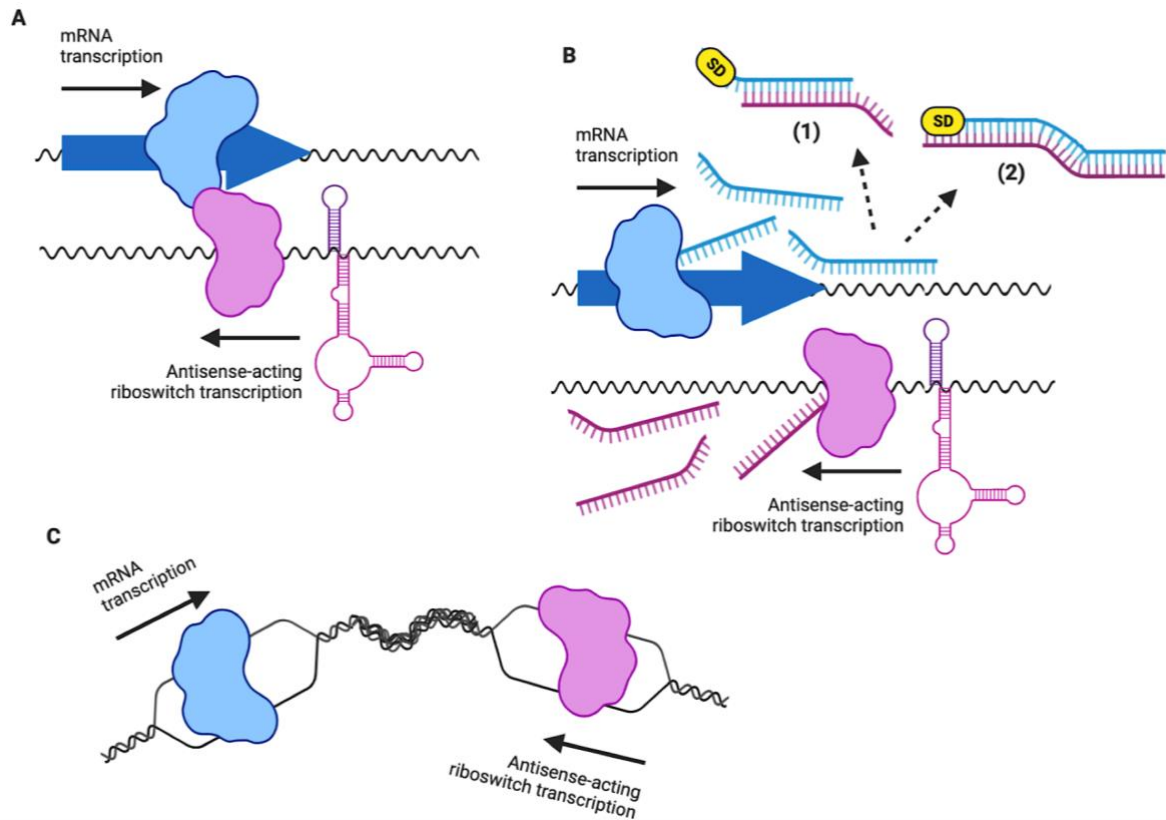


Figura 8. Modelos propuestos de los mecanismos regulatorios de las RNA polimerasas que transcriben en direcciones opuestas: A. Colisión de RNA polimerasas transcribiendo en direcciones opuestas, B. Formación de dúplex (1)mRNA-(2)asRNA, C. Superenrollamiento (Tomado de Serrano-Gutiérrez & Merino, 2023).

Riboswitches antisentido (*Antisense-acting* riboswitches)

En el año 2004, se describió por primera vez la existencia de un riboswitch colocado en el extremo 3' de su gen blanco de regulación y cuya dirección de transcripción era opuesta al mismo, en el genoma de *Clostridium acetobutylicum* (Rodionov et al., 2004). Posteriormente, en el año 2008, y gracias al estudio de la regulación del operón *ubiG/mccBA* en este organismo, se comprobó la existencia y funcionalidad de dicho elemento (Figura 9) (André et al., 2008). A este tipo de riboswitches se les denominó como riboswitches antisentido.

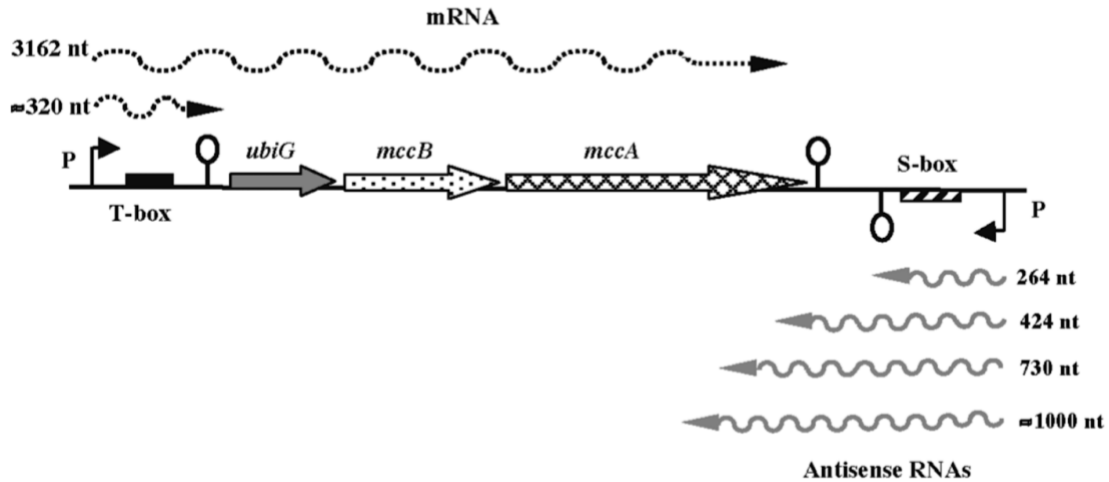


Figura 9. Mecanismo de regulación del riboswitch antisentido S-box que regula el operón *ubiG/mccAB* en *C. acetobutylicum* (Tomado de André et al., 2008).

El mecanismo de acción de los diferentes elementos regulatorios del operón *ubiG/mccAB*, el cual contiene genes requeridos para la conversión de cisteína a metionina, está basado principalmente en la interferencia transcripcional. En ausencia de metionina, la transcripción del gen blanco de regulación del riboswitch S-box (SAM) antisentido ocurre de manera regular, debido a que no existe transcripción a partir del promotor del mismo que interfiera sobre la transcripción de su gen blanco. De manera contraria, en presencia de metionina, existe transcripción a partir del riboswitch SAM antisentido al generarse un terminador transcripcional, por lo que el transcrito generado interfiere directamente sobre el proceso transcripcional de su gen blanco de regulación. El mecanismo de regulación podría deberse a la competencia entre RNA polimerasas que inician su transcripción en promotores convergentes. El análisis experimental de este riboswitch demostró la producción de RNA antisentido en respuesta a la fluctuación de los niveles de metionina (André et al., 2008; Breaker, 2012; Courtney & Chatterjee, 2014).

En el año 2013, se describió la existencia de un segundo riboswitch actuando en antisentido, mismo que es dependiente de vitamina B₁₂ (cobalamina) localizado en antisentido al gen que codifica para el regulador transcripcional *pocR* en *Listeria monocytogenes* (Figura 10). Se determinó que este riboswitch controla la transcripción de un RNA antisentido, evitando la generación de dicho transcrito, mediante la regulación positiva de la expresión de *pocR* cuando la cobalamina está disponible. En ausencia de cobalamina, el RNA antisentido es transcrito por completo e inhibe la expresión de *pocR*, interfiriendo con su proceso de transcripción. El factor transcripcional *pocR* se encarga de la regulación de los genes *pdu*, involucrados en la utilización de propanediol, haciendo uso de la Cobalamina como cofactor (Lebreton & Cossart, 2017; Mellin et al., 2013).

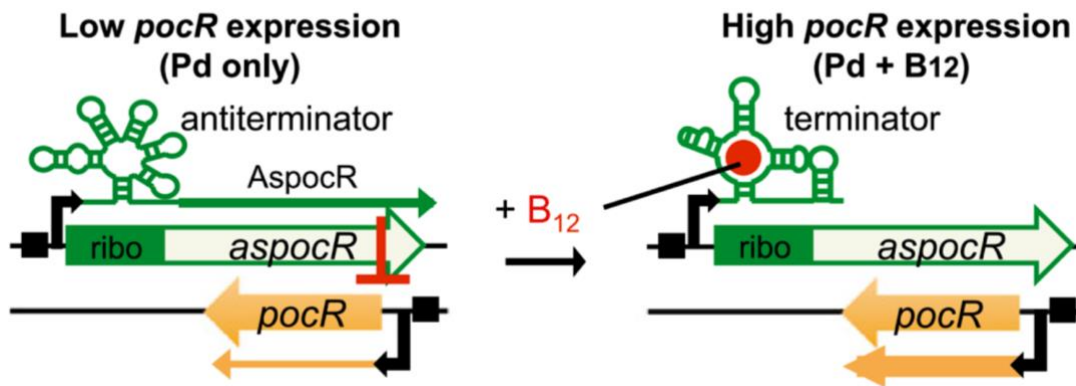


Figura 10. Mecanismo de regulación del riboswitch antisentido de Cobalamina que regula al regulador transcripcional *pocR* en *L. monocytogenes* (Tomado de Mellin et al., 2013).

HIPÓTESIS

La regulación de la expresión genética mediada por riboswitches antisentido no se limita a los dos casos descritos hasta el momento en la literatura, sino que se trata de una ventaja adaptativa en el proceso de regulación por metabolitos, por lo que podrían existir más ejemplos de este tipo de riboswitches en los genomas de bacterias y arqueas.

OBJETIVO

Análisis de los riboswitches antisentido existentes en los genomas de bacterias y arqueas disponibles hasta el momento.

Objetivos particulares

- Revisión actualizada de los riboswitches canónicos y los procesos metabólicos que estos regulan.
- Clasificación de los riboswitches antisentido encontrados por su clase, filogenia y función mediante el uso de una herramienta de procesamiento de datos.
- Contextualización los genes regulados por riboswitches antisentido con base en su filogenia y función dentro de un contexto bioquímico.
- Establecimiento de una estrategia metodológica para la identificación de errores en la anotación de secuencias genómicas basado en la identificación de riboswitches.

METODOLOGÍA

Identificación *in silico* de riboswitches en los genomas de bacterias y arqueas disponibles

La búsqueda de riboswitches se llevó a cabo en los 7,272 genomas de bacterias y arqueas a nivel de especie, disponibles en la base de datos KEGG del año 2022 (Kanehisa et al., 2023). Para ello, se generó un programa en lenguaje Perl que permitiera hacer la búsqueda mediante el uso del software CmSearch perteneciente a la paquetería Infernal (versión 1.1.3) (Chang et al., 2009; Cui et al., 2016). Previamente, se llevó a cabo la creación de una base de datos con los modelos de covarianza de las 50 clases de riboswitches descritos, hasta el momento del análisis, en la base de datos Rfam (<https://rfam.org/>) (Kalvari et al., 2021). Debido a que existen diferentes modelos de covarianza relacionados entre sí, con la capacidad de reconocer secuencias de RNA similares, se creó un programa en lenguaje Perl que permitiera seleccionar el modelo más adecuado para cada riboswitch identificado, con base en su correspondiente bitscore, obtenido mediante el análisis previo que se hizo con el programa CMsearch. Además, tomando en cuenta las similitudes entre las estructuras de los metabolitos sensados por los riboswitches, se clasificaron las 50 diferentes clases en 27 familias.

La búsqueda de estos motivos conservados se llevó a cabo en las regiones 5' y 3' intergénicas de los genomas de las bacterias y arqueas. Posteriormente, se procedió a distinguir entre los resultados si se trataba de riboswitches localizados de manera canónica, que para efectos de este proyecto denominamos “en sentido”, o bien de riboswitches localizados en antisentido, considerando para ello las

anotaciones genómicas disponibles en la base de datos KEGG, con respecto a la orientación de cada uno de los genes identificados como blanco de riboswitches. Además, se consideró que estos elementos regulatorios se localizaran a no más de 500 nucleótidos del mismo.

Un segundo análisis se llevó a cabo, considerando los mismos parámetros que ya se mencionaron previamente, pero en este caso la búsqueda de los motivos se hizo en todo el genoma, es decir que se consideraron tanto regiones anotadas como codificantes, como regiones intergénicas.

Selección de organismos representativos

Del total de las secuencias genómicas analizadas (7,272), 6,892 corresponden a organismos bacterianos y 380 a arqueas. Una vez realizado el primer análisis, se decidió seleccionar únicamente una secuencia genómica por especie analizada; esto con el propósito de evitar la sobrerrepresentación de secuencias genómicas redundantes. Para la selección del genoma representativo dentro del conjunto de genomas de cepas de la misma especie, se consideró aquel con el mayor número de marcos abiertos de lectura. Luego de esta selección, el número de secuencias genómicas de nuestro estudio se redujo a un total de 5,418, de las cuales 5,086 corresponden a bacterias y 332 a arqueas.

Análisis estadístico de los resultados

El análisis estadístico de los resultados del proyecto se llevó a cabo haciendo uso del software RStudio (Version 1.2.5033).

Además, como parte del análisis, se evaluaron las frecuencias absolutas para las familias de riboswitches en sentido y antisentido, así como la probabilidad hipergeométrica para cada familia de riboswitches antisentido. Finalmente, se analizó la distribución por phylum de bacteria o arquea para cada familia de riboswitches en sentido y antisentido.

Análisis del contexto genómico

El sentido biológico de los riboswitches antisentido encontrados, se analizó mediante una revisión inicial del contexto genómico de los genes regulados por su correspondiente riboswitch, haciendo uso del servidor GeConT (<http://biocomputo.ibt.unam.mx:8080/GeConT/>) (Martinez-Guerrero et al., 2008).

Análisis de las rutas metabólicas

Los resultados obtenidos se analizaron también desde un punto de vista bioquímico, para lo cual se utilizó la información de las rutas metabólicas en las que se encontraba implicado el gen blanco de regulación de cada uno de los riboswitches antisentido identificados disponible en la base de datos de rutas metabólicas de KEGG (<https://www.genome.jp/kegg/pathway.html>) (Kanehisa et al., 2023). Además, también se realizó su búsqueda en el servidor STRING (<https://string-db.org>) (Szklarczyk et al., 2019), con el fin de obtener mayor información acerca de las redes de interacción de la proteína codificada por cada gen.

Análisis de la función de los genes

La función de los genes regulados por riboswitches antisentido se analizó considerando tanto la descripción existente para el gen en la base de datos KEGG, como la descripción de su Cluster of Orthologous Genes (COG) correspondiente. Esta asignación fue determinada con base en una búsqueda de Hidden Markov Models (HMMs) usando el programa hmmsearch (Eddy, 2011) y un set de modelos HMM previamente construidos, mismos que representan cada uno de los 4,873 COGs existentes (Taboada et al., 2010; Tatusov et al., 2001).

Análisis del contexto sinténico de los genes relacionados con riboswitches

La comparación de las secuencias de los genes adyacentes a riboswitches, con respecto a sus homólogos en los organismos filogenéticamente más cercanos, se llevó a cabo mediante la realización de alineamientos para lo cual se utilizó el servidor MUSCLE (<https://www.ebi.ac.uk/Tools/msa/muscle/>). Con los resultados obtenidos se crearon esquemas que mostrarán los análisis pertinentes.

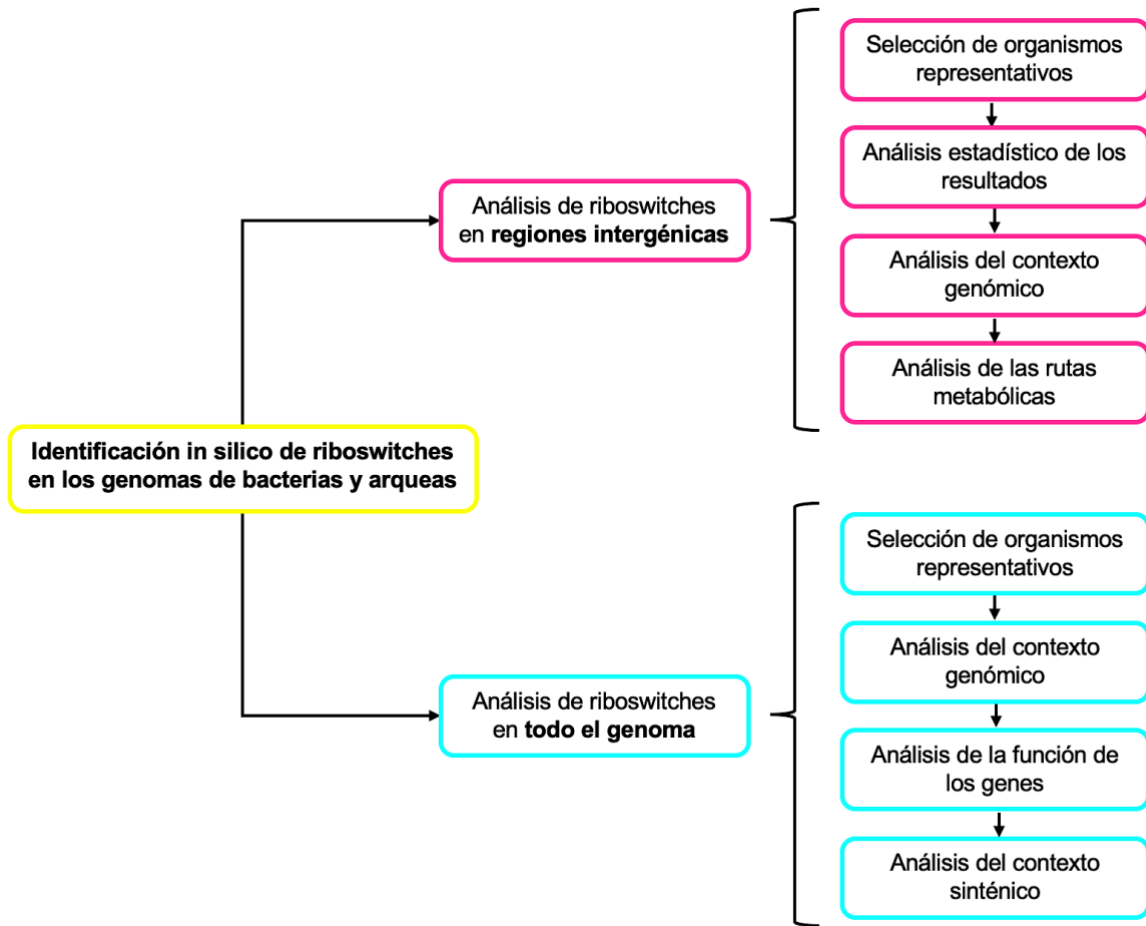


Figura 11. Diagrama general de trabajo.

RESULTADOS

Identificación *in silico* de riboswitches en el genoma de bacterias y arqueas

Se identificaron un total de 99,267 riboswitches en las regiones intergénicas del genoma de las bacterias y arqueas analizadas, mismas que fueron clasificadas como se estableció en la sección de materiales y métodos de acuerdo con su orientación respecto a la de su gen blanco de regulación, de los cuales el 98.8% (98,082) corresponde a riboswitches en sentido (Figura 12) y el 1.2% (1,185) a riboswitches antisentido (Tabla 2).

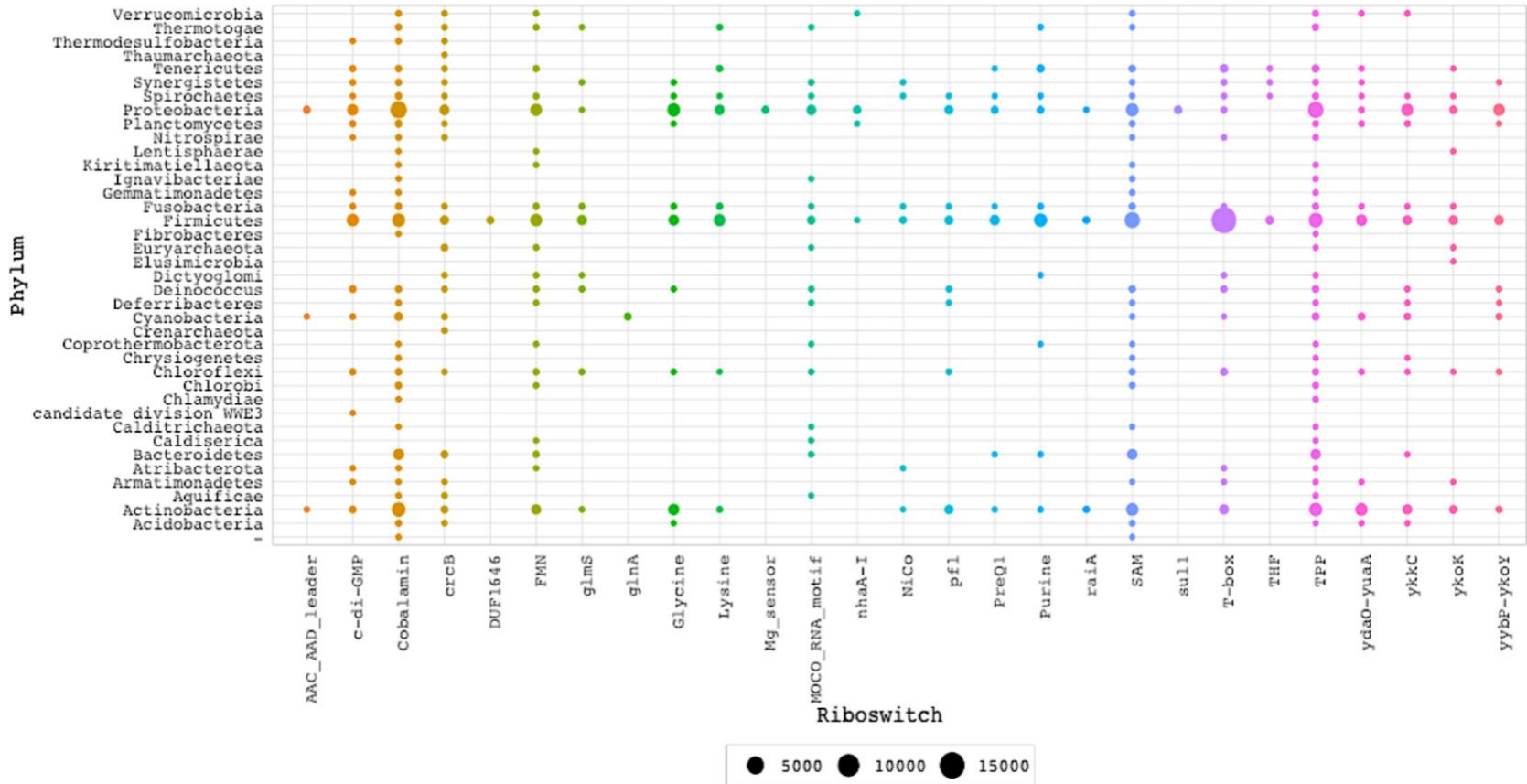


Figura 12. Distribución por phylum y familia de los riboswitches en sentido identificados en el genoma de bacterias y arqueas

Tabla 2. Riboswitches antisentido identificados en el genoma de bacterias y arqueas (Extracto de Serrano-Gutiérrez & Merino, 2023) <https://doi.org/10.1371/journal.pone.0281744.s002>.

Organismo				Gen				Riboswitch				
Representativo	Nombre	Phylum	Clave	ID	Inicio	Fin	Cadena	Familia	E-value	Inicio	Fin	Cadena
*	<i>Amycolatopsis albispota</i> WP1	Actinobacteria	aab	aab-A4R43_11230	2430457	2430813	+	Glycine	1.00E-12	2430955	2431056	-
*	<i>Amycolatopsis acidiphila</i> KCTC 39523	Actinobacteria	aacd	aacd-LWP59_14940	3073830	3074192	-	Glycine	2.10E-11	3073576	3073673	+
			aacd	aacd-LWP59_14940	3073830	3074192	-	Glycine	1.40E-05	3073674	3073791	+
	<i>Alicyclobacillus acidocaldarius</i> subsp. <i>acidocaldarius</i> Tc-4-1	Firmicutes	aad	aad-TC41_2316	2212150	2212347	+	ykoK / M-box	7.50E-36	2212400	2212563	-
			aad	aad-TC41_0140	141606	141806	-	c-di-GMP-I	8.10E-16	141354	141440	+
			aad	aad-TC41_0445	449833	450018	+	c-di-GMP-I	1.40E-15	450166	450248	-
*	<i>Acetohalobium arabaticum</i> DSM 5501	Firmicutes	aar	aar-Acear_1243	1323739	1324179	+	FMN	2.80E-18	1324244	1324394	-
*	<i>Agromyces archimandritae</i> G127AT	Actinobacteria	aarc	aarc-G127AT_15645	3387717	3388505	+	ykkC / Guanidine	3.50E-05	3388506	3388555	-
*	<i>Acetobacter ascendens</i> LMG 1590	Proteobacteria	aasc	aasc-A4S02_08380	1746203	1746397	-	Cobalamin	2.50E-28	1745586	1745819	+
*	<i>Acidisarcina polymorpha</i> SBC82	Acidobacteria	abas	abas-ACPOL_5155	5496575	5496697	+	Cobalamin	1.00E-08	5496697	5496788	-
	<i>Azospirillum brasilense</i> Sp7	Proteobacteria	abf	abf-AMK58_27890	16977	17459	-	Cobalamin	2.50E-32	16452	16670	+
*	<i>Aromatoleum bremense</i> Pbn1	Proteobacteria	abre	abre-pbn1_04840	526502	526852	-	Cobalamin	3.80E-23	526269	526409	+
*	<i>Anaerocolumna cellulolytica</i> SN021	Firmicutes	acel	acel-acsn021_04420	488476	489207	-	T-box	4.00E-05	488253	488376	+
*	<i>Aerococcus christensenii</i> CCUG28831	Firmicutes	acg	acg-AWM71_03460	734568	735755	-	TPP	2.90E-14	734234	734330	+
*	<i>Acidianus</i> sp. HS-5	Crenarchaeota	acih	acih-HS5_04810	465760	466200	+	ykoK / M-box	1.40E-05	466289	466433	-
*	<i>Alistipes communis</i> 5CBH24	Bacteroidetes	acou	acou-A5CBH24_08850	1121700	1121933	+	Cobalamin	8.60E-24	1121933	1122137	-
*	<i>Acidithiobacillus caldus</i> SM-1	Proteobacteria	acu	acu-Atc_2286	2349145	2349219	+	TPP	4.00E-06	2349219	2349276	-
*	<i>Aequorivita</i> sp. H23M31	Bacteroidetes	aev	aev-EI546_04345	963027	964134	+	TPP	4.00E-07	964191	964289	-
*	<i>Anoxybacter fermentans</i> DY22613	Firmicutes	aft	aft-BBF96_13995	3115267	3116469	+	c-di-GMP-I	1.50E-12	3116661	3116761	-
*	<i>Alicyclobacillus ferroxydans</i> S30H14	Firmicutes	afx	afx-JZ786_18910	3988711	3988848	-	c-di-GMP-I	3.60E-14	3988563	3988645	+
			afx	afx-JZ786_05900	1264717	1265508	-	c-di-GMP-I	7.70E-07	1264601	1264689	+
*	<i>Anseongella ginsenosidimitans</i> Gsoil 524	Bacteroidetes	agd	agd-FRZ59_08895	2138282	2138848	-	TPP	2.30E-12	2138124	2138219	+
	<i>Agrobacterium fabacearum</i> H13-3	Proteobacteria	agr	agr-AGROH133_06158	1440191	1441231	+	SAM	1.00E-08	1441271	1441349	-
*	<i>Actinomadura graeca</i> 32-07	Actinobacteria	agra	agra-AGRA3207_005183	5880813	5882000	+	ykkC / Guanidine	1.80E-06	5882024	5882086	-
			agra	agra-AGRA3207_000905	934260	935630	-	ykkC / Guanidine	5.00E-06	934104	934165	+
*	<i>Agromyces aureus</i> AR33	Actinobacteria	agy	agy-ATC03_16660	3726902	3732256	+	SAM	2.90E-12	3732592	3732708	-
*	<i>Adlercreutzia hattorii</i> 8CFCBH1	Actinobacteria	ahat	ahat-ADCFC_16390	1952469	1952678	+	FMN	8.70E-25	1952731	1952882	-
*	<i>Anaerobacillus isosaccharinicus</i> NB2006	Firmicutes	aia	aia-AWH56_001000	203329	204828	+	T-box	4.00E-27	205079	205327	-
			aia	aia-AWH56_019820	3904933	3905109	-	Lysine	8.10E-31	3904756	3904931	+
			aia	aia-AWH56_024220	4798658	4800451	-	c-di-GMP-I	1.80E-13	4798412	4798497	+
*	<i>Aequorivita iocasae</i> KX20305	Bacteroidetes	aic	aic-JK629_13680	2968127	2969638	+	TPP	7.60E-09	2969672	2969772	-

Posteriormente, con base en la frecuencia absoluta y la distribución hipergeométrica para cada una de las familias de riboswitches, evaluamos el significado estadístico de la tendencia de los riboswitches para actuar en antisentido (Tabla 3). Nuestro estudio identificó que la familia de riboswitches que tiene una mayor tendencia para actuar en antisentido es la familia c-di-GMP, mismos que no son tan frecuentes en sentido. Esta familia de riboswitches se encuentra comúnmente involucrada en la regulación de procesos de señalización celular y definición del destino celular (motilidad, virulencia y formación de biofilm). Además, de manera contrastante a lo que ocurre para los riboswitches en sentido donde prevalecen de manera muy importante los elementos de la familia T-box, las familias de riboswitches de cobalamina, ykoK/M-box (magnesio) y MOCO (molibdeno) también presentan una alta tendencia a actuar en antisentido. Resulta interesante el hecho de que un elemento regulado por un segundo mensajero, como el c-di-GMP, tenga la mayor probabilidad de aparición en antisentido; además, es importante resaltar el hecho de que la cobalamina, el magnesio y el molibdeno son importantes cofactores ampliamente utilizados en reacciones indispensables en diferentes procesos metabólicos que requieren una estricta regulación. Por otro lado, de manera interesante, ningún riboswitch antisentido perteneciente a las familias THF (Tetrahidrofolato) y NiCo (Níquel-Cobalto) fue identificado en nuestro análisis, a pesar de no ser los menos abundantes en la frecuencia general (en sentido).

Tabla 3. Tendencia de las familias de riboswitches para actuar en antisentido (Modificado de Serrano-Gutiérrez & Merino, 2023).

Familia	Riboswitches en sentido Frecuencia absoluta	Riboswitches antisentido Frecuencia absoluta	P-value
c-di-GMP	2,554	152	4.55E-62
cobalamina	10,769	190	1.05E-10
ykoK / M-box	828	27	1.67E-06
MOCO	834	24	4.34E-05
AAC_AAD	97	3	0.10
glnA	97	2	0.30
nhaA-I	180	2	0.60
sul1	195	2	0.64
Mg_sensor	104	1	0.69
glicina	3,153	32	0.74
DUF1646	125	1	0.75
raiA	127	1	0.76
ydaO-yuaA	2,799	27	0.81
FMN	3,472	32	0.89
ykkC / guanidina	2,526	20	0.96
crcB / fluoruro	1,446	9	0.98
pfl / ZMP-ZTP	1,001	5	0.99
TPP	9,766	83	1.00
purina	2,289	13	1.00
PreQ1	771	2	1.00
glmS	635	1	1.00
lisina	1,974	8	1.00
SAM	8,039	56	1.00
yybP-ykoY	1,907	5	1.00
THF	290	0	1.00
NiCo	100	0	1.00
T-box	15,851	107	1.00

Aunado al análisis previo, se estudió la distribución de los riboswitches antisentido agrupándolos por el phylum en el que fueron identificados y considerando la familia de riboswitches a la que pertenecen. Como resultado de esto se observó que el phylum con mayor cantidad de riboswitches antisentido corresponde a los Firmicutes, mientras que los phyla con riboswitches antisentido distribuidos en una mayor variedad de familias son Proteobacteria y Firmicutes,

mismo resultado a lo que fue observado en el caso de los riboswitches en sentido (Figura 12). Además, se encontró que los riboswitches antisentido de la familia cobalamina presentan distribución en una mayor variedad de phyla (Figura 13).

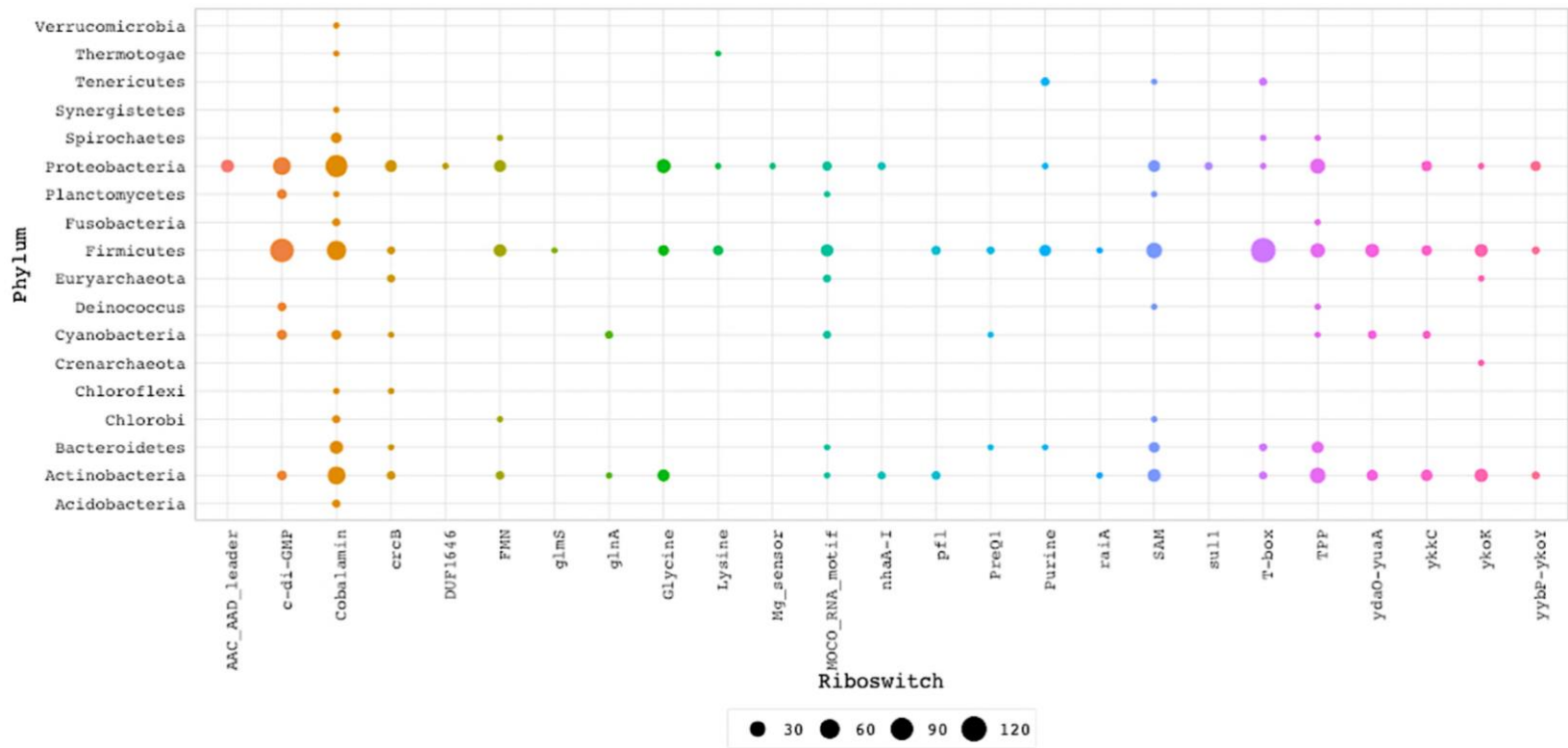


Figura 13. Distribución por phylum y familia de los riboswitches antisentido identificados en el genoma de bacterias y arqueas.

Clasificación de tipos de genes regulados por riboswitches antisentido

Posteriormente, y con el propósito de identificar posibles tendencias en las que los riboswitches antisentido son capaces de regular la expresión de genes, así como entender la naturaleza de las implicaciones bioquímicas de tal regulación, clasificamos los genes probablemente regulados por riboswitches antisentido en diez grupos diferentes (Tabla 3), haciendo para ello uso de un sistema híbrido que consideraba tanto una clasificación automatizada (mediante la identificación de términos clave en la información obtenida), como una profunda curación manual de los datos (mediante la obtención de información disponible en literatura y en bases de datos). Dado lo cual se lograron establecer las siguientes clases, de acuerdo a las tendencias observadas, y se describirá brevemente un ejemplo de un modelo regulatorio para cada una de las mismas:

Tabla 4. Clasificación de los genes regulados por riboswitches antisentido (Extracto de Serrano-Gutiérrez & Merino, 2023)
<https://doi.org/10.1371/journal.pone.0281744.s003>.

Clasificación del gen	Organismo	Phylum	ID Organismo	ID Gen	Familia del riboswitch	COG	Descripción del COG	Descripción del gen
Interconversion	<i>Clostridium acetobutylicum</i> ATCC 824	Firmicutes	cac	cac-CA_C0931	SAM	COG0031	Cysteine synthase	Cysteine synthase
Interconversion	<i>Actinomadura graeca</i> 32-07	Actinobacteria	agra	agra-AGRA3207_000905	ykkC / Guanidine	COG1486	Alpha-galactosidases/6-phospho-beta-glucosidases, family 4 o	6-phospho-beta-glucosidase
Interconversion	<i>Actinomadura graeca</i> 32-07	Actinobacteria	agra	agra-AGRA3207_005183	ykkC / Guanidine	COG1404	Subtilisin-like serine proteases	S8 family serine peptidase
Cofactor	<i>Amycolatopsis</i> sp. AA4	Actinobacteria	amyc	amyc-CU254_21355	SAM	COG0500	SAM-dependent methyltransferases	class I SAM-dependent methyltransferase
Cofactor	<i>Corynebacterium argentoratense</i> DSM 44202	Actinobacteria	caz	caz-CARG_08305	Cobalamin	COG1045	Serine acetyltransferase	hypothetical protein
Cofactor	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350	Actinobacteria	sgr	sgr-SGR_1654	Cobalamin	COG2265	SAM-dependent methyltransferases related to tRNA	putative RNA methyltransferase
Transporters	<i>Brevibacterium linens</i> BS258	Actinobacteria	bly	bly-A2T55_02635	ykkC / Guanidine	COG0747	ABC-type dipeptide transport system, periplasmic component	hypothetical protein
Transporters	<i>Acidipropionibacterium acidipropionici</i> ATCC 4875	Actinobacteria	pbo	pbo-PACID_00450	Cobalamin	COG4597	ABC-type amino acid transport system, permease component	hypothetical protein
Transporters	<i>Conexibacter woesei</i> DSM 14684	Actinobacteria	cwo	cwo-Cwoe_1787	ydaO-yuaA	COG2059	Chromate transport protein ChrA	hypothetical protein
Signaling	<i>Granulicella mallensis</i> MP5ACTX8	Actinobacteria	gma	gma-AciX8_0323	Cobalamin	COG2205	Osmosensitive K+ channel histidine kinase	integral membrane sensor signal transduction histidine kinas
Signaling	<i>Gramella forsetii</i> KT0803	Bacteroidetes	gfo	gfo-GFO_0733	TPP	COG0589	Universal stress protein UspA and related nucleotide-binding	universal stress protein family protein
Signaling	<i>Alkalibacter</i> sp. ES005	Firmicutes	alka	alka-J0B03_03500	MOCO_RNA_motif	COG0589	Universal stress protein UspA and related nucleotide-binding	universal stress protein
TFs	<i>Listeria monocytogenes</i> EGD-e	Firmicutes	lmo	lmo-lmo1150	Cobalamin	COG4936	Predicted sensor domain	transcriptional regulator Pocr
TFs	<i>Streptomyces aquilus</i> GGCR-6	Actinobacteria	saqu	saqu-EJC51_15435	TPP	COG1309	Transcriptional regulator	TetR/AcrR family transcriptional regulator
TFs	<i>Streptomyces cadmisioli</i> ZFG47	Actinobacteria	scad	scad-DN051_26805	TPP	COG1309	Transcriptional regulator	TetR/AcrR family transcriptional regulator
Transposase	<i>Nocardia seriolae</i> EM150506	Actinobacteria	nsr	nsr-NS506_02699	ykoK / M-box	COG5659	FOG: Transposase	hypothetical protein
Transposase	<i>Thermus scotoductus</i> SA-01	Deinococcus	tsc	tsc-TSC_c17270	c-di-GMP-I	COG3335	Transposase and inactivated derivatives	transposase family protein
Transposase	<i>Allobaculum mucolyticum</i> 128	Firmicutes	amuo	amuo-KWG62_01805	T-box	COG3385	FOG: Transposase and inactivated derivatives	IS4 family transposase
Pseudogene	<i>Candidatus Frankia datiscaae</i> Dg1	Actinobacteria	fsy	fsy-FsymDg_2117	Cobalamin	---	---	pseudogene
Pseudogene	<i>Micromonospora carbonacea aurantiaca</i>	Actinobacteria	mcab	mcab-HXZ27_03575	Cobalamin	---	---	pseudogene
Pseudogene	<i>Micromonospora tulbaghiae</i> CNY-010	Actinobacteria	mtua	mtua-CSH63_05160	Cobalamin	---	---	pseudogene
Uncharacterized	<i>Mycobacterium canettii</i> CIPT 140060008	Actinobacteria	mcq	mcq-BN44_40063	ykoK / M-box	---	---	Conserved protein of unknown function
Uncharacterized	<i>Serinicoccus hydrothermalis</i> JLT9	Actinobacteria	serj	serj-SGUI_0533	Cobalamin	COG3778	Uncharacterized protein conserved in bacteria	hypothetical protein
Uncharacterized	<i>Mucilaginibacter ginsenosidivorax</i> KHI28	Bacteroidetes	mgk	mgk-FSB76_21060	SAM	---	---	DUF3943 domain-containing protein
Hypothetical	<i>Acidiscarcina polymorpha</i> SBC82	Acidobacteria	abas	abas-ACPOL_5155	Cobalamin	---	---	hypothetical protein
Hypothetical	<i>Adlercreutzia hattorii</i> 8CFCBH1	Actinobacteria	ahat	ahat-ADCFC_16390	FMN	---	---	hypothetical protein
Hypothetical	<i>Acidipropionibacterium acidipropionici</i> ATCC 4875	Actinobacteria	pbo	pbo-PACID_11190	Cobalamin	---	---	hypothetical protein
Other	<i>Agromyces aureus</i> AR33	Actinobacteria	agy	agy-ATC03_16660	SAM	COG1205	Distinct helicase family with a unique C-terminal domain inc	hypothetical protein
Other	<i>Amycolatopsis acidiphila</i> KCTC 39523	Actinobacteria	aacd	aacd-LWP59_14940	Glycine	---	---	nuclear transport factor 2 family protein
Other	<i>Amycolatopsis</i> sp. AA4	Actinobacteria	amyc	amyc-CU254_20220	Glycine	---	---	nuclear transport factor 2 family protein

a) **Genes que codifican enzimas involucradas en la interconversión de compuestos relacionados con el metabolito detectado por el riboswitch.**

Se identificó un riboswitch antisentido de la clase ykkC-ykkD o Guanidino-I en el genoma de la bacteria *Paenibacillus polymyxa* SC2 regulando el gen ppm-PPSC2_12215, mismo que codifica para la proteína amidasa AmyE (Figura 14). Esta enzima, mejor conocida como 4-guanidinobutanamida amidohidrolasa, está involucrada en la ruta de metabolismo de la arginina, catalizando la reacción de conversión de 4-guanidinobutanamida a 4-guanidinobutanoato, resultado en la conversión de un grupo funcional amido en un ácido carboxílico (Nelson et al., 2017; Reiss et al., 2017).

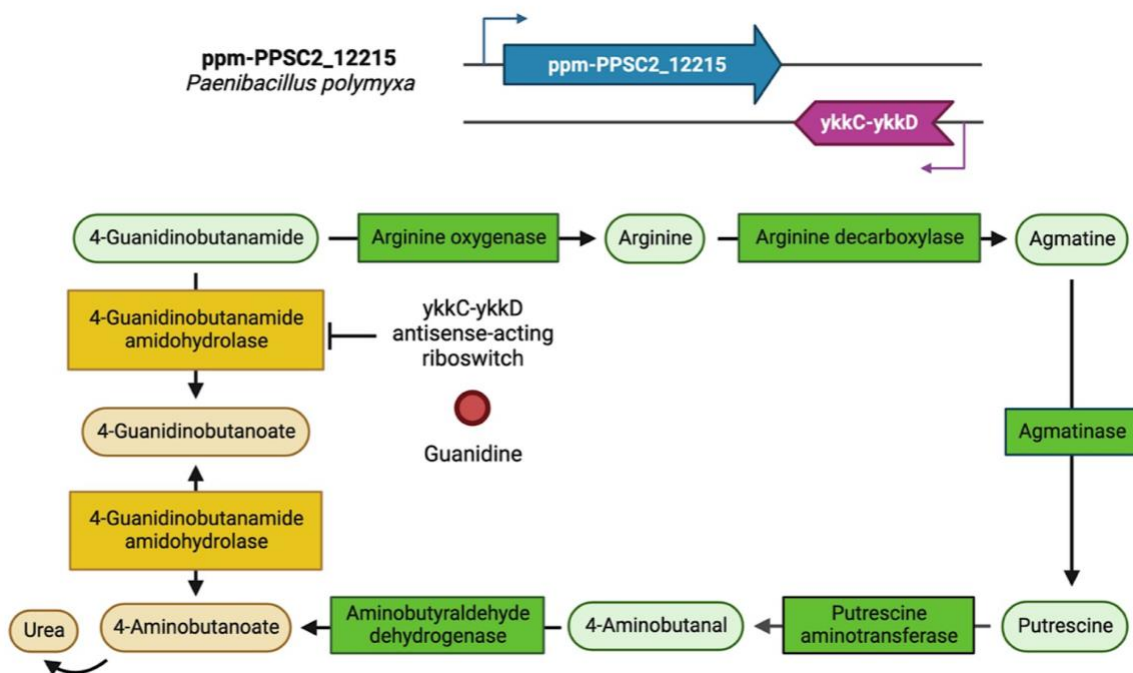


Figura 14. Gen codificante de una enzima involucrada en la interconversión de un compuesto relacionado al metabolito sensorizado por el riboswitch (Tomado de Serrano-Gutiérrez & Merino, 2023).

La arginina es el compuesto central de esta vía, y actúa como un precursor en la biosíntesis de muchas poliaminas, tales como la putrescina y la espermidina, los cuales son esenciales en diversos procesos fisiológicos. También pueden ser una pobre fuente de nitrógeno. Además, la arginina es el único aminoácido con un grupo guanidino en su composición, a pesar del hecho de que constituye compuestos que son abundantes al interior de las células bacterianas, incluyendo la creatinina y metabolitos secundarios tales como la estreptomina (Caldara et al., 2008; Sinn et al., 2021).

El riboswitch antisentido identificado en el extremo 3' del gen ppm-PPSC2_12215 podría permitir la síntesis de la enzima 4-guanidinobutanamida amidohidrolasa cuando el guanidino está presente; en este caso, su disponibilidad permite el avance a lo largo de la ruta mientras está formando parte de la 4-guanidinobutanamida, un compuesto que es catabolizado a butanoato y urea en un paso subsecuente. Por otro lado, cuando el riboswitch sensa la ausencia de guanidino, la ruta es interrumpida con el fin de mantener los niveles transitorios de guanidino mientras es parte del compuesto 4-guanidinobutanamida.

Como era esperado, los genes pertenecientes a este grupo que son regulados por riboswitches antisentido son aquellos que codifican enzimas con actividades de transferasa tales como aminotransferasas, fosfatasas, aciltransferasas, sintetasas y deshidrogenasas. La regulación de los riboswitches antisentido de este grupo inhiben la síntesis de sus correspondientes genes blanco cuando detectan bajas concentraciones del metabolito involucrado. Las familias más importantes de

riboswitches antisentido en este grupo incluyen principalmente Cobalamina, TPP, T-box, SAM, FMN y MOCO.

- b) Genes codificantes para enzimas que requieren un cofactor detectado por el riboswitch para ser activas o que participan en una ruta metabólica en la que otra enzima requiere dicho cofactor.**

En este caso, un riboswitch antisentido de cobalamina regulando el gen *clj-CLJU_c22210* que codifica para la enzima fosforibosilaminoimidazolcarboxamida formiltransferasa/IMP ciclohidrolasa bifuncional, mejor conocida como la enzima AICARFT/IMPCHasa fue identificado en el genoma de *Clostridium ljungdahlii* DSM 13528 mediante análisis bioinformático (Figura 15). *C. ljungdahlii* es una bacteria anaeróbica acetogénica que fermenta azúcar o CO₂, sintetiza gas (CO/H₂), y puede utilizar CO como sustrato. Además, este organismo posee un número significativo de proteínas que requieren de cobalamina como cofactor y también es capaz de sintetizar cobalamina mediante una ruta anaeróbica (Köpke et al., 2010; Romine et al., 2017).

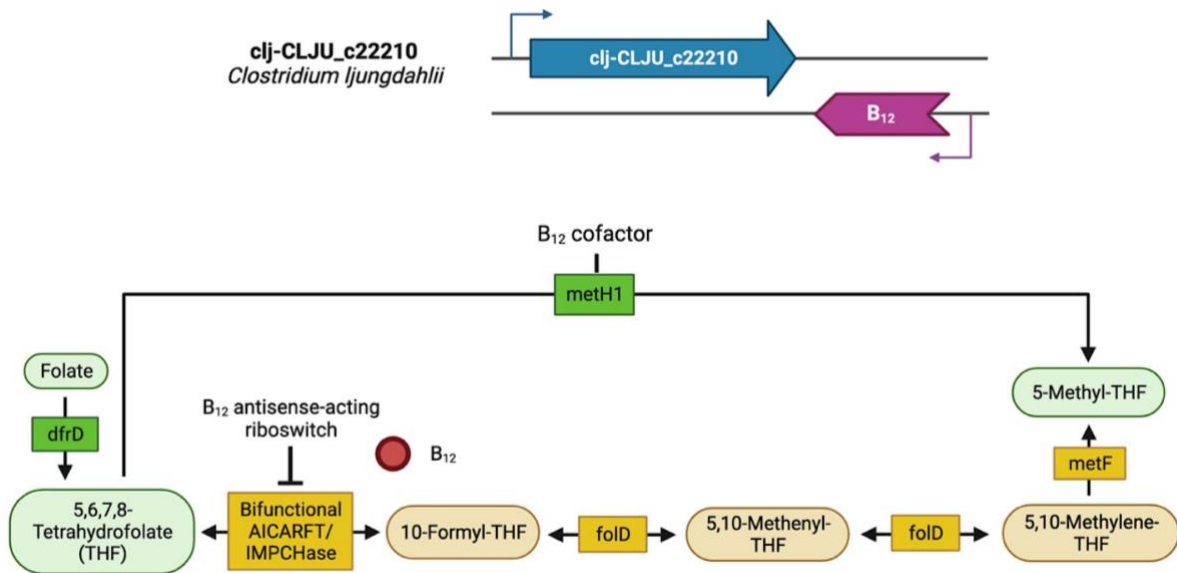


Figura 15. Gen que codifica para una enzima que requiere un cofactor sentido por el riboswitch para activarse o que participa en una ruta metabólica en la cual otra enzima requiere de dicho cofactor (Tomado de Serrano-Gutiérrez & Merino, 2023).

La enzima AICARFT/IMPCHasa está involucrada en la ruta metabólica de un carbono, la cual comprende los ciclos de folato y metionina, y permite la generación de grupos metilo que son usados en diferentes reacciones de metilación y en la biosíntesis de importantes precursores anabólicos, uno de los cuales es el 5-metil-tetrahydrofolato (5-metilTHF), cuya síntesis puede llevarse a cabo a través de dos diferentes rutas. El punto de partida de ambos es la síntesis de tetrahydrofolato (THF) a partir de folato gracias a la actividad de la enzima dihydrofolato reductasa DfrD. En una ruta, THF es convertida en 10-formilTHF por la enzima AICARFT/IMPCHasa, codificada por el gen regulado por el riboswitch antisentido de cobalamina identificado. A continuación, la 10-formilTHF sirve como el sustrato para la síntesis de 5,10-metenilTHF and 5,10-metilenTHF en dos reacciones consecutivas catalizadas por la enzima metilentetrahydrofolato deshidrogenasa

(NADP⁺)/meteniltetrahidrofolato ciclohidrolasa bifuncional, FOLD. Este último compuesto es el precursor para la síntesis de 5-metilTHF catalizado por la enzima metilTHF reductasa, MetF, con la participación de NADPH y flavoproteínas. Adicionalmente, mediante otra ruta, el 5-metilTHF es directamente sintetizado a partir de THF por la enzima metionina sintasa Meth1, cuya actividad es dependiente de cobalamina.

El riboswitch antisentido de cobalamina identificado regula el gen clj-CLJU_c22210, mismo que codifica para la AICARFT/IMPCHasa bifuncional, lo cual permite la selección de la ruta que será usada para la síntesis de 5-metilTHF, de acuerdo a la disponibilidad bacteriana de cobalamina. De tal manera, cuando la concentración intracelular de cobalamina es alta, el 5-metilTHF puede ser directamente sintetizado a partir de THF por la enzima dependiente de cobalamina Meth1, mientras que cuando existen cambios en la concentración intracelular de cobalamina, el mecanismo de regulación dependiente del riboswitch favorecerá la actividad de la enzima AICARFT/IMPCHasa bifuncional. De esta manera, la regulación del gen clj-CLJU_c22210 por el riboswitch antisentido de cobalamina garantiza la biosíntesis de 5-metilTHF y reasegura la correcta utilización de cobalamina debido a que su biosíntesis y consumo resultan metabólicamente muy costosos. Además, los intermediarios metabólicos de esta ruta tienen un papel esencial en la biosíntesis de moléculas vitales tales como purinas, DNA, CoA y serina (Romine et al., 2017; Yi et al., 2021).

Debido a la naturaleza de los genes regulados en este grupo, los cuales se caracterizan como enzimas codificadas que requieren de un cofactor para ser

activas, la familia de riboswitches antisentido cobalamina (cobalamina, AdoCbl, and AdoCbl-variant) es el miembro más importante de este grupo, debido a que representa el 38% del número total de riboswitches en esta categoría. De manera general, las familias de enzimas con actividades de isomerasa y metiltransferasa son las que más frecuentemente requieren cobalamina, o alguno de sus derivados, como cofactores. Otra familia importante de riboswitches antisentido en este grupo es la familia TPP. El cofactor TPP consiste de un anillo de pirimidina, un anillo tiazol y un grupo funcional pirofosfato, los cuales están conectados al otro, con el anillo tiazol siendo la porción comúnmente involucrada en reacciones enzimáticas, debido a que es la parte más reactiva de la molécula. También se encontraron ejemplos de riboswitches antisentido del cofactor molibdeno (MOCO) que regulan genes involucrados en el metabolismo de los cofactores molibdeno y tungsteno.

c) Genes que codifican proteínas transportadoras para compuestos relacionados al metabolito detectado por el riboswitch.

Un riboswitch antisentido de la clase yybP-ykoY, cuyo metabolito de unión es manganeso (Mn^{2+}), regula el gen cthm-CFE_2050 que codifica para una proteína transportadora de Mn^{2+} en el genoma de *Carboxydocella thermautotrophica* 019 (Figura 16) (Dambach et al., 2015). Mn^{2+} es un ión esencial en el metabolismo bacteriano que participa en múltiples procesos metabólicos, actuando principalmente como un cofactor para varias enzimas y como un agente protector contra especies reactivas de oxígeno (ROS) mediante la sustitución transitoria de moléculas de Mn^{2+} por hierro (Fe) en los sitios activos de algunas enzimas para prevenir el daño oxidativo de las proteínas (L. S. Waters, 2020). Por otra parte, el

exceso de Mn^{2+} puede resultar tóxico debido a que su acumulación en bacteria está directamente relacionada a la deficiencia de otros iones, lo cual puede entonces provocar la metalación errónea de los factores de transcripción y enzimas críticas, afectando el desarrollo celular, la virulencia y la sensibilidad a ROS (Kehres & Maguire, 2003; Richer et al., 2003; L. S. Waters et al., 2011; Zeinert et al., 2018).

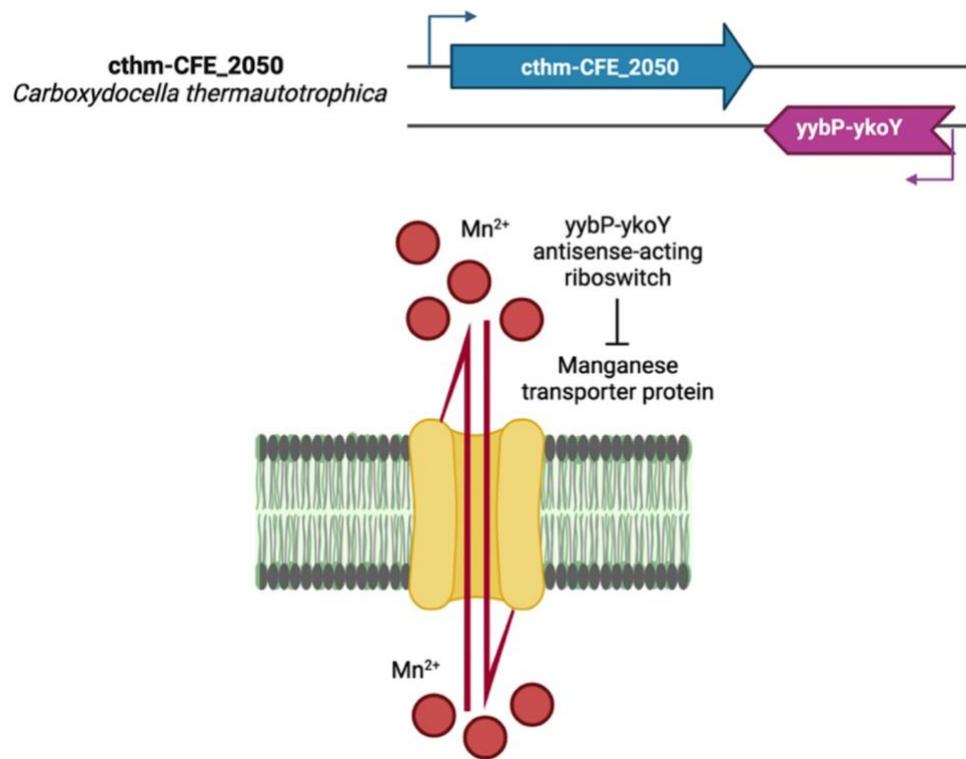


Figura 16. Gen codificante de una proteína transportadora de un compuesto relacionado al metabolito sentido por el riboswitch (Tomado de Serrano-Gutiérrez & Merino, 2023).

El gen identificado como sujeto a regulación por el riboswitch antisentido se encuentra extensamente distribuido a lo largo del phylum Firmicutes y tiene un papel esencial en el metabolismo de *C. thermautotrophica* 019. Esta bacteria ha sido descrita como carboxidotrófica, pero sus capacidades metabólicas permanecen mayormente desconocidas. Sin embargo, como la habilidad de transportar Mn^{2+} ha

sido observada en otras especies, es sabido que estas especies pueden proliferar en ambientes hostiles, incluyendo la superficie de sustratos minerales compuestos por Mn^{2+} , tales como diatomita y glauconita, en la ausencia de una fuente de carbono (Slobodkina et al., 2012; Toshchakov et al., 2018). El riboswitch antisentido identificado en el extremo 3' del gen *cthm-CFE_2050* podría prevenir la acumulación interna excesiva de Mn^{2+} que de otra manera podría tener efectos tóxicos.

Muchos genes de este tercer grupo son miembros de sistemas de transporte de tipo ABC o codifican para diferentes proteínas transportadoras, tales como los transportadores de cationes $Co/Zn/Cd$, simportadores de dicarboxilato Na^+/H^+ , o antiportadores Ca^{2+}/Na^+ . Los riboswitch antisentido más comunes de este grupo son de las familias cobalamina, c-di-GMP, T-box, SAM, *ykoK*, TPP y MOCO.

d) Genes codificantes de proteínas de señalización celular.

Un riboswitch antisentido de la clase AAC-AAD, el cual sensa y responde a concentraciones de aminoglicósidos, se identificó regulando el gen *paeu-BN889_07203* que codifica para la síntesis de la enzima hidroxiaxilglutación hidrolasa en la bacteria *Pseudomonas aeruginosa* PA38182 (Figura 17). Esta enzima, mejor conocida como glioxalasa II (GlxII), participa en la ruta de metabolismo de metilglioxal (MG) a través del sistema glioxalasa (Glx), el cual es un complejo enzimático para la detoxificación crítica. Este sistema comprende dos metaloenzimas a cargo de descomponer MG citotóxico, metabolizándolo hasta sus correspondientes α -hidroxiácidos (He et al., 2013; Sukdeo & Honek, 2007).

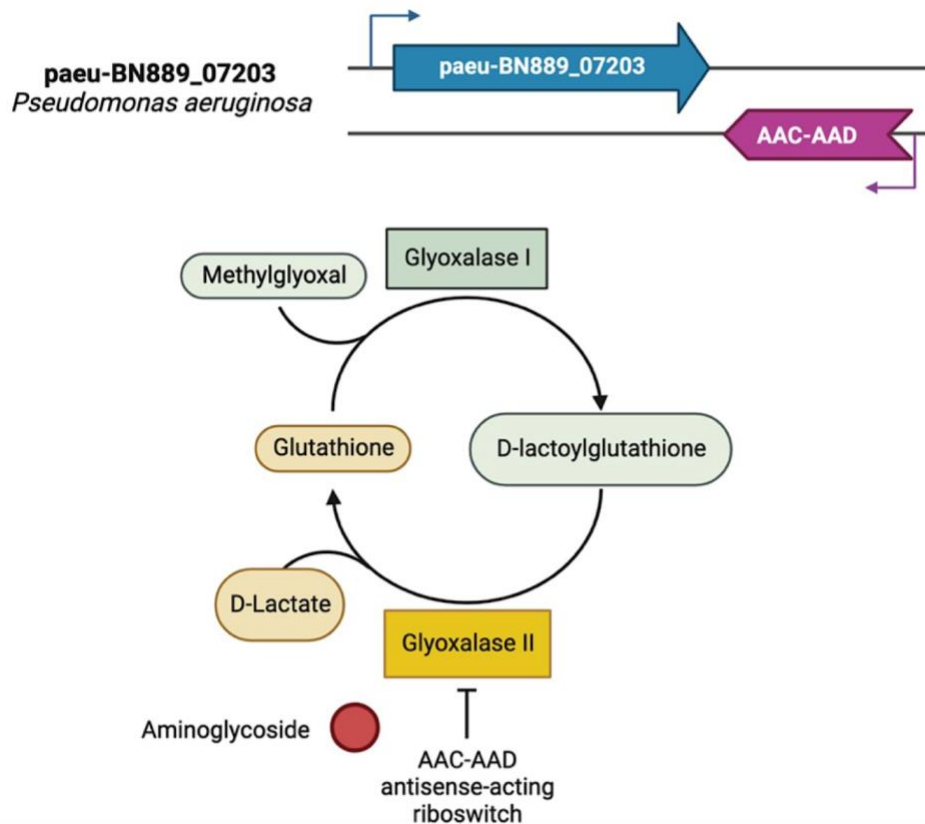


Figura 17. Gen que codifica para una enzima involucrada en un sistema de señalización celular relacionado con el metabolito sensado por el riboswitch (Tomado de Serrano-Gutiérrez & Merino, 2023).

La primera enzima de este sistema es la glioxilasa I (GlxI) o S-D-lactoil-glutación metilglioxal liasa, misma que es una reductasa isomerizante divalente que puede ser Zn^{2+} dependiente o no Zn^{2+} dependiente (activado por Ni^{2+}/Co^{2+}). GlxI acepta MG como un glutación hemiacetal (GSH) y cataliza la transferencia de protones, transformando GSH hacia su tioéster correspondiente, lactoglutación. La segunda enzima, GlxII, regenera el cosustrato tiol GSH y el α -hidroxiácido lactato, en el caso de la detoxificación de MG. El MG es altamente tóxico para las moléculas productoras de especies electrofílicas celulares, y producen estrés celular mediante reacciones con macromoléculas nucleofílicas. El MG también actúa como un

activador transcripcional y participa en la cascada de activación de cinasas, teniendo un papel primario en procesos de regulación o ajuste a condiciones de estrés, haciéndolo una molécula señalizadora esencial (Sukdeo et al., 2004; Sukdeo & Honek, 2007).

Durante el proceso patogénico de *P. aeruginosa*, la bacteria es primero atacada por el sistema inmune innato, primeramente, vía fagocitos. Al interior de los fagocitos, se generan ROS mediante un proceso dependiente de NADPH oxidasa como sustancias bactericidas. Entonces, las células bacterianas son expuestas a estrés oxidativo, induciendo daño debido a las interacciones con elementos celulares tales como lípidos, DNA y proteínas. Este proceso lleva a la peroxidación lipídica, mutación de DNA, entrecruzamiento de DNA-proteína, oxidación y ruptura de proteínas. *P. aeruginosa* ha evolucionado mecanismos para protegerse y sobrevivir bajo estas condiciones hostiles. Una de ellas es la producción de GSH, la molécula antioxidante más abundante dentro de las células, la cual juega un papel relevante en la remoción de ROS, actuando como un donador de electrones. El GSH también está involucrado en la preservación de la homeostasis celular, la regulación del transporte de azufre, la combinación de metabolitos, la detoxificación de xenobióticos, la resistencia a antibióticos, la regulación enzimática y la expresión de genes de respuesta a estrés (Lee & Park, 2017; Wongsaroj et al., 2018).

Los compuestos que producen estrés oxidativo celular, tal como los antibióticos, y sus efectos sobre las diferentes etapas del desarrollo bacteriano han sido ampliamente estudiados. Estos compuestos incrementan la sensibilidad de *P. aeruginosa* a los aminoglicósidos afectando la permeabilidad de la membrana,

principalmente durante la producción de biofilm, misma que es estimulada por concentraciones subletales de antibióticos tales como los aminoglicósidos tobramicina y gentamicina. Además, el estrés oxidativo actúa como una señal para la expresión de sistemas de eflujo, lo cual promueve la resistencia bacteriana a antibióticos (da Cruz Nizer et al., 2021; Hall et al., 2016; Lochab et al., 2020; Pan et al., 2017).

El riboswitch antisentido identificado en el extremo 3' del gen para-BN889_07203 podría permitir la expresión de la enzima GlxII cuando se sense un aminoglicósido en el ambiente y las células bacterianas se encuentran bajo condiciones de estrés. Por otro lado, cuando no hay aminoglicósido presente en el medio, la expresión de la enzima es reprimida, ya que la bacteria no es sujeta a estrés oxidativo provocado por la presencia del antibiótico.

Este quinto grupo de genes codifican para proteínas que están principalmente involucradas en diferentes respuestas a estrés y procesos de desarrollo, tales como la síntesis de autoinductores de quorum sensing y diferentes tipos de histidinas transductoras de señales. Los principales riboswitches antisentido de este grupo son aquellos de las familias sensoras de segundos mensajeros c-di-GMP y ydaO-yuaA (Jenal et al., 2017).

e) Genes que codifican para reguladores transcripcionales.

Se identificó un riboswitch antisentido de la clase c-di-GMP-I regulando el gen paen-P40081_02330 que codifica un regulador transcripcional de la familia XRE en el genoma de *Paenibacillus sp.* FSL P4 (Figura 18). Este regulador transcripcional, cuyo homólogo más cercano es SinR, un represor transcripcional conocido como el

regulador maestro para la formación de biofilm en *Bacillus subtilis*, controla la expresión de genes involucrados en producción de exopolisacáridos, proteínas de matriz y motilidad (Chen et al., 2012; Sudarsan et al., 2008).

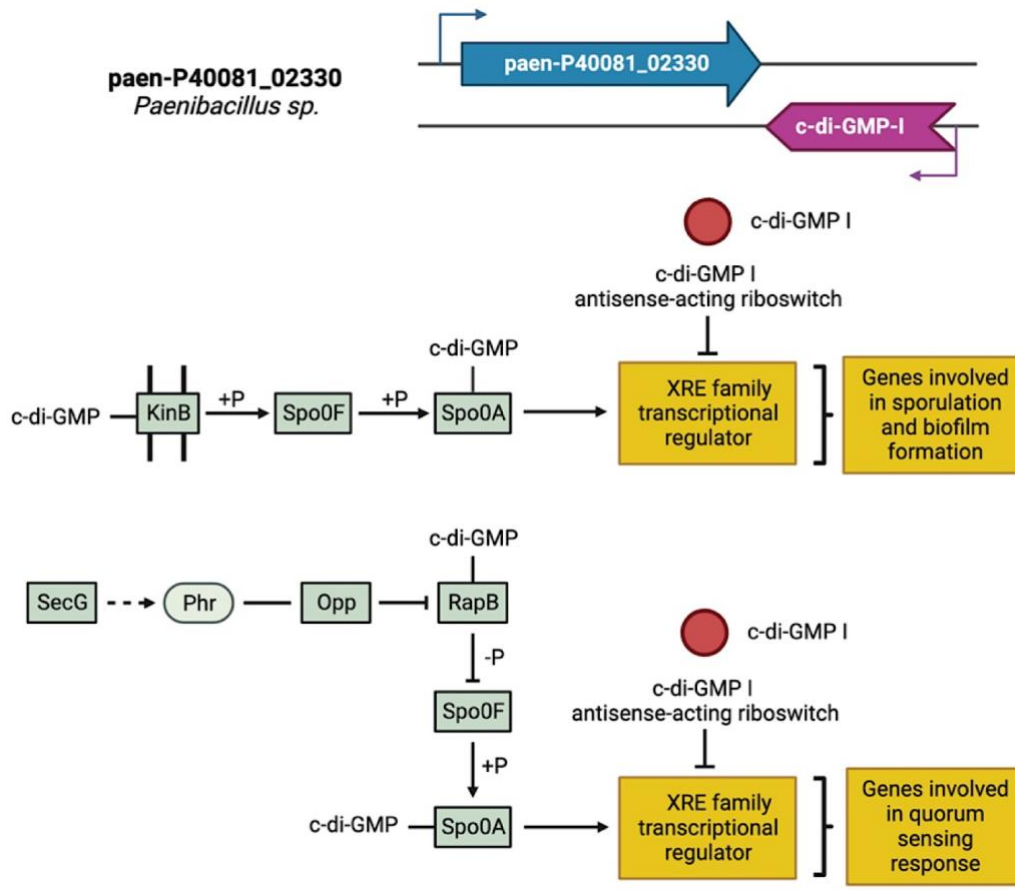


Figura 18. Gen codificante de un regulador transcripcional relacionado con el metabolito sentido por el riboswitch (Tomado de Serrano-Gutiérrez & Merino, 2023).

La síntesis de moléculas señaladoras es una estrategia que las bacterias emplean para sentir alteraciones en su ambiente con el propósito de ajustarse rápidamente de acuerdo a tales cambios. Una de estas moléculas es el bis-(3'-5')-GMP dimérico cíclico (c-di-GMP), el cual es capaz de inducir la transición de un estado móvil unicelular a un estado sésil multicelular mediante sistemas de

fosforelevo que controlan la formación de biofilm y esporas (Reyes-Darias & Krell, 2017; Sornchuer et al., 2022; Tamayo, 2019).

Para muchas bacterias, el mecanismo de comunicación intercelular de quorum sensing es esencial para la coordinación de tendencias de expresión de genes a nivel poblacional. Además, el quorum sensing es fundamental para los procesos de motilidad, producción de matriz de biofilm y formación de esporas, ya que es una de las principales estrategias para sensor y responder a cambios ambientales (Georg & Hess, 2011; Rahman et al., 2021; Voichek et al., 2020; Weiss et al., 2019).

La regulación del factor transcripcional codificado por el gen *paen-P40081_02330* mediante el riboswitch antisentido de c-di-GMP-I que se propone está basada en la detección de esta molécula, desencadenando la expresión de genes involucrados en la formación de biofilm y esporas, tal como en los procesos de quorum sensing, y teniendo un efecto general sobre su expresión cuando no es sentido gracias a su estrecha relación con el regulador maestro de estos procesos, Spo0A (Kearns et al., 2005).

Tal como previamente se mencionó, en la literatura, existen únicamente dos ejemplos de regulación de la expresión de genes mediada por riboswitches que actúan en antisentido, uno de los cuales es el riboswitch en antisentido de cobalamina identificado al extremo 3' del gen que codifica para el regulador transcripcional PocR de las enzimas involucradas en el metabolismo de propanediol en *L. monocytogenes*, las cuales requieren de cobalamina como un cofactor para ser activas. En este proyecto, se identificaron riboswitches antisentido de cobalamina regulado genes ortólogos de *pocR* en otras Listerias y en el Enterococo

Vagococcus carniphilus. La regulación de genes que codifican factores transcripcionales (TFs) hace posible extender el efecto regulatorio de los riboswitches que actúan en antisentido hasta múltiples genes funcionalmente relacionados en diferentes operones, que tienen en común el sitio operador de sus correspondientes TFs.

Las familias de riboswitches antisentido más frecuentes en este grupo fueron c-di-GMP, Cobalamina y TPP.

f) Genes codificantes de transposasas.

A diferencia de los ejemplos anteriormente mencionados, donde la regulación de la expresión de genes mediada por riboswitches antisentido representa una ventaja selectiva, se propone que en el caso de los genes que codifican para transposones, la existencia de riboswitches antisentido corresponde a inserciones aleatorias del elemento transponible en las regiones del cromosoma donde los riboswitches se encontraban previamente regulando otros genes de manera canónica (Tabla 4). Consistente con esta hipótesis, se llevó a cabo la evaluación del contexto genómico de los genes regulados por riboswitches antisentido en los organismos filogenéticamente más cercanos. En los genomas de todos esos organismos, el riboswitch antisentido de la clase T-box regula el gen río abajo que codifica para un transportador de triptófano de manera canónica (Figura 19).

Tabla 5. Genes codificantes para transposasas adyacentes a riboswitches antisentido
(Extracto de Serrano-Gutiérrez & Merino, 2023)
<https://doi.org/10.1371/journal.pone.0281744.s004>.

Gen de transposasa	Riboswitch antisentido	Gen adyacente	Dirección de transcripción	COG	Descripción de COG	Consistencia
bthu-YBT1518_08170	T-box	bthu-YBT1518_08175	Sense	COG0733	Sodium-dependent transporter, putative	Yes
bthu-YBT1518_10950	T-box	bthu-YBT1518_10955	Sense	COG0460	Homoserine dehydrogenase	Yes
bthu-YBT1518_24240	T-box	bthu-YBT1518_24245	Sense	COG3186	Phenylalanine 4-monooxygenase	Yes
dfg-B0537_08990	T-box	dfg-B0537_08995	Sense	COG2984	ABC transporter substrate-binding protein	Yes
nth-Nther_0169	T-box	nth-Nther_0170	Sense	COG1045	Serine O-acetyltransferase	Yes
pdm-ADU72_1747	T-box	pdm-ADU72_1748	Sense	COG1135	Methionine ABC transporter ATP-binding protein	Yes
pio-PI20285_09230	T-box	pio-PI20285_09225	Sense	COG4720	DUF3816 family protein	Yes
tbz-BK011_04395	T-box	tbz-BK011_04390	Sense	COG0441	Threonine--tRNA ligase	Yes
tbz-BK011_08290	T-box	tbz-BK011_08285	Sense	COG0162	Tyrosine--tRNA ligase	Yes
tep-TepRe1_2063	T-box	tep-TepRe1_2062	Sense	COG0162	Tyrosyl-tRNA synthetase	Yes
too-C7K38_05890	T-box	too-C7K38_05895	Sense	COG0620	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	Yes
twi-Thewi_1651	T-box	twi-Thewi_1650	Sense	COG0016	Phenylalanyl-tRNA synthetase alpha chain	Yes
efl-EF62_0727	T-box	efl-EF62_0728	Sense	COG0620	Cobalamin-independent synthase, Catalytic domain protein	Yes
lko-ABN16_04995	T-box	lko-ABN16_05000	Sense	COG4166	Peptide ABC transporter substrate-binding protein	Yes
eac-EAL2_c00320	TPP	eac-EAL2_c00330	Sense	COG2104	Sulfur transfer protein involved in thiamine biosynthesis	Yes
fpd-CTM68_07135	TPP	fpd-CTM68_07140	Sense	COG0351	Bifunctional hydroxymethylpyrimidine kinase/phosphomethylpyrimidine	Yes
rid-Rldsm_05440	TPP	rid-Rldsm_05439	Sense	COG0422	Phosphomethylpyrimidine synthase	Yes
fpe-Ferpe_1699	Cobalamin	fpe-Ferpe_1700	Sense	COG2038	Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	Yes
cki-Calkr_1509	SAM	cki-Calkr_1508	Sense	COG0192	S-adenosylmethionine synthetase	Yes
tbz-BK011_03805	SAM	tbz-BK011_03800	Sense	COG0192	methionine adenosyltransferase	Yes

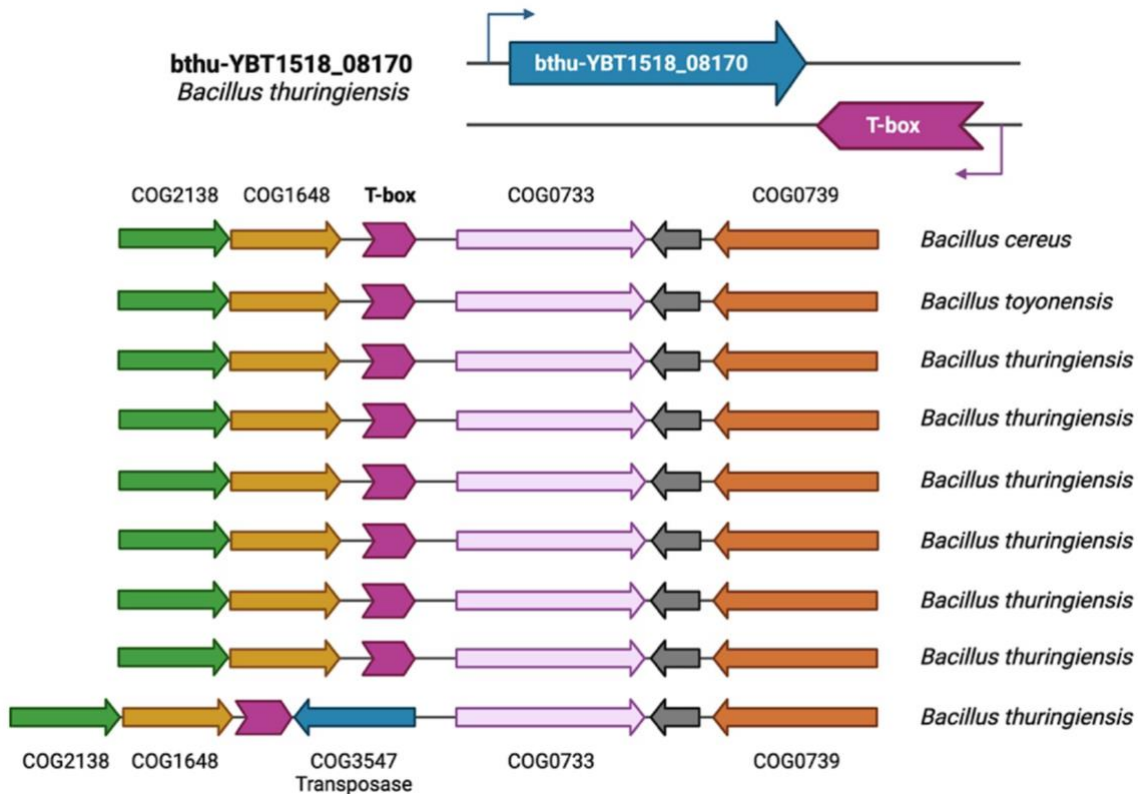


Figura 19. Gen codificante para una transposasa adyacente a un riboswitch antisentido (Tomado de Serrano-Gutiérrez & Merino, 2023).

De acuerdo con lo anterior, evaluamos la frecuencia con la cual el gen localizado río arriba de la trasposasa posee un gen ortólogo regulado por el mismo riboswitch antisentido. Se muestra que la hipótesis es consistente con su contexto genómico en más del 50 % de los casos de genes que codifican para transposasas.

g) Pseudogenes.

Otra importancia tendencia observada en nuestro análisis fue la presencia de riboswitches antisentido adyacentes a ORFs anotados como pseudogenes (Tabla 5). De manera similar a lo mencionado anteriormente, se llevó a cabo un análisis del contexto genómico de los genes regulados en antisentido y se compararon con

los genes en sus organismos filogenéticamente más cercanos (Figura 20). Con base en la conservación de la sintenia observada en todos los genes, excepto en el gen regulado en antisentido, se propone que la existencia de un riboswitch antisentido no corresponde a una ventaja selectiva para los organismos, sino a un evento fortuito que podría estar relacionado al origen del pseudogen.

Tabla 6. Pseudogenes adyacentes a riboswitches antisentido (Extracto de Serrano-Gutiérrez & Merino, 2023) <https://doi.org/10.1371/journal.pone.0281744.s005>.

Pseudogen	Riboswitch antisentido	Gen adyacente	Dirección de transcripción	COG	Descripción de COG	Consistencia
bacp-SB24_06455	Purine	bacp-SB24_06450	Sense	COG0041	N5-carboxyaminoimidazole ribonucleotide mutase	Yes
crs-FQB35_11485	TPP	crs-FQB35_11490	Sense	COG0351	bifunctional hydroxymethylpyrimidine kinase/phosphomethylpyrimidine	Yes
Inn-F0161_06915	FMN	Inn-F0161_06910	Sense	COG0117	bifunctional diaminoxyphosphoribosylaminopyrimidine deaminase/5-a	Yes
lpak-GDS87_22790	SAM	lpak-GDS87_22795	Sense	COG1878	cyclase family protein	Yes
mesm-EJ066_09470	Cobalamin	mesm-EJ066_09465	Sense	COG4638	aromatic ring-hydroxylating dioxygenase subunit alpha	Yes
mlb-MLBr02668	ykoK / M-box	mlb-MLBr02667	Sense	COG1914	possible membrane transport protein	Yes
mmal-CKJ54_12460	ykoK / M-box	mmal-CKJ54_12455	Sense	COG5651	PPE-repeat proteins	Yes
mpag-C0J29_15460	ykoK / M-box	mpag-C0J29_15465	Sense	COG5651	PPE family protein	Yes
nfv-FAH67_07590	Glycine	nfv-FAH67_07595	Sense	COG0404	glycine cleavage system aminomethyltransferase GcvT	Yes
nsi-A6J88_06455	Glycine	nsi-A6J88_06460	Sense	COG0404	glycine cleavage system protein T	Yes
oek-FFI11_001160	FMN	oek-FFI11_001165	Sense	COG3201	nicotinamide mononucleotide transporter	Yes
plp-Ple7327_1638	ykkC / Guanidine	plp-Ple7327_1637	Sense	COG0010	arginase family hydrolase, arginase/agmainase/formiminoglutamate hydro	Yes
slp-Slip_0637	ydaO-yuaA	slp-Slip_0636	Sense	COG1125	glycine betaine/L-proline ABC transporter, ATPase subunit	Yes
tbo-Thebr_0876	MOCO_RNA_motif	tbo-Thebr_0877	Sense	COG1763	molybdopterin-guanine dinucleotide biosynthesis protein B	Yes
tex-Teth514_1384	MOCO_RNA_motif	tex-Teth514_1385	Sense	COG1763	molybdopterin-guanine dinucleotide biosynthesis protein B	Yes
thx-Thet_1517	MOCO_RNA_motif	thx-Thet_1516	Sense	COG1763	molybdopterin-guanine dinucleotide biosynthesis protein B	Yes
twi-Thewi_0380	MOCO_RNA_motif	twi-Thewi_0381	Sense	COG1763	molybdopterin-guanine dinucleotide biosynthesis protein B	Yes
txy-Thexy_2017	c-di-GMP	txy-Thexy_2016	Sense	COG2206	diguanylate cyclase and metal dependent phosphohydrolase	Yes
avm-JQX13_05915	TPP	avm-JQX13_05920	Antisense	COG2059	chromate efflux transporter	Yes
bnt-GSN03_15970	ydaO-yuaA	bnt-GSN03_15975	Antisense	COG0500	methyltransferase domain-containing protein	Yes

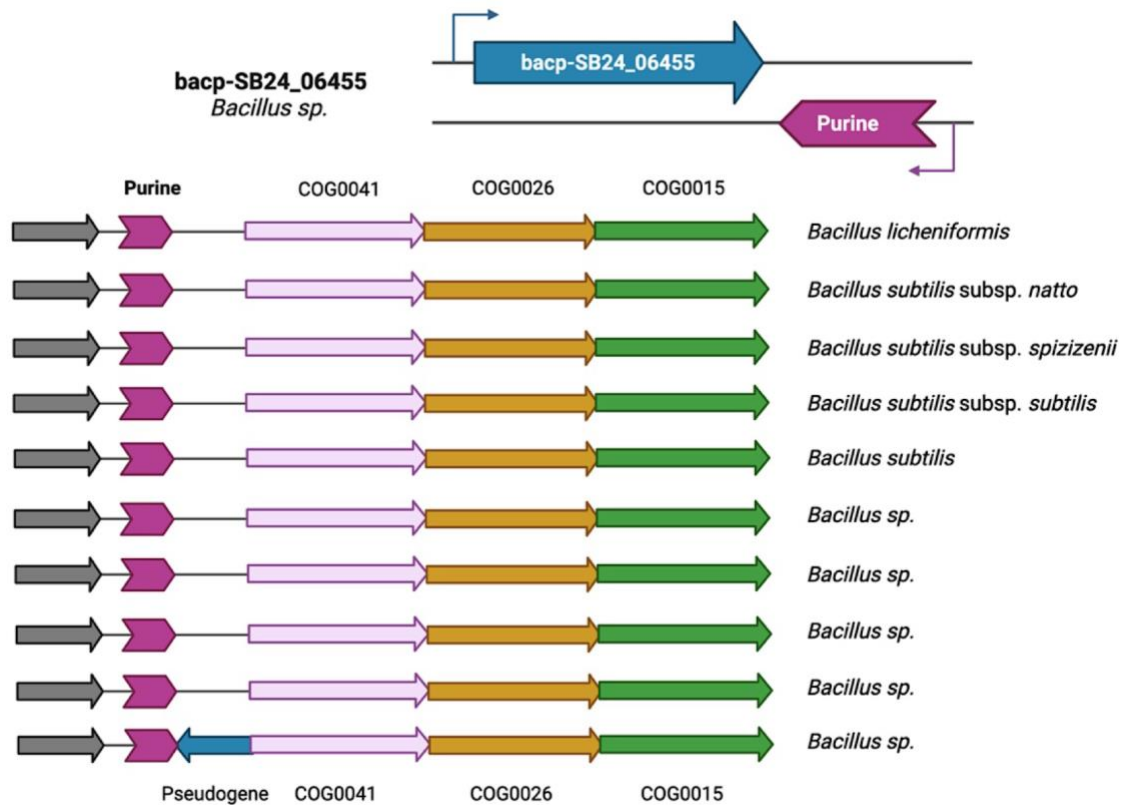


Figura 20. Pseudogen adyacente a un riboswitch antisentido (Tomado de Serrano-Gutiérrez & Merino, 2023).

h) ORFs no caracterizados que poseen ortólogos en múltiples genomas.

La siguiente categoría de genes regulada por riboswitches antisentido identificada en este trabajo corresponde a ORFs anotados como no caracterizados pero con ortólogos *bona fide* en otros genomas. En este grupo se incluyeron sólo genes más extensos que 300 nucleótidos. Se muestra un ejemplo de este gen, cuya función es aún desconocida (Figura 21).

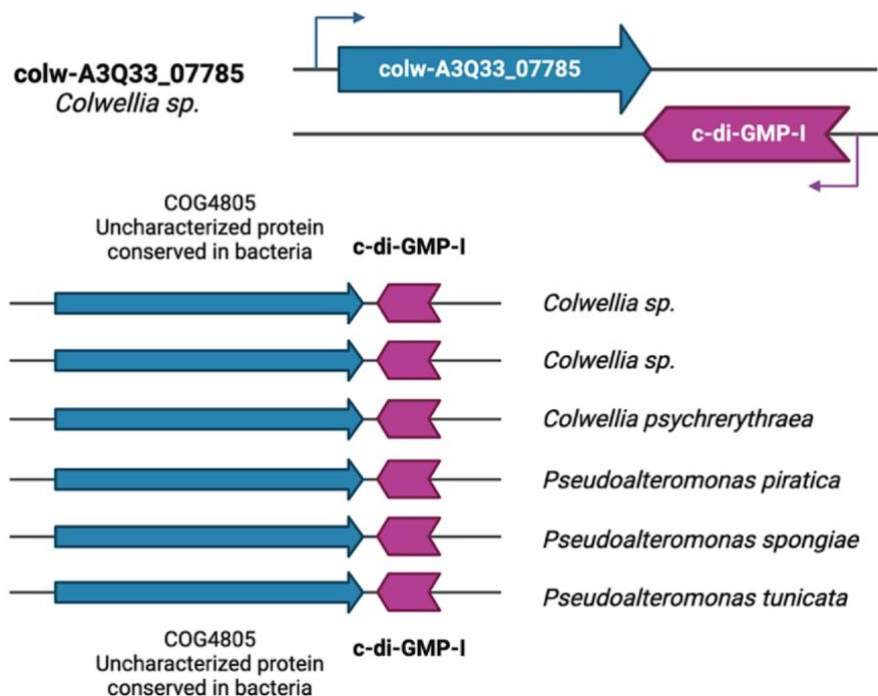


Figura 21. Gen codificante para un ORF no caracterizado adyacente a un riboswitch antisentido (Tomado de Serrano-Gutiérrez & Merino, 2023).

i) Pequeños genes hipotéticos sin ortólogos identificables.

En esta categoría se incluyen genes identificados como potencialmente regulados por riboswitches antisentido con la característica de tener un tamaño pequeño de menos de 300 bases y careciendo de ortólogos aparentes en el set de genomas secuenciados, los cuales comprenden 34% de los elementos (Tabla 6). Estas características nos permiten considerar que estos “genes hipotéticos” corresponden a probables errores de anotación de marcos abiertos de lectura que, en la mayoría de los casos, son incorrectamente anotados como genes. Tal como en nuestro previo grupo, el análisis comparativo de regiones sinténicas en organismos filogenéticamente más cercanos confirma la hipótesis acerca de que la mayoría de aquellos riboswitches (72% de los casos estudiados) regulan los genes

adyacentes río arriba de los genes hipotéticos, de acuerdo a la consistencia de su contexto genómico. Este es el caso para el pequeño gen hipotético spi-MGAS10750_Spy1129 en el genoma de *Streptococcus pyogenes* MGAS10750 con un riboswitch de Glicina en su extremo 3' y un gen D-alanina glicina permeasa adyacente, cuyos ortólogos equivalentes son regulados por riboswitches de glicina en organismos filogenéticamente cercanos (Figura 22). Se enlistan todos los genes hipotéticos así como su gen vecino regulado más probable.

Tabla 7. Genes codificantes para proteínas hipotéticas adyacentes a riboswitches antisentido (Extracto de Serrano-Gutiérrez & Merino, 2023)
<https://doi.org/10.1371/journal.pone.0281744.s006>.

Proteína hipotética	Riboswitch antisentido	Tamaño del gen	Gen adyacente	Dirección de transcripción	COG	Descripción de COG	Consistencia
lsg-lse_1049	Cobalamin	189	lsg-lse_1050	Sense	COG1454	propanol dehydrogenase	Yes
rpy-Y013_10415	Cobalamin	159	rpy-Y013_10410	Sense	COG0614	ABC transporter substrate-binding protein	Yes
rpy-Y013_14535	Cobalamin	159	rpy-Y013_14530	Sense	COG0614	ABC transporter substrate-binding protein	Yes
amed-B224_0672	c-di-GMP	129	amed-B224_0671	Sense	COG3325	chitinase	Yes
dec-DCF50_p1698	c-di-GMP	114	dec-DCF50_p1699	Sense	COG2206	HD-GYP hydrolase domain containing protein	Yes
dov-DSCO28_70960	Cobalamin	168	dov-DSCO28_70950	Sense	COG3539	P pilus assembly protein, pilin FimA	Yes
elm-ELI_4146	c-di-GMP	207	elm-ELI_4145	Sense	COG2165	Type II secretory pathway, pseudopilin PulG	Yes
pib-BBD41_14935	c-di-GMP	189	pib-BBD41_14940	Sense	COG4719	Uncharacterized protein conserved in bacteria	Yes
abas-ACPOL_5155	Cobalamin	123	abas-ACPOL_5154	Sense	COG2087	Adenosylcobinamide-phosphate guanylyltransferase	Yes
bhl-Bache_2223	Cobalamin	159	bhl-Bache_2224	Sense	COG1629	TonB-dependent receptor	Yes
bvl-BF3285c2_0291	Cobalamin	156	bvl-BF3285c2_0290	Sense	COG0614	iron compound ABC transporter periplasmic iron compound-binding protein	Yes
cko-CKO_00803	Cobalamin	153	cko-CKO_00802	Sense	COG1797	Cobyrinic acid a,c-diamide synthase	Yes
dhy-DESAM_21377	Cobalamin	189	dhy-DESAM_21378	Sense	COG0209	Ribonucleoside-diphosphate reductase	Yes
dmx-X792_03180	Cobalamin	105	dmx-X792_03185	Sense	COG2038	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	Yes
dti-Desti_5588	Cobalamin	162	dti-Desti_5589	Sense	COG2087	adenosyl cobinamide kinase/adenosyl cobinamide phosphate guanylyltrans	Yes
gak-X907_0913	Cobalamin	123	gak-X907_0914	Sense	COG0614	iron compound ABC transporter periplasmic substrate-binding protein	Yes
gdi-GDI2573	Cobalamin	165	gdi-GDI2572	Sense	COG0620	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferas	Yes
hmc-HYPMC_0280	Cobalamin	228	hmc-HYPMC_0281	Sense	COG4206	TonB-dependent receptor plug	Yes
ibu-IB211_00811	Cobalamin	249	ibu-IB211_00810c	Sense	COG1520	FOG: WD40-like repeat	Yes
lab-LA76x_4124	Cobalamin	279	lab-LA76x_4125	Sense	COG0620	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferas	Yes

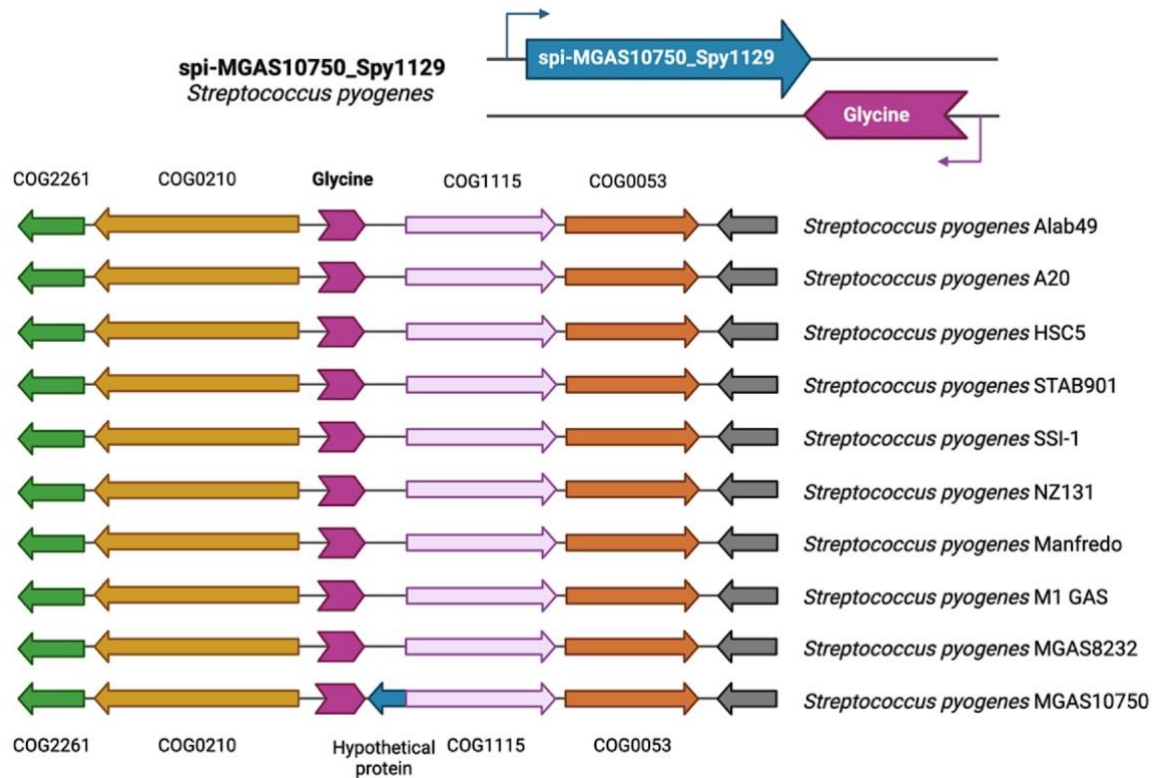


Figura 22. Gen codificante para una proteína hipotética adyacente a un riboswitch antisentido (Tomado de Serrano-Gutiérrez & Merino, 2023).

j) Otro tipo de genes.

Finalmente, el último grupo del análisis está compuesto por riboswitches antisentido, identificados adyacentes a genes que no pudieron ser clasificados en ninguna de los grupos previamente descritos. Los genes de esta categoría representan aproximadamente 9% de los riboswitches antisentido identificados en este estudio y las implicaciones bioquímicas de su regulación no resultan obvias a primera instancia.

Para cada una de las clases establecidas y tomando en cuenta toda la información disponible en literatura y en diversas bases de datos como GeConT (<http://biocomputo.ibt.unam.mx:8080/GeConT/>),

KEGG

(<https://www.genome.jp/kegg/>), STRING (<https://string-db.org/>), y PFAM (<http://pfam.xfam.org/>), se realizaron modelos regulatorios que pudieran ejemplificarlas, considerando las posibles implicaciones bioquímicas de su regulación.

Por otro lado, también se identificaron algunos genes regulados en antisentido por riboswitches en tándem. Este término corresponde a aquellos en los que existe una o más copias de la misma o diferentes clases de este elemento regulatorio con el propósito de tener una regulación más compleja que responde a más de una señal intracelular.

Apéndice 1. Localización de probables errores de anotación en el genoma de bacterias y arqueas basado en la identificación de riboswitches en regiones codificantes

Durante el proceso de búsqueda de riboswitches antisentido, se obtuvieron resultados interesantes, mismos que eran sugerentes de la identificación de probables errores de anotación en determinados genes de los organismos evaluados. Tomando en cuenta lo anterior, se estableció un protocolo de búsqueda que nos permitió identificar motivos de riboswitches que se encontraban localizados de manera total o parcial en regiones anotadas como codificantes, mismo que resulta incompatible con la definición canónica de un riboswitch.

Al proceder con la realización de un análisis minucioso de estos riboswitches, fue posible encontrar ciertas tendencias entre ellos, con base en la información obtenida para cada uno de estos elementos. Por ello, se tomaron estos resultados con el fin de clasificarlos considerando su localización, es decir, si fueron identificados de manera total o parcial en una región anotada como codificante para un gen, y se establecieron tres categorías de riboswitches localizados en regiones codificantes potencialmente mal anotadas. Estas categorías fueron establecidas considerando tanto la proporción de la secuencia del riboswitch compartida con la secuencia del ORF en el cual se identificó, como su ubicación en el extremo 5' o 3'; gracias a lo cual pudimos también subclasificar estos elementos sugerentes de probables errores de anotación en dichos genes.

Errores en la anotación de genomas

Como antecedente del análisis que se propone, se sabe que durante las últimas décadas ha existido un incremento exponencial en el número de genomas secuenciados y anotados en diferentes bases de datos, mismas que se encuentran disponibles para quien desee hacer uso de dicha información. Dada la necesidad del uso de herramientas computacionales con propósito de llevar a cabo la anotación o reanotación de secuencias genómicas, también ha existido un incremento en el desarrollo de métodos que tienen el fin de hacer de este un proceso automatizado y con alta calidad; sin embargo, las técnicas de secuenciación y procesamiento de datos pueden dar lugar a la existencia de errores (Devos & Valencia, 2001; Jones et al., 2007; Lin et al., 2008).

Aunado al avance en la secuenciación y anotación de genomas, se ha dado también un incremento en los estudios dedicados a conocer la tasa de error en los mismos, trayendo a la luz evidencia que demuestra la existencia de una gran prevalencia de errores en la secuenciación y/o anotación de genomas. Algunas de las características observables en secuencias genómicas que pueden ser evidencia de la existencia de errores de anotación puede ser la gran cantidad de ORFs cortos y sin ortólogos identificables anotados en las bases de datos, la identificación de nuevos motivos conservados en secuencias anotadas como codificantes, entre otros. A pesar de esto, al no llevarse a cabo métodos experimentales para sustentar la información obtenida mediante predicciones computacionales, se consideraba que no era posible afirmar la identidad de dichas únicamente haciendo uso de la

identificación *in silico* (Amin et al., 2018; Brenner, 1999; Linial, 2003; Starmer et al., 2005; Weinmaier et al., 2016).

Así pues, considerando las evidencias que es posible identificar en las secuencias genómicas y el hecho de que la disponibilidad de los datos abre un amplio panorama en el análisis de distintos elementos genéticos, se ha propuesto el desarrollo de diferentes métodos mediante los cuales es posible encontrar una solución a la existencia de dichos errores, partiendo desde el establecimiento de protocolos más estrictos en el análisis y anotación de secuencias, continuando con la identificación de errores en las secuencias ya anotadas y finalmente, llegar a la propuesta de un método para la reanotación de dichas secuencias corregidas. El establecimiento de estos métodos ya se ha dado en algunos casos, ya sea de manera fortuita y como resultado de la búsqueda de diferentes motivos genéticos conservados, pero también mediante el desarrollo de protocolos con el fin de buscar una solución a este problema (Amin et al., 2018; Lin et al., 2008; Linial, 2003; Starmer et al., 2005; Weinmaier et al., 2016).

Con el propósito de tener un mejor entendimiento acerca del contexto de los riboswitches identificados, se procedió a establecer una estrategia para su análisis a mayor profundidad, tomando en cuenta toda la información disponible concerniente a:

a) el tamaño del gen que se encontraran regulando o interrumpiendo, respectivamente;

b) la frecuencia con la que genes ortólogos son regulados en otros organismos por el riboswitch de estudio.

Este criterio, correspondiente a la conservación de genes regulados por el mismo tipo de elemento regulador, ha sido previamente considerado en estudios que tienen como objetivo la identificación de errores tomando en cuenta la anotación funcional de secuencias. En nuestro estudio, la identificación de genes ortólogos se realizó haciendo uso de sus anotaciones dentro de la base de datos COG.

En el caso de los riboswitches (en sentido y antisentido) identificados en regiones intergénicas (Figura 23), el tamaño de los genes regulados por los mismos, presenta una distribución que va desde secuencias cortas, hasta ORFs con un tamaño aproximado de 3000 pb; sin embargo, la gran mayoría poseen un tamaño de entre 1000 y 1500 pb. Por otra parte, al analizar estas secuencias con respecto a su asignación en el sistema COG, el 92%, posee una clasificación dentro del mismo.

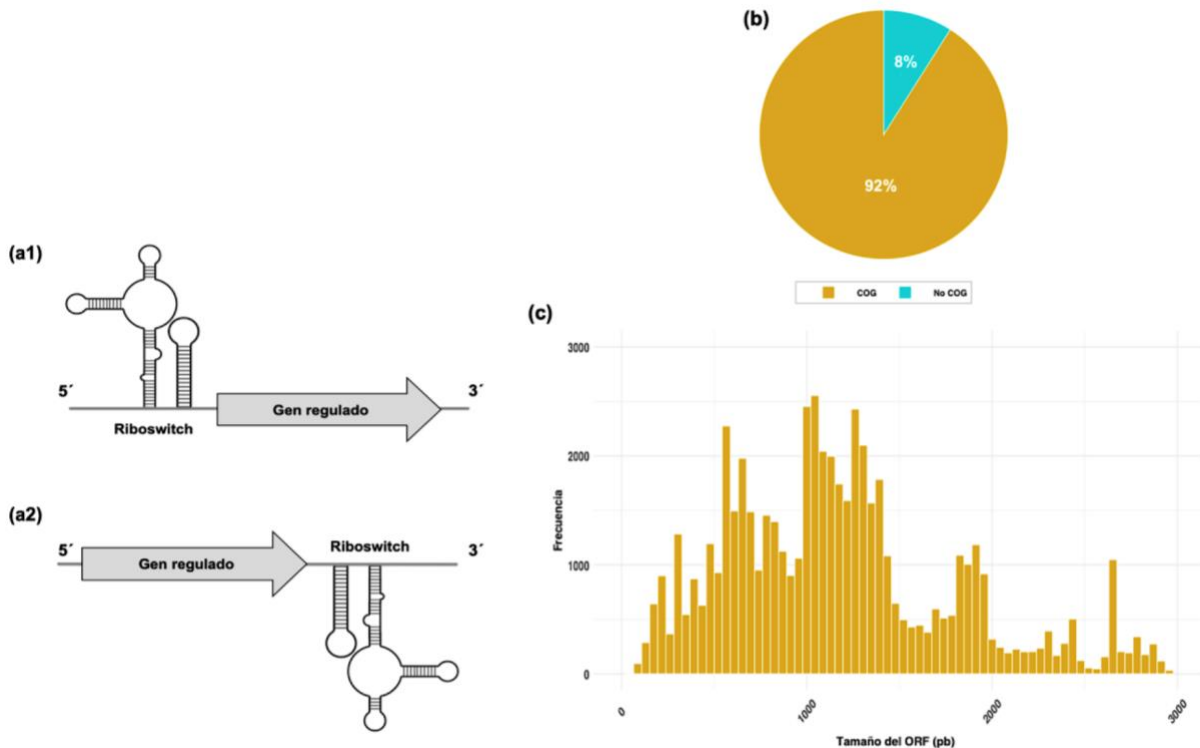


Figura 23. Análisis de riboswitches canónicos. (a1) Riboswitch en sentido, (a2) Riboswitch antisentido, (b) Proporción de asignación COG de genes regulados por riboswitches canónicos, (c) Distribución de tamaño de los genes regulados por riboswitches canónicos.

De manera contrastante, cuando estas dos evaluaciones se realizaron considerando los resultados obtenidos para los riboswitches identificados en regiones codificantes, se apreciaron dos interesantes variantes. El primero corresponde a aquellas secuencias que presentan un comportamiento similar al observado para los genes regulados por riboswitches en regiones intergénicas, es decir, con una prevalencia de tamaño de entre 1000 y 1500 pb; y una proporción de asignación en el sistema COG de arriba del 90% (semi-internos) (Figura 24) o alrededor del 50% (internos) (Figura 25). El otro tipo de comportamiento observado corresponde a secuencias que poseen tamaños menores a 300 pb, mismos que pueden ser considerados como ORFs cortos. Es importante resaltar que estas

secuencias tienen en común una baja tasa de asignación en el sistema COG, siendo ésta de tan sólo 11%, en el caso de los semi-internos, 5% en los internos y 0% en aquellos que contienen un ORF (Figura 26). En análisis previos, la existencia de ORFs cortos se ha relacionado con la existencia de errores de anotación. Aunado a lo anterior, la falta de conservación entre los organismos con dichas secuencias, observado por la ausencia de asignación dentro de la base de datos COG, constituye evidencia importante que apoya nuestra hipótesis acerca de que la mayoría de los genes aparentemente interrumpidos por riboswitches, poseen errores de anotación.

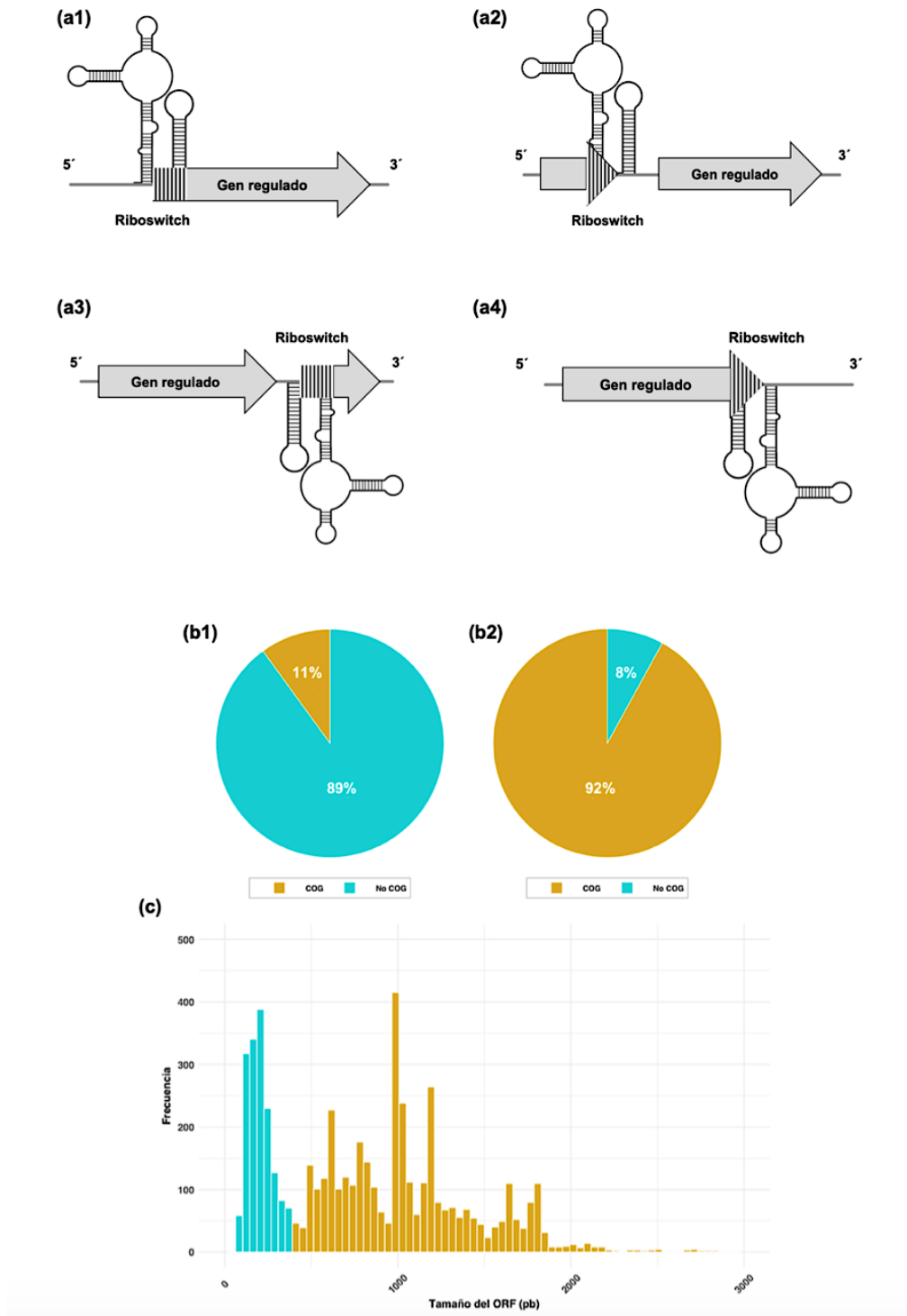


Figura 24. Análisis de riboswitches semi-internos. (a1) Riboswitch en sentido sobre el inicio de su gen regulado, (a2) Riboswitch en sentido en el gen adyacente a su gen regulado, (a3) Riboswitch antisentido en el gen adyacente a su gen regulado, (a4) Riboswitch antisentido sobre el fin de su gen regulado, (b) Proporción de asignación COG de genes regulados por riboswitches semi-internos, (c) Distribución de tamaño de los genes regulados por riboswitches semi-internos.

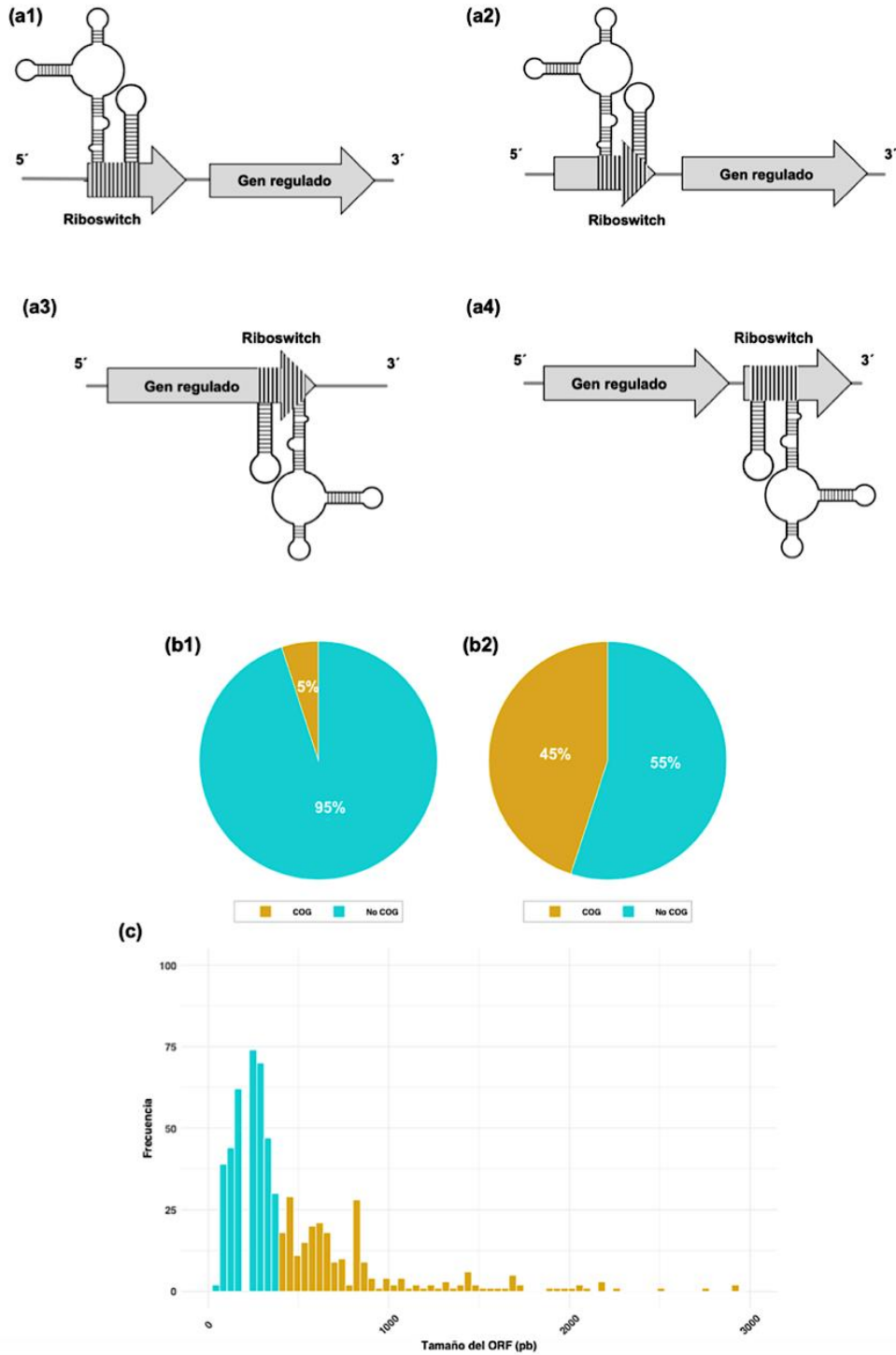


Figura 25. Análisis de riboswitches internos. (a1) Riboswitch en sentido sobre el inicio de su gen regulado, (a2) Riboswitch en sentido en el gen adyacente a su gen regulado, (a3) Riboswitch antisentido en el gen adyacente a su gen regulado, (a4) Riboswitch antisentido sobre el fin de su gen regulado, (b) Proporción de asignación COG de genes regulados por riboswitches internos, (c) Distribución de tamaño de los genes regulados por riboswitches internos.

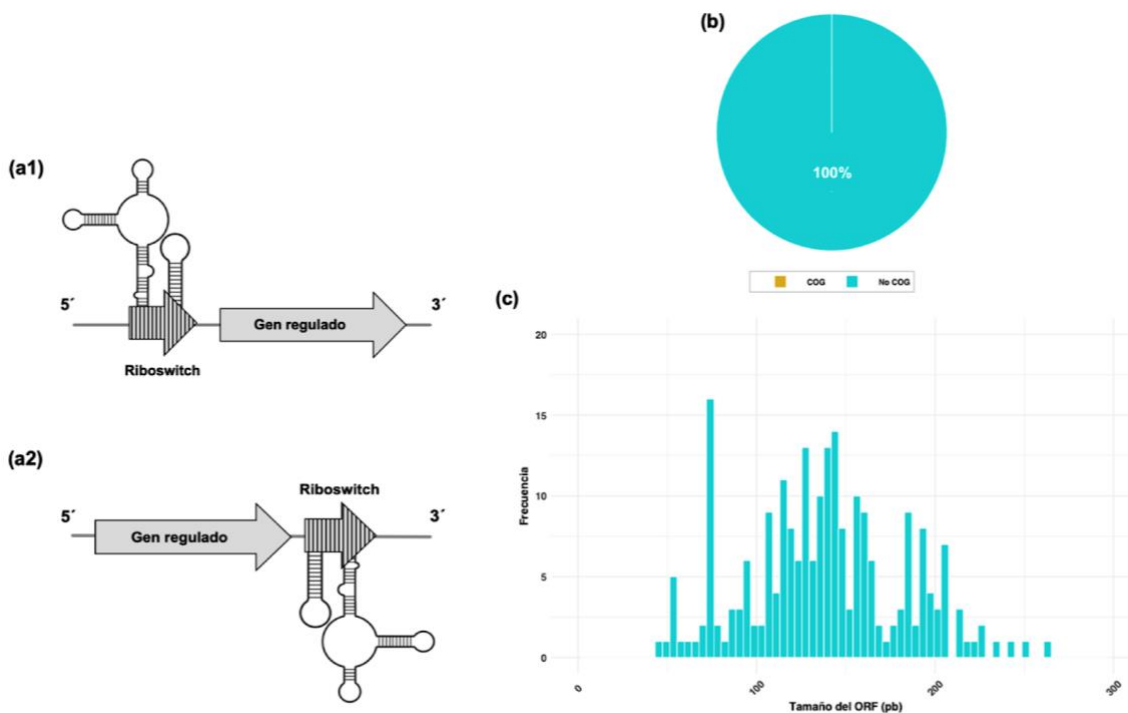


Figura 26. Análisis de riboswitch conteniendo un ORF. (a1) Riboswitch en sentido en el gen adyacente a su gen regulado, (a2) Riboswitch antisentido en el gen adyacente a su gen regulado, (b) Proporción de asignación COG de genes regulados por riboswitches conteniendo un ORF, (c) Distribución de tamaño de los genes regulados por riboswitches conteniendo un ORF.

Al englobar los resultados correspondientes a riboswitches identificados en regiones codificantes, se procedió a hacer un análisis de la frecuencia de los probables errores encontrados con respecto a la familia de riboswitch a la cual pertenecen. Lo anterior se llevó a cabo con el propósito de conocer las familias que, debido a la existencia de dichos errores, se encontraban subrepresentados al compararlos con el total de riboswitches identificados. Este comportamiento resultó ser prevalente en dos familias de riboswitches: MOCO y AAC-AAD (Aminoglicósidos).

Posteriormente, se decidió considerar ejemplos que fuesen representativos de los diferentes tipos de probables errores de anotación con base en la identificación de motivos de cada clase de riboswitch y, mediante el uso del servidor GeConT se hizo una comparación gráfica de su contexto genómico con el de otros genes pertenecientes a organismos filogenéticamente cercanos. Una vez seleccionados, se procedió a realizar alineamientos de las secuencias contrastantes en el servidor MUSCLE (<https://www.ebi.ac.uk/Tools/msa/muscle/>), gracias a lo que se logró identificar un desfase entre las secuencias de dichos genes ortólogos, lo cual se traduce en una asignación incorrecta en el codón de inicio de la transcripción en el caso de los genes *lhl-LBHH_1649* (Figura 27), *msg-MSMEI_0774* y *msg-MSMEI_0775* (Figura 29), o bien del codón de paro de la transcripción del gen *aamy-GFC30_98* (Figura 28).

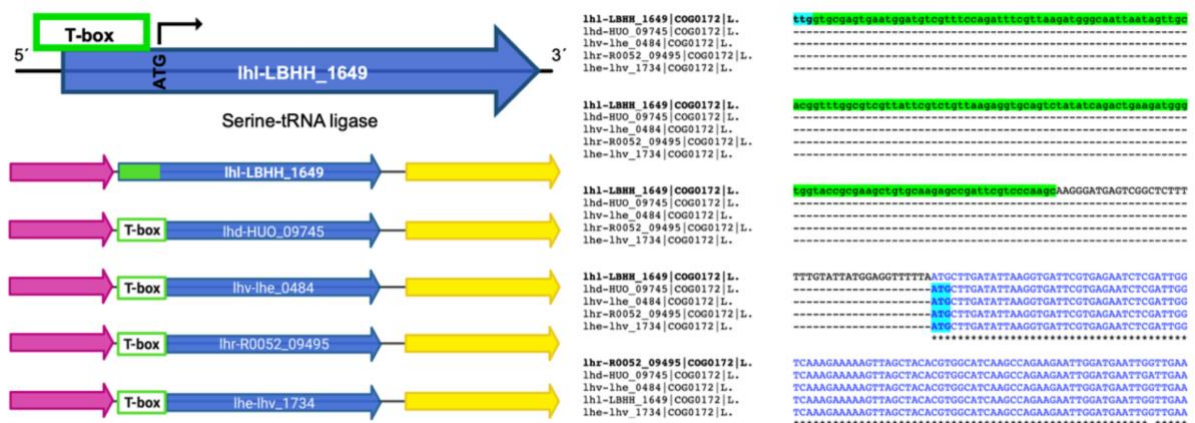


Figura 27. Localización de un probable error de anotación en el genoma basado en la identificación de un riboswitch semi-interno en el inicio de su gen regulado.

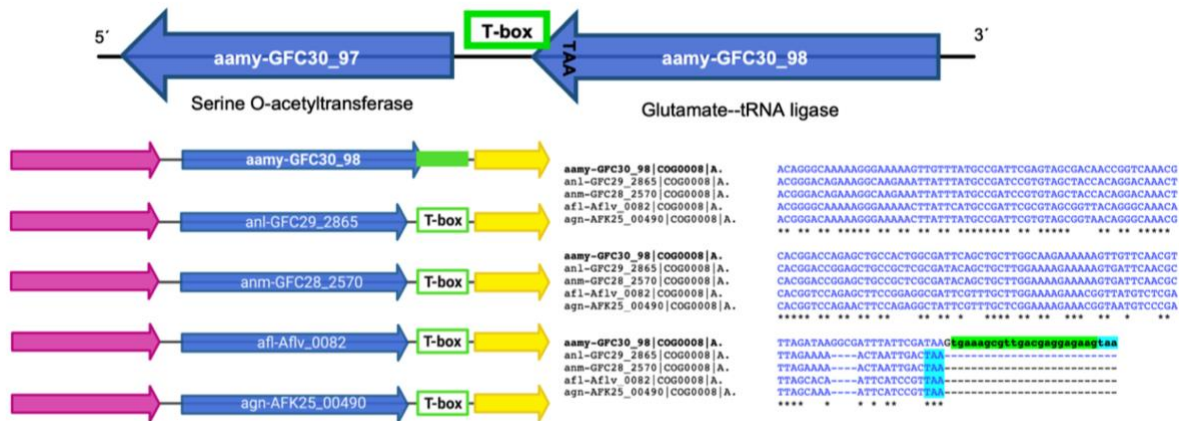


Figura 28. Localización de un probable error de anotación en el genoma basado en la identificación de un riboswitch semi-interno en el fin del gen adyacente a su gen regulado.

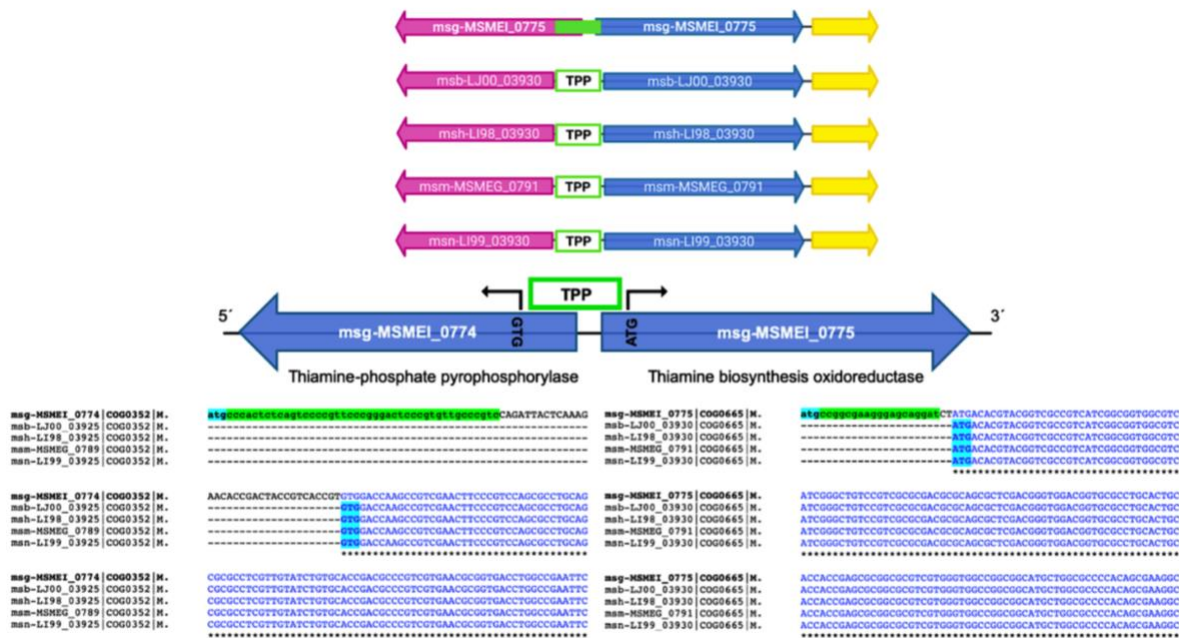


Figura 29. Localización de un probable error de anotación en el genoma basado en la identificación de un riboswitch semi-interno en el inicio de su gen regulado y el inicio del gen adyacente al mismo.

DISCUSIÓN

Durante el presente proyecto se llevó a cabo la identificación y el análisis de un tipo muy particular de motivo regulador de RNA sumamente conservado, los riboswitches. El trabajo se llevó a cabo con la particularidad de que la búsqueda de los mismos se realizó cuando estos se encuentran localizados en antisentido a su gen blanco de regulación y llevando a cabo su función de manera no canónica. Lo anterior es debido a que se ha descrito que los riboswitches que se localizan en sentido reprimen la expresión de sus genes blanco de regulación en presencia de su metabolito específico y, por el contrario, en el caso de los riboswitches antisentido esta expresión se ve favorecida cuando se sensa la presencia de su metabolito afín (Serrano-Gutiérrez & Merino, 2023; W. C. Winkler, 2005). Adicionalmente, los riboswitches antisentido tienen el efecto de bloquear la expresión de los genes mediante un mecanismo de interferencia con la elongación de la transcripción.

El proceso de trabajo que llevamos a cabo, hizo uso de todas las herramientas con las que cuenta el grupo de trabajo, además se establecieron protocolos que fuesen pertinentes para cumplir con los objetivos establecidos, lo cual sin duda representó un reto importante para el proyecto al trabajar con una clase especial de elemento genético del cual únicamente se contaba con 2 descripciones previas (André et al., 2008; Mellin et al., 2013).

Una vez completada la primera fase del proyecto, que constaba de la identificación primaria de los riboswitches en sentido y antisentido, se logró tener un incremento del número de estos últimos, el cual pasó de los 2 elementos conocidos a un total de más de 1000 nuevos riboswitches antisentido identificados. Sin

embargo, al continuar con el proceso de análisis, fue posible notar que una importante proporción de los genes que serían regulados por dichos riboswitches antisentido, se trataban de genes codificantes para proteínas anotadas como hipotéticas sin ortólogos identificables. Lo anterior resultó una de las primeras pautas para el curso del análisis y la clasificación de los genes regulados por estos elementos.

De tal manera, se decidió agrupar los genes regulados por riboswitches antisentido en 10 diferentes clases, mismas que se mencionaron con anterioridad en el apartado de resultados de este trabajo. En estas clases se consideran las funciones de los genes, lo cual se logró identificar utilizando toda la información disponible, tanto a nivel bioinformático como en la literatura que pudo encontrarse disponible para cada uno de los elementos, razón por la que este fue uno de los puntos más complejos del proyecto. Por otra parte, también se agruparon aquellos genes que no tenían una función aparente que justificara la razón biológica de la existencia de un riboswitch antisentido adyacente al mismo, tal como el caso de las transposasas o los pseudogenes.

Al llevar a cabo un análisis más profundo de cada una de las clases de genes reguladas por riboswitches antisentido, resulta interesante resaltar diversos puntos, tal como lo es el caso de la gran relevancia de la regulación de genes involucradas en la interconversión de compuestos. Este podría comprender el caso más clásico de la regulación por un riboswitch antisentido; lo anterior debido a que estaríamos apreciando la regulación opuesta a lo que se observa en un riboswitch en sentido. Esto ocurre cuando se enciende la expresión del gen de interconversión del

compuesto con la presencia del metabolito específico. Además, este tipo de gen comprende la primera descripción hecha de un riboswitch antisentido (André et al., 2008).

Otro punto interesante a resaltar es lo que ocurre en el caso de los genes o rutas metabólicas que requieren de la participación de cofactores para que su función se lleve a cabo. La especial importancia de esto radica en que, dichos cofactores suelen ser moléculas cuya biosíntesis o consumo es sumamente costoso metabólicamente para el organismo en cuestión, y por tanto resulta primordial tener un sistema de regulación sumamente bien controlado (Nijland et al., 2022; Romine et al., 2017).

Un grupo de genes que resultó muy interesante encontrar regulado por riboswitches antisentido fue el caso de aquellos que codifican para proteínas transportadoras, lo anterior debido a que, existen algunos transportadores que pueden ser regulados por riboswitches en sentido (Sherlock & Breaker, 2020; Xu et al., 2023), sin embargo el tipo de transportadores que se identificaron fueron en especial aquellos que pertenecen al tipo ABC o que transportan cationes relevantes para la función de diversas enzimas (Akhtar & Turner, 2022; Davidson & Chen, 2004; Porcheron et al., 2013).

En ese sentido, otro de los ejemplos a destacar es el sistema que involucra genes que regulan moléculas señalizadoras; se propone que en presencia de riboswitches antisentido en estos genes resulta de gran importancia para contar con mecanismos que funcionen como sistemas de alarma cuando existen procesos que ponen en peligro la viabilidad o las funciones básicas de la célula ya que la

existencia de un riboswitch antisentido que sense de manera específica la existencia de un metabolito involucrado en dichos procesos, es una manera de controlar la regulación de genes vitales con el importante propósito de encender el mecanismo regulatorio de una manera sumamente fina (Camilli & Bassler, 2006; Yim et al., 2007).

Otro grupo de genes sumamente interesante y que además fue el segundo ejemplo descrito para un riboswitch antisentido, es el caso de aquellos que codifican para reguladores transcripcionales (Mellin et al., 2013). Es bien conocido que regulación de este tipo de elementos genéticos debe siempre encontrarse sumamente bien controlada, y es por ello que se propone que la existencia de un riboswitch antisentido, en un punto tan crítico de una red de regulación, permite asegurar que la función de los genes río abajo en la cascada regulatoria de la misma, únicamente ocurrirá cuando las condiciones sean las pertinentes (Weirauch & Hughes, 2011).

Además, en el caso de los grupos que no nos permiten dar una razón biológica para la existencia de un riboswitch antisentido, resulta importante resaltar que en este proyecto se propone que la existencia de los riboswitches antisentido fuese previa a la existencia de aquellos elementos, debido a que como ya se ha mencionado en los puntos anteriores, resulta sumamente importante resaltar que la existencia de los riboswitches antisentido, al analizar los resultados obtenidos, puede ser explicada gracias a la necesidad de un sistema de control o regulación mayor que el que se requiere en otros puntos del metabolismo.

Como una parte adicional al proyecto (Apéndice 1), se decidió ahondar en el descubrimiento de probables errores de anotación gracias a la identificación de motivos de riboswitches en regiones anotadas como codificantes. Esta parte del trabajo surgió gracias a que durante el proceso de identificación de riboswitches antisentido, una gran parte de los genes sujetos a regulación por este tipo de elementos, se trataba de pequeños ORFs (menores a 300 pb) y que no poseían ortólogos en los organismos que se evaluaron. Al llevar a cabo una revisión bibliográfica de dicho fenómeno, se encontró que estas características constituyen evidencia importante de la existencia de errores de anotación en el genoma, además se encontraron referencias que apoyaron la utilización de la búsqueda de motivos conservados para lograr la identificación de errores de anotación (Amin et al., 2018; Lin et al., 2008; Linial, 2003).

Gracias a este método, nos fue posible identificar errores de anotación de secuencias de dos tipos: parciales, con un codón de inicio o paro de la transcripción erróneo, y totales, con ORFs cortos que no poseen ortólogos identificables; lo anterior gracias a la relación de los genes adyacentes con motivos de riboswitches. Dicha relación aporta certeza en la identificación debido a las propiedades de conservación tanto de secuencia y estructura que poseen los riboswitches, así como gracias al método de análisis utilizado.

Debido al incremento de en la secuenciación de genomas, se ha dado también un incremento en el número de errores de anotación y esto es principalmente debido al proceso que se sigue para ello. Se considera que este es un método relevante para la búsqueda y futura corrección de errores de anotación,

o bien que este sea un método de verificación en las secuencias, que pudiera utilizarse previo a la anotación de las secuencias (Devos & Valencia, 2001; Jones et al., 2007).

Finalmente, se desea hacer énfasis que una de las principales implicaciones del proyecto es que el método que seguimos pudo haber sido encontrado con anterioridad y así también la identificación de los riboswitches antisentido, y se considera que este descubrimiento puede abrir las puertas a la investigación de este elemento regulatorio, así como de otro tipo de nuevos elementos no canónicos que pudiesen cambiar el paradigma.

CONCLUSIONES

- Se identificaron 99,267 riboswitches en las regiones intergénicas del genoma de las bacterias y arqueas analizadas.
- Un total de 98,082 riboswitches (98.8%) corresponde a riboswitches en sentido.
- Los riboswitches antisentido identificados corresponden a un total de 1,185 (1.2%).
- Las familias de riboswitches cobalamina, c-di-GMP, ykoK/M-box y MOCO presentan un enriquecimiento en tener regulación antisentido.
- Se estableció una estrategia para la clasificación de los genes regulados por riboswitches antisentido.
- Se logró incrementar el número de riboswitches antisentido reportados a más de 100 veces con respecto a los datos previos a este análisis.
- La búsqueda de riboswitches en regiones codificantes en el genoma de bacterias y arqueas permite la identificación de probables errores de anotación.

PERSPECTIVAS

- Comprobación experimental de al menos uno de los riboswitches antisentido identificados.
- Análisis transcriptómico de secuencias para la búsqueda de riboswitches antisentido, cuando los datos disponibles sean óptimos para tal fin.
- Creación de un servidor para la búsqueda de riboswitches antisentido.
- Creación de un servidor para la búsqueda de errores de anotación basados en la identificación de riboswtiches.

REFERENCIAS

- Akhtar, A. A., & Turner, D. P. (2022). The role of bacterial ATP-binding cassette (ABC) transporters in pathogenesis and virulence: Therapeutic and vaccine potential. *Microbial Pathogenesis*, 171(September 2021), 105734. <https://doi.org/10.1016/j.micpath.2022.105734>
- Amin, M. R., Yurovsky, A., Chen, Y., Skiena, S., & Futcher, B. (2018). Re-annotation of 12,495 prokaryotic 16s rrna 3' ends and analysis of shine-dalgarno and anti-shine-dalgarno sequences. *PLoS ONE*, 13(8). <https://doi.org/10.1371/journal.pone.0202767>
- André, G., Even, S., Putzer, H., Burguière, P., Croux, C., Danchin, A., Martin-Verstraete, I., & Soutourina, O. (2008). S-box and T-box riboswitches and antisense RNA control a sulfur metabolic operon of *Clostridium acetobutylicum*. *Nucleic Acids Research*, 36(18), 5955–5969. <https://doi.org/10.1093/nar/gkn601>
- Antunes, D., Jorge, N. A. N., Caffarena, E. R., & Passetti, F. (2018). Using RNA sequence and structure for the prediction of riboswitch aptamer: A comprehensive review of available software and tools. *Frontiers in Genetics*, 8(231). <https://doi.org/10.3389/fgene.2017.00231>
- Barrick, J. E., & Breaker, R. R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biology*, 8(11). <https://doi.org/10.1186/gb-2007-8-11-r239>
- Batey, R. T. (2015). Riboswitches: Still a lot of undiscovered country. *Rna*, 21(4), 560–563. <https://doi.org/10.1261/rna.050765.115>
- Bervoets, I., & Charlier, D. (2019). Diversity, versatility and complexity of bacterial gene regulation mechanisms: Opportunities and drawbacks for applications in synthetic biology. *FEMS Microbiology Reviews*, 43(3), 304–339. <https://doi.org/10.1093/femsre/fuz001>
- Brantl, S. (2007). Regulatory mechanisms employed by cis-encoded antisense RNAs. *Current Opinion in Microbiology*, 10(2), 102–109. <https://doi.org/10.1016/j.mib.2007.03.012>
- Breaker, R. R. (2011). Prospects for Riboswitch Discovery and Analysis. In *Molecular Cell* (Vol. 43, Issue 6, pp. 867–879). <https://doi.org/10.1016/j.molcel.2011.08.024>
- Breaker, R. R. (2012). Riboswitches and the RNA world. *Cold Spring Harbor Perspectives in Biology*, 4(2). <https://doi.org/10.1101/cshperspect.a003566>
- Brenner, S. E. (1999). Errors in genome annotation. *Trends in Genetics*, 15(4), 132–133. [https://doi.org/10.1016/S0168-9525\(99\)01706-0](https://doi.org/10.1016/S0168-9525(99)01706-0)
- Caldara, M., Dupont, G., Leroy, F., Goldbeter, A., De Vuyst, L., & Cunin, R. (2008). Arginine biosynthesis in *Escherichia coli*: Experimental perturbation and mathematical modeling. *Journal of Biological Chemistry*, 283(10), 6347–6358. <https://doi.org/10.1074/jbc.M705884200>

- Camilli, A., & Bassler, B. L. (2006). Bacterial Small-Molecule Signaling Pathways. *Science (New York, N.Y.)*, *311*(5764), 1113–1116. <https://doi.org/doi:10.1126/science.1121357>
- Chang, T. H., Huang, H. Da, Wu, L. C., Yeh, C. T., Liu, B. J., & Horng, J. T. (2009). Computational identification of riboswitches based on RNA conserved functional sequences and conformations. *RNA*, *15*(7), 1426–1430. <https://doi.org/10.1261/rna.1623809>
- Chen, Y., Chai, Y., Guo, J. hua, & Losick, R. (2012). Evidence for cyclic Di-GMP-mediated signaling in *Bacillus subtilis*. *Journal of Bacteriology*, *194*(18), 5080–5090. <https://doi.org/10.1128/JB.01092-12>
- Courtney, C., & Chatterjee, A. (2014). cis-Antisense RNA and Transcriptional Interference: Coupled Layers of Gene Regulation. *Journal of Gene Therapy*, *1*(1), 1–9. <https://doi.org/10.13188/2381-3326.1000004>
- Cui, X., Lu, Z., Wang, S., Jing-Yan Wang, J., & Gao, X. (2016). CMsearch: Simultaneous exploration of protein sequence space and structure space improves not only protein homology detection but also protein structure prediction. *Bioinformatics*, *32*(12), i332–i340. <https://doi.org/10.1093/bioinformatics/btw271>
- da Cruz Nizer, W. S., Inkovskiy, V., Versey, Z., Stempel, N., Cassol, E., & Overhage, J. (2021). Oxidative stress response in *Pseudomonas aeruginosa*. *Pathogens*, *10*(9), 1–32. <https://doi.org/10.3390/pathogens10091187>
- Dambach, M., Sandoval, M., Updegrove, T. B., Anantharaman, V., Aravind, L., Waters, L. S., & Storz, G. (2015). The Ubiquitous yybP-ykoY Riboswitch Is a Manganese-Responsive Regulatory Element. *Molecular Cell*, *57*(6), 1099–1109. <https://doi.org/10.1016/j.molcel.2015.01.035>
- Davidson, A. L., & Chen, J. (2004). ATP-binding cassette transporters in bacteria. *Annual Review of Biochemistry*, *73*, 241–268. <https://doi.org/10.1146/annurev.biochem.73.011303.073626>
- Devos, D., & Valencia, A. (2001). Intrinsic errors in genome annotation. *Trends in Genetics*, *17*(8), 429–431. [https://doi.org/10.1016/S0168-9525\(01\)02348-4](https://doi.org/10.1016/S0168-9525(01)02348-4)
- Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, *7*(10). <https://doi.org/10.1371/journal.pcbi.1002195>
- Edwards, A. L., & Batey, R. T. (2010). Riboswitches: A Common RNA Regulatory Element. *Nature Education*, *3*(9)(9).
- Erdem, M., Cicek, M., & Erson-Bensan, A. E. (2023). Versatile RNA: overlooked gems of the transcriptome. *FEBS Journal*, 1–9. <https://doi.org/10.1111/febs.16742>
- Garst, A. D., Edwards, A. L., & Batey, R. T. (2011). Riboswitches: Structures and mechanisms. In *Cold Spring Harbor Perspectives in Biology* (Vol. 3, Issue 6, pp. 1–13). Cold Spring Harbor Laboratory Press. <https://doi.org/10.1101/cshperspect.a003533>

- Georg, J., & Hess, W. R. (2011). cis -Antisense RNA, Another Level of Gene Regulation in Bacteria . *Microbiology and Molecular Biology Reviews*, 75(2), 286–300. <https://doi.org/10.1128/mubr.00032-10>
- Grundy, F. J., & Henkin, T. M. (1998). The S box regulon: A new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Molecular Microbiology*, 30(4), 737–749. <https://doi.org/10.1046/j.1365-2958.1998.01105.x>
- Hall, S., McDermott, C., Anoopkumar-Dukie, S., McFarland, A. J., Forbes, A., Perkins, A. V., Davey, A. K., Chess-Williams, R., Kiefel, M. J., Arora, D., & Grant, G. D. (2016). Cellular effects of pyocyanin, a secreted virulence factor of *Pseudomonas aeruginosa*. *Toxins*, 8(8), 1–14. <https://doi.org/10.3390/toxins8080236>
- Haller, A., Soulière, M. F., & Micura, R. (2011). The dynamic nature of RNA as key to understanding riboswitch mechanisms. In *Accounts of Chemical Research* (Vol. 44, Issue 12, pp. 1339–1348). <https://doi.org/10.1021/ar200035g>
- He, W., Zhang, X., Zhang, J., Jia, X., Zhang, J., Sun, W., Jiang, H., Chen, D., & Murchie, A. I. H. (2013). Riboswitch control of induction of aminoglycoside resistance acetyl and adenylyl-transferases. *RNA Biology*, 10(8), 1266–1273. <https://doi.org/10.4161/rna.25757>
- Jenal, U., Reinders, A., & Lori, C. (2017). Cyclic di-GMP: Second messenger extraordinaire. *Nature Reviews Microbiology*, 15(5), 271–284. <https://doi.org/10.1038/nrmicro.2016.190>
- Jones, C. E., Brown, A. L., & Baumann, U. (2007). Estimating the annotation error rate of curated GO database sequence annotations. *BMC Bioinformatics*, 8. <https://doi.org/10.1186/1471-2105-8-170>
- Kalvari, I., Nawrocki, E. P., Ontiveros-Palacios, N., Argasinska, J., Lamkiewicz, K., Marz, M., Griffiths-Jones, S., Toffano-Nioche, C., Gautheret, D., Weinberg, Z., Rivas, E., Eddy, S. R., Finn, R. D., Bateman, A., & Petrov, A. I. (2021). Rfam 14: Expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Research*, 49(D1), D192–D200. <https://doi.org/10.1093/nar/gkaa1047>
- Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., & Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, 51(D1), D587–D592. <https://doi.org/10.1093/nar/gkac963>
- Kavita, K., & Breaker, R. R. (2023). Discovering riboswitches: the past and the future. *Trends in Biochemical Sciences*, 48(2), 119–141. <https://doi.org/10.1016/j.tibs.2022.08.009>
- Kearns, D. B., Chu, F., Branda, S. S., Kolter, R., & Losick, R. (2005). A master regulator for biofilm formation by *Bacillus subtilis*. *Molecular Microbiology*, 55(3), 739–749. <https://doi.org/10.1111/j.1365-2958.2004.04440.x>
- Kehres, D. G., & Maguire, M. E. (2003). Emerging themes in manganese transport, biochemistry and pathogenesis in bacteria. *FEMS Microbiology Reviews*,

27(2–3), 263–290. [https://doi.org/10.1016/S0168-6445\(03\)00052-4](https://doi.org/10.1016/S0168-6445(03)00052-4)

- Köpke, M., Held, C., Hujer, S., Liesegang, H., Wiezer, A., Wollherr, A., Ehrenreich, A., Liebl, W., Gottschalk, G., & Dürre, P. (2010). Clostridium ljungdahlii represents a microbial production platform based on syngas. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29), 13087–13092. <https://doi.org/10.1073/pnas.1004716107>
- Lebreton, A., & Cossart, P. (2017). RNA- and protein-mediated control of Listeria monocytogenes virulence gene expression. *RNA Biology*, 14(5), 460–470. <https://doi.org/10.1080/15476286.2016.1189069>
- Lee, C., & Park, C. (2017). Bacterial responses to glyoxal and methylglyoxal: Reactive electrophilic species. *International Journal of Molecular Sciences*, 18(1). <https://doi.org/10.3390/ijms18010169>
- Lin, Y. H., Chang, B. C. H., Chiang, P. W., & Tang, S. L. (2008). Questionable 16S ribosomal RNA gene annotations are frequent in completed microbial genomes. *Gene*, 416(1–2), 44–47. <https://doi.org/10.1016/j.gene.2008.02.023>
- Linial, M. (2003). How incorrect annotations evolve - The case of short ORFs. In *Trends in Biotechnology* (Vol. 21, Issue 7, pp. 298–300). Elsevier Ltd. [https://doi.org/10.1016/S0167-7799\(03\)00139-2](https://doi.org/10.1016/S0167-7799(03)00139-2)
- Lochab, V., Jones, T. H., Dusane, D. H., Peters, C. W., Stoodley, P., Wozniak, D. J., Subramaniam, V. V., & Prakash, S. (2020). Ultrastructure imaging of Pseudomonas aeruginosa lawn biofilms and eradication of the tobramycin-resistant variants under in vitro electroceutical treatment. *Scientific Reports*, 10(1), 1–12. <https://doi.org/10.1038/s41598-020-66823-y>
- Martinez-Guerrero, C. E., Ciria, R., Abreu-Goodger, C., Moreno-Hagelsieb, G., & Merino, E. (2008). GeConT 2: gene context analysis for orthologous proteins, conserved domains and metabolic pathways. *Nucleic Acids Research*, 36(Web Server issue), 176–180. <https://doi.org/10.1093/nar/gkn330>
- Mccown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E., & Breaker, R. R. (2017). Riboswitch diversity and distribution. *RNA*, 23(7), 995–1011. <https://doi.org/10.1261/rna.061234>
- Mehdizadeh Aghdam, E., Hejazi, M. S., & Barzegar, A. (2016). Riboswitches: From living biosensors to novel targets of antibiotics. *Gene*, 591(2), 244–259. <https://doi.org/10.1016/j.gene.2016.07.035>
- Mellin, J. R., Tiensuu, T., Bécavin, C., Gouin, E., Johansson, J., & Cossart, P. (2013). A riboswitch-regulated antisense RNA in Listeria monocytogenes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(32), 13132–13137. <https://doi.org/10.1073/pnas.1304795110>
- Miranda-Ríos, J., Navarro, M., & Soberón, M. (2001). A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 98(17), 9736–9741. <https://doi.org/10.1073/pnas.161168098>

- Miranda-Ríos, Juan, Morera, C., Taboada, H., Dávalos, A., Encarnación, S., Mora, J., & Soberón, M. (1997). Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiotic terminal oxidase cbb3 in *Rhizobium etli*. *Journal of Bacteriology*, 179(22), 6887–6893. <https://doi.org/10.1128/jb.179.22.6887-6893.1997>
- Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., Perumov, D. A., & Nudler, E. (2002). Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria. *Cell*, 111(5), 747–756. [https://doi.org/10.1016/S0092-8674\(02\)01134-0](https://doi.org/10.1016/S0092-8674(02)01134-0)
- Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., & Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. *Chemistry and Biology*, 9(9), 1043–1049. [https://doi.org/10.1016/S1074-5521\(02\)00224-7](https://doi.org/10.1016/S1074-5521(02)00224-7)
- Nelson, J., Atilho, R., Sherlock, M., Stockbridge, R., & Breaker, R. (2017). Metabolism of Free Guanidine in Bacteria is Regulated by a Widespread Riboswitch Class. *Mol Cell*, 65(2), 220–230. <https://doi.org/10.1016/j.molcel.2016.11.019>
- Nijland, M., Martínez Felices, J. M., Slotboom, D. J., & Thangaratnarajah, C. (2022). Membrane transport of cobalamin. In *Vitamins and Hormones* (1st ed., Vol. 119). Elsevier Inc. <https://doi.org/10.1016/bs.vh.2022.01.008>
- Nou, X., & Kadner, R. J. (2000). Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proceedings of the National Academy of Sciences of the United States of America*, 97(13), 7190–7195. <https://doi.org/10.1073/pnas.130013897>
- Pan, X., Dong, Y., Fan, Z., Liu, C., Xia, B., Shi, J., Bai, F., Jin, Y., Cheng, Z., Jin, S., & Wu, W. (2017). In vivo host environment alters *Pseudomonas aeruginosa* susceptibility to aminoglycoside antibiotics. *Frontiers in Cellular and Infection Microbiology*, 7(MAR), 1–12. <https://doi.org/10.3389/fcimb.2017.00083>
- Pavlova, N., Kaloudas, D., & Penchovsky, R. (2019). Riboswitch distribution, structure, and function in bacteria. In *Gene* (Vol. 708, pp. 38–48). Elsevier B.V. <https://doi.org/10.1016/j.gene.2019.05.036>
- Porcheron, G., Garénaux, A., Proulx, J., Sabri, M., & Dozois, C. M. (2013). Iron, copper, zinc, and manganese transport and regulation in pathogenic Enterobacteria: Correlations between strains, site of infection and the relative importance of the different metal transport systems for virulence. *Frontiers in Cellular and Infection Microbiology*, 3(DEC), 1–24. <https://doi.org/10.3389/fcimb.2013.00090>
- Rahman, F. Bin, Sarkar, B., Moni, R., & Rahman, M. S. (2021). Molecular genetics of surfactin and its effects on different sub-populations of *Bacillus subtilis*. *Biotechnology Reports*, 32, e00686. <https://doi.org/10.1016/j.btre.2021.e00686>
- Reiss, C., Xiong, Y., & Strobel, S. (2017). Structural basis for ligand binding to the guanidine-I riboswitch. *Structure*, 25(1), 195–202. <https://doi.org/10.1016/j.str.2016.11.020>

- Reyes-Darias, J. A., & Krell, T. (2017). Riboswitches as Potential Targets for the Development of Anti-Biofilm Drugs. *Current Topics in Medicinal Chemistry*, 17(17), 1945–1953. <https://doi.org/10.2174/1568026617666170407163517>
- Richer, E., Courville, P., Bergevin, I., & Cellier, M. F. M. (2003). Horizontal Gene Transfer of “Prototype” Nramp in Bacteria. *Journal of Molecular Evolution*, 57(4), 363–376. <https://doi.org/10.1007/s00239-003-2472-z>
- Rodionov, D. A., Vitreschak, A. G., Mironov, A. A., & Gelfand, M. S. (2004). Comparative genomics of the methionine metabolism in Gram-positive bacteria: A variety of regulatory systems. *Nucleic Acids Research*, 32(11), 3340–3353. <https://doi.org/10.1093/nar/gkh659>
- Romine, M. F., Rodionov, D. A., Maezato, Y., Anderson, L. N., Nandhikonda, P., Rodionova, I. A., Carre, A., Li, X., Xu, C., Clauss, T. R. W., Kim, Y. M., Metz, T. O., & Wright, A. T. (2017). Elucidation of roles for vitamin B12 in regulation of folate, ubiquinone, and methionine metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 114(7), E1205–E1214. <https://doi.org/10.1073/pnas.1612360114>
- Serrano-Gutiérrez, M., & Merino, E. (2023). Antisense-acting riboswitches: A poorly characterized yet important model of transcriptional regulation in prokaryotic organisms. *Plos One*, 18(2), e0281744. <https://doi.org/10.1371/journal.pone.0281744>
- Sesto, N. et al. (2013). The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nature Reviews Microbiology*, 11(2), 75–82.
- Sherlock, M. E., & Breaker, R. R. (2020). Former orphan riboswitches reveal unexplored areas of bacterial metabolism, signaling, and gene control processes. *Rna*, 26(6), 675–693. <https://doi.org/10.1261/rna.074997.120>
- Sherwood, A. V., & Henkin, T. M. (2016). Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annual Review of Microbiology*, 70(1), 361–374. <https://doi.org/10.1146/annurev-micro-091014-104306>
- Sinn, M., Hauth, F., Lenkeit, F., Weinberg, Z., & Hartig, J. S. (2021). Widespread bacterial utilization of guanidine as nitrogen source. *Molecular Microbiology*, 116(1), 200–210. <https://doi.org/10.1111/mmi.14702>
- Sinumvayo, J. P., Zhao, C., & Tuyishime, P. (2018). Recent advances and future trends of riboswitches: attractive regulatory tools. *World Journal of Microbiology and Biotechnology*, 34(11), 1–15. <https://doi.org/10.1007/s11274-018-2554-0>
- Slobodkina, G. B., Panteleeva, A. N., Sokolova, T. G., Bonch-Osmolovskaya, E. A., & Slobodkin, A. I. (2012). Carboxydocella manganica sp. nov., a thermophilic, dissimilatory Mn(IV)- and Fe(III)-reducing bacterium from a Kamchatka hot spring. *International Journal of Systematic and Evolutionary Microbiology*, 62(4), 890–894. <https://doi.org/10.1099/ijs.0.027623-0>
- Sornchuer, P., Saninjuk, K., Prathaphan, P., Tiengtip, R., & Wattanaphansak, S.

- (2022). Antimicrobial Susceptibility Profile and Whole-Genome Analysis of a Strong Biofilm-Forming *Bacillus* Sp. B87 Strain Isolated from Food. *Microorganisms*, 10(2). <https://doi.org/10.3390/microorganisms10020252>
- Starmer, J., Stomp, A., Vouk, M., & Bitzer, D. (2005). Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Computational Biology*, preprint(2006), e57. <https://doi.org/10.1371/journal.pcbi.0020057.eor>
- Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., & Breaker, R. R. (2008). Riboswitches in eubacteria sense the second messenger cyclic Di-GMP. *Science*, 321(5887), 411–413. <https://doi.org/10.1126/science.1159519>
- Sukdeo, N., Clugston, S. L., Daub, E., & Honek, J. F. (2004). Distinct classes of glyoxalase I: Metal specificity of the *Yersinia pestis*, *Pseudomonas aeruginosa* and *Neisseria meningitidis* enzymes. *Biochemical Journal*, 384(1), 111–117. <https://doi.org/10.1042/BJ20041006>
- Sukdeo, N., & Honek, J. F. (2007). *Pseudomonas aeruginosa* contains multiple glyoxalase I-encoding genes from both metal activation classes. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1774(6), 756–763. <https://doi.org/10.1016/j.bbapap.2007.04.005>
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1), D607–D613. <https://doi.org/10.1093/nar/gky1131>
- Taboada, B., Verde, C., & Merino, E. (2010). High accuracy operon prediction method based on STRING database scores. *Nucleic Acids Research*, 38(12). <https://doi.org/10.1093/nar/gkq254>
- Tamayo, R. (2019). Cyclic diguanylate riboswitches control bacterial pathogenesis mechanisms. *PLoS Pathogens*, 15(2), 1–7. <https://doi.org/10.1371/journal.ppat.1007529>
- Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., Kiryutin, B., Galperin, M. Y., Fedorova, N. D., & Koonin, E. V. (2001). The COG database: New developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Research*, 29(1), 22–28. <https://doi.org/10.1093/nar/29.1.22>
- Toshchakov, S. V., Lebedinsky, A. V., Sokolova, T. G., Zavarzina, D. G., Korzhenkov, A. A., Teplyuk, A. V., Chistyakova, N. I., Rusakov, V. S., Bonch-Osmolovskaya, E. A., Kublanov, I. V., & Gavrilov, S. N. (2018). Genomic insights into energy metabolism of *Carboxydocella therrautotrophica* coupling hydrogenogenic co oxidation with the reduction of Fe(III) minerals. In *Frontiers in Microbiology* (Vol. 9, Issue AUG). <https://doi.org/10.3389/fmicb.2018.01759>
- Vitreschak, A. G., Rodionov, D. A., Mironov, A. A., & Gelfand, M. S. (2002).

- Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Research*, 30(14), 3141–3151. <https://doi.org/10.1093/nar/gkf433>
- Vitreschak, A. G., Rodionov, D. A., Mironov, A. A., & Gelfand, M. S. (2003). Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element. *Rna*, 9(9), 1084–1097. <https://doi.org/10.1261/rna.5710303>
- Voichek, M., Maaß, S., Kroniger, T., Becher, D., & Sorek, R. (2020). Peptide-based quorum sensing systems in *Paenibacillus polymyxa*. *Life Science Alliance*, 3(10), 1–12. <https://doi.org/10.26508/LSA.202000847>
- Waters, L. S. (2020). Bacterial manganese sensing and homeostasis. *Current Opinion in Chemical Biology*, 55(Box 1), 96–102. <https://doi.org/10.1016/j.cbpa.2020.01.003>
- Waters, L. S., Sandoval, M., & Storz, G. (2011). The *Escherichia coli* MntR miniregulon includes genes encoding a small protein and an efflux pump required for manganese homeostasis. *Journal of Bacteriology*, 193(21), 5887–5897. <https://doi.org/10.1128/JB.05872-11>
- Waters, L., & Storz, G. (2009). Regulatory RNAs in Bacteria. *Cell*, 136(4), 615–628. <https://doi.org/10.1016/j.cell.2009.01.043>
- Weinmaier, T., Platzer, A., Frank, J., Hellinger, H. J., Tischler, P., & Rattei, T. (2016). ConsPred: A rule-based (re-)annotation framework for prokaryotic genomes. *Bioinformatics*, 32(21), 3327–3329. <https://doi.org/10.1093/bioinformatics/btw393>
- Weirauch, M. T., & Hughes, T. R. (2011). A Handbook of Transcription Factors. In T. R. Hughes (Ed.), *Sub-Cellular Biochemistry* (Vol. 52). Springer Netherlands. <https://doi.org/10.1007/978-90-481-9069-0>
- Weiss, C. A., Hoberg, J. A., Liu, K., Tu, B. P., & Winkler, W. C. (2019). *Single-Cell Microscopy Reveals That Levels of Cyclic di-GMP Vary among Bacillus subtilis Subpopulations*. *May*, 1–17.
- Winkler, W. C. (2005). Riboswitches and the role of noncoding RNAs in bacterial metabolic control. *Current Opinion in Chemical Biology*, 9(6), 594–602. <https://doi.org/10.1016/j.cbpa.2005.09.016>
- Winkler, W. C., & Breaker, R. R. (2005). REGULATION OF BACTERIAL GENE EXPRESSION BY RIBOSWITCHES. *Annual Review of Microbiology*, 59(1), 487–517. <https://doi.org/10.1146/annurev.micro.59.030804.121336>
- Winkler, W. C., Cohen-Chalamish, S., & Breaker, R. R. (2002). An mRNA structure that controls gene expression by binding FMN. *Proceedings of the National Academy of Sciences of the United States of America*, 99(25), 15908–15913. <https://doi.org/10.1073/pnas.212628899>
- Winkler, W., Nahvi, A., & Breaker, R. R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*,

419(6910), 952–956. <https://doi.org/10.1038/nature01145>

- Wongsaroj, L., Saninjuk, K., Romsang, A., Duang-nkern, J., Trinachartvanit, W., Vattanaviboon, P., & Mongkolsuk, S. (2018). *Pseudomonas aeruginosa* glutathione biosynthesis genes play multiple roles in stress protection, bacterial virulence and biofilm formation. *PLoS ONE*, 13(10), 1–18. <https://doi.org/10.1371/journal.pone.0205815>
- Xu, J., Hou, J., Ding, M., Wang, Z., & Chen, T. (2023). Riboswitches, from cognition to transformation. *Synthetic and Systems Biotechnology*, 8(3), 357–370. <https://doi.org/10.1016/j.synbio.2023.05.008>
- Yi, J., Huang, H., Liang, J., Wang, R., Liu, Z., Li, F., & Wang, S. (2021). A Heterodimeric Reduced-Ferredoxin-Dependent Methylenetetrahydrofolate Reductase from Syngas-Fermenting *Clostridium ljungdahlii*. *Microbiology Spectrum*, 9(2), e00958-21.
- Yim, G., Wang, H. H., & Davies, J. (2007). Antibiotics as signalling molecules. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1483), 1195–1200. <https://doi.org/10.1098/rstb.2007.2044>
- Zeinert, R., Martinez, E., Schmitz, J., Senn, K., Usman, B., Anantharaman, V., Aravind, L., & Waters, L. S. (2018). Structure–function analysis of manganese exporter proteins across bacteria. *Journal of Biological Chemistry*, 293(15), 5715–5730. <https://doi.org/10.1074/jbc.M117.790717>

RESEARCH ARTICLE

Antisense-acting riboswitches: A poorly characterized yet important model of transcriptional regulation in prokaryotic organisms

Mariela Serrano-Gutiérrez, Enrique Merino*

Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

* enrique.merino@ibt.unam.mx

OPEN ACCESS

Citation: Serrano-Gutiérrez M, Merino E (2023) Antisense-acting riboswitches: A poorly characterized yet important model of transcriptional regulation in prokaryotic organisms. PLoS ONE 18(2): e0281744. <https://doi.org/10.1371/journal.pone.0281744>

Editor: Michael F. Minnick, University of Montana, UNITED STATES

Received: November 7, 2022

Accepted: January 30, 2023

Published: February 21, 2023

Copyright: © 2023 Serrano-Gutiérrez, Merino. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: Our work was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT) (Grant 2040) and DGAPA-Universidad Nacional Autónoma de México (UNAM) PAPIIT (Grant IN202120); both grants were awarded to E.M. M.S. is a Ph.D. student enrolled in the Programa de Doctorado en Ciencias Bioquímicas, Universidad

Abstract

Riboswitches are RNA elements involved in regulating genes that participate in the biosynthesis or transport of essential metabolites. They are characterized by their ability to recognize their target molecules with high affinity and specificity. Riboswitches are commonly cotranscribed with their target genes and are located at the 5' end of their transcriptional units. To date, only two exceptional cases of riboswitches being situated at the 3' end and transcribing in the antisense direction of their regulated genes have been described. The first case involves a SAM riboswitch located at the 3' end of the *ubiG-mccB-mccA* operon in *Clostridium acetobutylicum* involved in converting methionine to cysteine. The second case concerns a Cobalamin riboswitch in *Listeria monocytogenes* that regulates the transcription factor PocR related to this organism's pathogenic process. In almost a decade since the first descriptions of antisense-acting riboswitches, no new examples have been described. In this work, we performed a computational analysis to identify new examples of antisense-acting riboswitches. We found 292 cases in which, according to the available information, we infer that the expected regulation of the riboswitch is consistent with the signaling molecule it senses and the metabolic function of the regulated gene. The metabolic implications of this novel type of regulation are thoroughly discussed.

Introduction

Riboswitches are RNA elements involved in the regulation of certain types of genes that participate in the biosynthesis or transport of essential metabolites, such as vitamins, cofactors, nucleotides, or amino acids, and are characterized by their ability to recognize their corresponding target molecules with high affinity and specificity without the participation of protein factors [1–5]. This high specificity of riboswitches to their ligands imposes three-dimensional restrictions that result in significant conservation in their primary and secondary structures, even in riboswitches of phylogenetically distant organisms. This conservation is essential to the *in silico* identification of riboswitches within currently sequenced genomes [6–9].

Nacional Autónoma de México (UNAM) and recipient of a CONACYT fellowship (702475). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Structurally and functionally, riboswitches are commonly composed of two domains: an aptamer domain, which is part of the riboswitch that selectively recognizes its ligand, and an expression or regulatory domain, which undergoes allosteric modifications due to metabolite binding [7–9]. The level of regulation at which riboswitches act depends on the type of regulatory elements of their expression domain, which mainly include i) premature termination of transcription, mediated by transcriptional attenuators; ii) inhibition of translation initiation by using the formation of mRNA secondary structures in the 5' UTR that sequester the Shine-Dalgarno (SD) sequence or by having the mRNA fragment that results from the premature transcription termination of a riboswitch act *in trans* as a small-antisense RNA that recognizes the SD of the target gene by complementary base pairing; iii) through the reduction in mRNA stability of the regulated gene when the expression platform is folded into an allosteric ribozyme that induces its self-cleavage; and iv) by splicing control, a mechanism that is present only in eukaryotic organisms. In addition to these regulatory mechanisms, riboswitches may block transcription elongation when they are located at the 3' end of their target genes and are transcribed in an antisense manner.

Antisense RNA-mediated gene regulation has been found to play an essential role in determining the outcome of different biological processes, such as plasmid segregational stability and the control of the plasmid copy number, transposition efficiency, regulation of the lysis/lysogeny phage cycles, and modulation of the toxin-antitoxin system [10]. More recently, the term “excludon” was coined to define a region of DNA whose transcription extends beyond the gene it codes for and continues in an antisense manner within the region of its corresponding adjacent gene, commonly with an opposing function, inhibiting its expression [11]. In principle, three different molecular mechanisms might be involved in the cis-antisense RNA regulatory outcome: a) RNA polymerase collision, in which the RNA polymerases from complementary DNA strands collide with each other during convergent transcription; b) duplex antisense RNA formation (asRNA/mRNA), which are subject to degradation by RNase III activity (partially synthesized double-stranded RNA) and inhibition of translation initiation when the SD sequences are sequestered by the antisense mRNA originated from the antisense-acting riboswitch promoter (totally synthesized double-stranded RNA); and c) accumulation of DNA supercoiling, which arises from the positive supercoils in the double DNA strand that are generated as the RNA polymerases co-transcribing in opposite directions move through the same DNA locus [12–14] (Fig 1).

To date, only two examples of antisense-acting riboswitches have been described. The first case, reported in 2008, involves a S-adenosylmethionine (SAM) riboswitch located at the 3' end that transcribes in the antisense direction of the *ubiG-mccB-mccA* operon in *Clostridium acetobutylicum*. This operon codes for putatively dependent methyltransferase enzymes involved in converting methionine to cysteine. This riboswitch, also known as the S-box, is found in Firmicutes and other Gram-positive bacteria and regulates the expression of genes involved in the biosynthesis, transport, and recycling of methionine [15]. SAM riboswitches commonly regulate their target genes at the level of transcription elongation in response to the intracellular availability of methionine, or its derivatives, through transcriptional attenuators located on their expression platforms. When the intracellular concentration of SAM is high, the regulatory domain is folded into secondary structures that include a transcription terminator, favoring the premature transcription of their regulated genes. Conversely, at low intracellular concentrations of SAM, the formation of a transcriptional antiterminator is induced, and the regulated gene transcription is completed. In the case of the *C. acetobutylicum ubiG-mccB-mccA* operon, an opposite regulatory outcome for the SAM antisense-acting riboswitch was described. When the intracellular concentration of SAM in *C. acetobutylicum* is low, the formation of a transcriptional antiterminator in the riboswitch expression platform is favored in

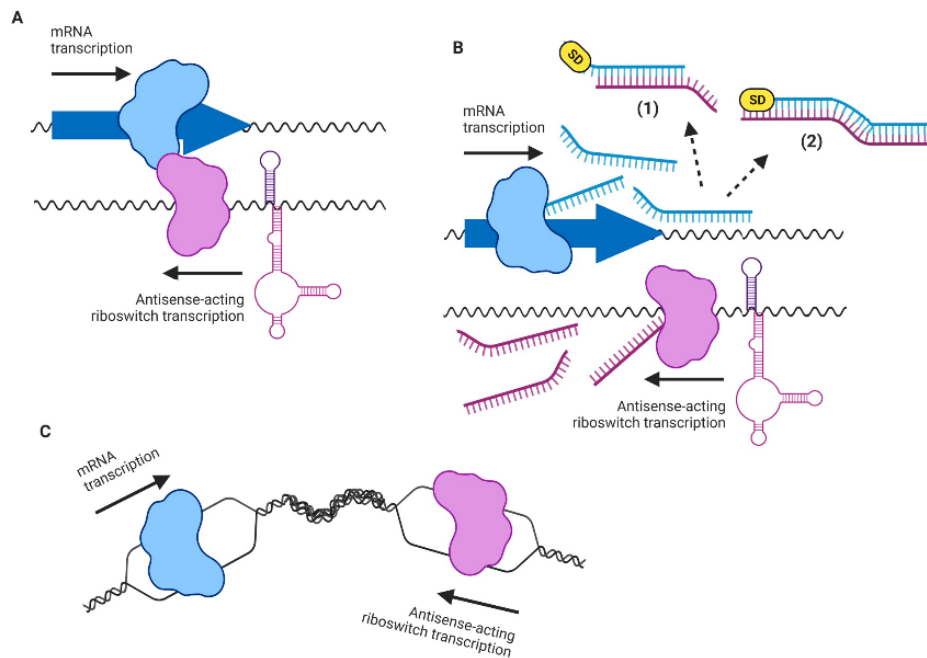


Fig 1. Models of the putative regulatory mechanisms of RNA polymerases transcribing in opposite directions. (A) RNA polymerase collision; (B) asRNA/mRNA duplex formation: (1) partially synthesized double-stranded RNA, (2) totally synthesized double-stranded RNA; (C) DNA supercoil accumulation. Representative elements are colored as follows: mRNA transcription complex, blue; asRNA transcription complex, pink; SD sequence, yellow. Created with BioRender.com.

<https://doi.org/10.1371/journal.pone.0281744.g001>

such a way that the RNA polymerase that transcribes the riboswitch continues toward the interior of the operon in the antisense direction and collides with the RNA polymerase transcribing in the sense direction, resulting in incomplete transcription of the operon and the consequent decrease in the synthesis of its corresponding protein products. In addition, the inhibitory outcome of the transcription coming from the SAM antisense-acting riboswitch might also be generated by the formation of the duplex asRNA/mRNA or the accumulation of DNA supercoiling within the *ubiG-mccB-mccA* locus when the genes are simultaneously transcribed in opposite directions. In contrast, when the intracellular SAM concentration is high, a Rho-independent transcription terminator is formed in the riboswitch's expression domain, eliminating the interference of antisense-transcribing RNA polymerases. This regulatory use of the SAM riboswitch that transcribes in the opposite direction guarantees that the synthesis of the *ubiG-mccB-mccA* operon occurs only when the intracellular SAM concentration is high [16].

The second reported case of an antisense-acting riboswitch is that of the cobalamin or B₁₂ riboswitch of *Listeria monocytogenes* that is located at the 3' end of the *pocR* gene of this organism. PocR is a transcriptional regulator of the genes coding for the enzymes involved in the metabolism of propanediol and thus plays a role in the pathogenic activities of this organism. Since these enzymes require vitamin B₁₂ as a cofactor to be active when the intracellular B₁₂

concentration is low, the cobalamin antisense-acting riboswitch represses *PocR* synthesis at the transcription level, due to RNA polymerase collision or the accumulation of DNA supercoiling within the *pocR* gene; at the mRNA stability level, by the formation of an asRNA/mRNA duplex which arises from the transcription of the *pocR* gene in the sense and antisense directions; and also interferes with *PocR* translation initiation. At a high concentration of vitamin B₁₂, a Rho-independent transcriptional terminator is formed in the regulatory domain of the SAM antisense-acting riboswitch, preventing the inhibition of *pocR* expression [17].

Here, we present the results of our comprehensive search for antisense-acting riboswitches in bacterial and archaeal representative genomes. Our findings show that regulation by antisense-acting riboswitches is much more widespread than initially reported, increasing the number from the two cases previously reported in the past decade to 292 biologically significant instances in which the regulatory outcome of antisense-acting riboswitches seems consistent with the metabolite recognized by the riboswitch and the metabolic function of the regulated gene product. Additionally, the genes we identified were classified according to their biological processes into ten categories. A representative example of each of these antisense-acting riboswitch groups is presented, and the metabolic implications of their regulation are discussed in depth.

Materials and methods

Identification of riboswitches in genome sequences

To search for riboswitches in the set of 5' and 3' intergenic genome sequences, we used the CMsearch software from the Infernal package (v.1.1.3) [18] and the covariance models of the 50 riboswitch classes described in the Rfam database [19]. Since different related covariance models could recognize the same RNA sequence, we used an ad hoc Perl program to select the most suitable model for each identified riboswitch based on their corresponding bit scores reported by the CMsearch program (i.e., *c-di-GMP-I*, *c-di-GMP-II*, *c-di-GMP-I-GGC*, *c-di-GMP-II-GAG*, *c-di-GMP-I-UAU*, and *c-di-GMP-II-GCG* covariance models), which tend to recognize the same set of RNA sequences. The transcription orientations of the riboswitches and their neighbor target genes were defined considering the genomic annotations reported in the KEGG database [20]. We only considered as antisense-acting riboswitches those elements located at no more than 500 nucleotides from their regulatory target genes.

Selection of representative organism genome sequences

Genome sequences were retrieved from the KEGG database 2022 (<https://www.genome.jp/kegg/>) with 6,892 bacterial and 380 archaeal organisms. We selected one genome sequence per species to avoid redundant genome sequences, considering those with the highest number of open reading frames. After this selection, the number of genome sequences was 5,086 and 332 for bacterial and archaeal genomes, respectively.

Statistical analysis of the results

The statistical analysis of the putative antisense-acting riboswitches identified in our study was performed using RStudio software (Version 1.2.5033).

Analysis of the genome context

The genomic context of the regulated genes and their corresponding antisense-acting riboswitches were analyzed using our GeConT webserver (<http://biocomputo.ibt.unam.mx:8080/GeConT/index.jsp>) [21].

Analysis of the metabolic pathways

The metabolic relationships of the enzymes encoded by genes regulated by antisense-acting riboswitches and the metabolites used as riboswitch signals were based on the metabolic pathways of the KEGG pathway database [20].

Analysis of the gene functions

The function of the regulated genes was analyzed considering the gene descriptions reported in the KEGG database and the description of their corresponding Cluster of Orthologous Genes (COGs). COGs assignments were determined based on a Hidden Markov Models (HMMs) search using the *hmmsearch* program [22] and a set of previously constructed HMM models that represent each of the 4,873 existing COGs [23,24].

Results and discussion

Within the repertoire of molecular elements that organisms possess to regulate gene expression, riboswitches undoubtedly have a prominent position since they are RNA elements that control the transcription or translation of their target genes in response to the highly specific recognition of their cognate molecules in the absence of protein components, an ability that was thought to be unique to regulatory proteins. To date, 50 classes of riboswitches have been described, according to their covariance models, which recognize small ligands such as metabolites, coenzymes, nucleotides, metal ions, amino acids, and even tRNA molecules, among others. In general, the nature of the genes regulated by riboswitches is related to the biosynthesis or transport of the molecules they recognize. Consequently, the presence of these molecules commonly results in the folding of riboswitch regulatory platforms into secondary structures in the 5' untranslated region of the mRNA which can promote a premature termination of transcription or inhibit the translation initiation.

To date, two examples of riboswitches have been described whose regulatory activity is contrary to that previously expected. That is, the expression of their corresponding target genes only occurs when the intracellular concentration of their signaling molecules is high. In the first case, a SAM riboswitch located at the 3' end and transcribing in antisense direction to the *ubiG-mccB-mccA* operon in *C. acetobutylicum* guarantees that the synthesis of their protein products occurs only when the intracellular SAM concentration is high. This response is consistent with the fact that the enzymes encoded by the operon above mentioned require this cofactor to be active. Therefore, the expression of these enzymes in metabolic conditions where the cofactor is limited results in futility [16]. The second case involves the regulation of the *L. monocytogenes* PocR transcription factor by a Cobalamin (B₁₂) riboswitch located at the 3' end and transcribing in opposite direction to that of the *pocR* gene. In such case, the inhibition of PocR synthesis in the absence of B₁₂ can be understood by the fact that the enzymes encoded by the target genes of this transcriptional activator are B₁₂-dependent, and the synthesis of cofactor-dependent enzymes is a result of the scarce use in absence of their corresponding cofactors [17]. In the two previous cases, where the riboswitches have inverse regulatory responses to the canonical synthesis/inhibition responses observed in their counterparts in the absence/presence of their target metabolites, two common features can be observed: they are located at the 3' untranslated region of their mRNAs and their transcription is in opposite direction to that of their target gene. Based on the latter characteristic, we refer to this kind of riboswitch as antisense-acting riboswitches.

Identification of riboswitches in prokaryotic genomic sequences

We selected a set of genomic sequences from the KEGG 2022 database from 6,892 bacteria and 380 archaea at the species phylogenetic level (see [Materials and Methods](#) section). We used the 50 covariance models defined in the Rfam database [25] and the CMsearch program [18] to search for riboswitch elements in the 5' and 3' intergenic sequences of the 7,272 prokaryotic organisms and found a total of 99,267 riboswitches (see [Materials and Methods](#) for details).

Identification of putative antisense-acting riboswitches

Considering the direction of transcription of the identified riboswitches in relation to that of their target genes, we initially determined that 98,082 (98.8%) correspond to sense-acting riboswitches and 1,185 (1.2%) to antisense-acting riboswitches. To avoid a statistical bias caused by the overrepresentation in the KEGG database of genomic sequences of microorganisms with several occurrences due to strains of the same species, all subsequent analyses of our study were performed considering only a representative genomic sequence per species (See [Materials and Methods](#) for details).

To find possible trends in which riboswitches might have antisense regulation, we grouped the 50 different classes of riboswitches described in the Rfam database into 27 families according to their shared recognition of their target molecules (S1 Table) (see [Materials and Methods](#) section). Subsequently, based on the absolute frequency for each one of the riboswitch families in the sense and antisense orientations, we evaluated the statistical significance of their tendencies to act as antisense-acting riboswitches using a hypergeometric distribution analysis. The results of this analysis are presented in [Table 1](#). What stands out in this table is that the *c*-di-GMP riboswitch family, commonly involved in cell signaling processes and definition of cell fates such as motility, virulence and biofilm formation, is the one with the highest tendency to regulate its target genes as antisense-acting riboswitches. We also identified a significant tendency for the Cobalamin, *ykoK/M*-box, and MOCO families, which respond to changing concentrations of cobalamin, magnesium ions and molybdenum, respectively to act as antisense-acting riboswitches. Additionally, the family of riboswitches in which we did not identify antisense regulation are the tetrahydrofolate (THF) and the nickel-cobalt (NiCo) sensing riboswitches.

Classification of the types of genes regulated by antisense-acting riboswitches

To elucidate possible trends in which antisense-acting riboswitches regulate the expression of genes, and to understand the nature of the biochemical implications of such regulation, we classified antisense-regulated genes into ten groups: a) genes coding enzymes involved in the interconversion of a compound related to the metabolite sensed by the riboswitch; b) genes coding enzymes that require a cofactor sensed by the riboswitch to become active or to participate in a metabolic pathway wherein another enzyme requires the cofactor; c) genes coding transporter proteins for compounds related to the metabolite sensed by the riboswitch; d) genes coding cell signaling proteins; e) genes coding transcriptional regulators; f) genes coding transposases; g) pseudogenes; h) uncharacterized ORFs with clear orthologs in multiple genomes; i) hypothetical small genes without orthologs; and j) other types of genes. [Table 2](#) shows the absolute and relative frequency of genes potentially regulated by the antisense-acting riboswitches which belong to the different groups above-mentioned.

Table 1. Tendency of riboswitch families to act in an antisense orientation.

Riboswitch family	Sense-acting riboswitches absolute frequency	Antisense-acting riboswitches absolute frequency	P-value
c-di-GMP	2,554	152	4.55E-62
Cobalamin	10,769	190	1.05E-10
ykoK / M-box	828	27	1.67E-06
MOCO	834	24	4.34E-05
AAC_AAD	97	3	0.10
glnA	97	2	0.30
nhaA-I	180	2	0.60
sulI	195	2	0.64
Mg_sensor	104	1	0.69
Glycine	3,153	32	0.74
DUF1646	125	1	0.75
raiA	127	1	0.76
ydaO-yuaA	2,799	27	0.81
FMN	3,472	32	0.89
ykkC / Guanidine	2,526	20	0.96
crcB / Fluoride	1,446	9	0.98
pfl / ZMP-ZTP	1,001	5	0.99
TPP	9,766	83	1.00
Purine	2,289	13	1.00
PreQ1	771	2	1.00
glmS	635	1	1.00
Lysine	1,974	8	1.00
SAM	8,039	56	1.00
yypP-ykoY	1,907	5	1.00
THF	290	0	1.00
NiCo	100	0	1.00
T-box	15,851	107	1.00

<https://doi.org/10.1371/journal.pone.0281744.t001>

Biochemical outcomes of antisense-acting riboswitch regulation

To exemplify each of the different groups of genes identified as being regulated by antisense-acting riboswitches, we chose a representative example of each group. Then, we analyzed in

Table 2. Frequency of genes potentially regulated by antisense-acting riboswitches.

Gene classification	Absolute frequency of genes regulated by antisense-acting riboswitches	Relative frequency (%)
Interconversion	116	14.41
Cofactor	25	3.11
Transporters	69	8.57
Signaling	12	1.49
TFs	70	8.70
Transposase	54	6.71
Pseudogene	50	6.21
Uncharacterized	57	7.08
Hypothetical	274	34.04
Other	78	9.69

<https://doi.org/10.1371/journal.pone.0281744.t002>

detail the possible biochemical implications of their regulation. S3 Table shows the genes identified and their corresponding groups.

a) Genes encoding enzymes involved in the interconversion of compounds related to the metabolite sensed by the riboswitch. We identified a ykkC-ykkD or Guanidine-I [26] antisense-acting riboswitch in the *Paenibacillus polymyxa* SC2 genome that regulates the ppm-PPSC2_12215 gene that encodes an amidase protein, AmyE (Fig 2A). This enzyme, better known as 4-guanidinobutanamide amidohydrolase, is involved in the arginine metabolism pathway catalyzing the conversion reaction from 4-guanidinobutanamide into 4-guanidinobutanate, resulting in the conversion of an amide into a carboxylic acid functional group.

Arginine is the core compound of this pathway, and it acts as a precursor in the biosynthesis of several polyamines, such as putrescine and spermidine, which are essential in several physiological processes. It can also be a poor nitrogen source [27]. Furthermore, arginine is the only amino acid with a guanidine group in its composition, despite the fact that it comprises compounds that are prominent within bacterial cells, including creatinine and secondary metabolites such as streptomycin [28–30].

The antisense-acting riboswitch identified at the 3' end of the ppm-PPSC2_12215 gene would enable the synthesis of the 4-guanidinobutanamide amidohydrolase enzyme when guanidine is present; in this case, its availability allows the progression along the pathway while being a part of 4-guanidinobutanamide, a compound that is catabolized into butanoate and urea in a subsequent step. On the other hand, when the riboswitch senses no guanidine available, the pathway is interrupted to maintain the transient guanidine levels while being a part of the 4-guanidinobutanamide compound.

As expected, the genes in this group that are regulated by active antisense-acting riboswitches are those that encode enzymes with transferase activities such as aminotransferases, phosphatases, acyltransferases, synthetases, and dehydrogenases. The regulation of the antisense-acting riboswitches of this group inhibits the synthesis of their corresponding target genes when they detect low concentrations of the substrates involved. The most important families of active antisense riboswitches in this group mainly include the cobalamin, TPP, T-box, SAM, FMN, and MOCO riboswitches (see S3 Table).

b) Genes coding enzymes that require a cofactor sensed by the riboswitch to become active or to participate in a metabolic pathway wherein another enzyme requires the cofactor. A cobalamin antisense-acting riboswitch regulating the clj-CLJU_c22210 gene that encodes the bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase enzyme, better known as the AICARFT/IMPCHase enzyme, was identified in the *Clostridium ljungdahlii* DSM 13528 genome through our bioinformatic analysis. *C. ljungdahlii* is an anaerobic acetogenic bacterium that ferments sugar or CO₂, that synthesizes gas (CO/H₂), and that can utilize CO as a substrate. Moreover, this organism possesses a significant number of proteins that require cobalamin as a cofactor and is also able to synthesize cobalamin via an anaerobic pathway [31,32].

The AICARFT/IMPCHase enzyme is involved in the one-carbon metabolism pathway, which comprises the folate and methionine cycles and allows the generation of methyl groups that are used in different methylation reactions and in the biosynthesis of important anabolic precursors, one of which is 5-methyl-tetrahydrofolate (5-methylTHF), whose synthesis can be carried out through two different routes. The starting point of both is tetrahydrofolate (THF) synthesis from folate by the activity of the dihydrofolate reductase enzyme DfrD. In one route, THF is converted into 10-formylTHF by the bifunctional AICARFT/IMPCHase enzyme, which encodes a gene regulated by the cobalamin antisense-acting riboswitch identified. Then, 10-formylTHF acts as the substrate for the synthesis of 5,10-methenylTHF and 5,10-methyleneTHF in two consecutive reactions catalyzed by the bifunctional methylenetetrahydrofolate

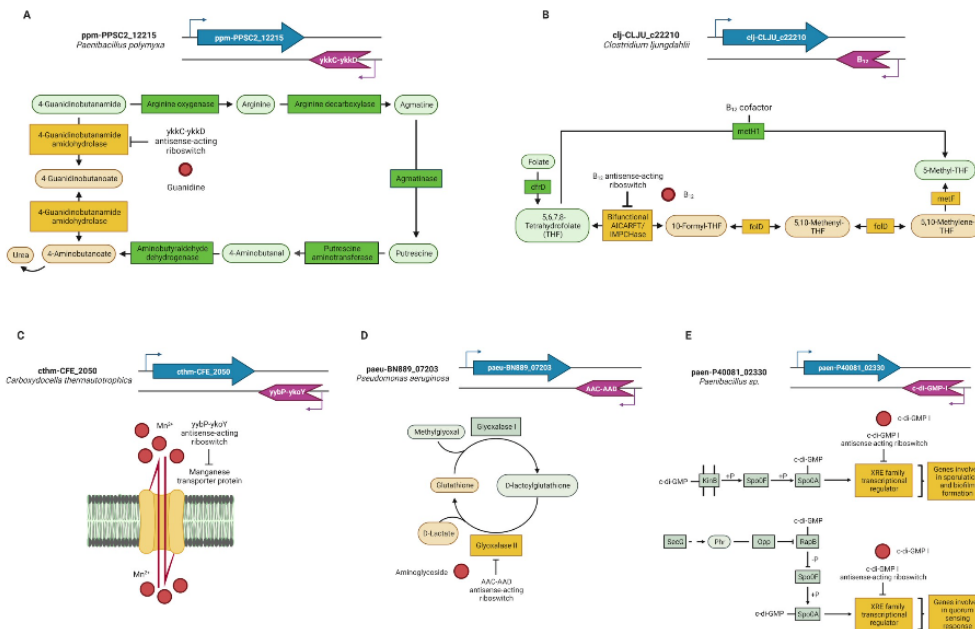


Fig 2. Representative pathways of genes regulated by antisense-acting riboswitches. (A) Gene encoding an enzyme involved in the interconversion of a compound related to the metabolite sensed by the riboswitch; (B) Gene encoding an enzyme that requires a cofactor sensed by the riboswitch to become active or participates in a metabolic pathway wherein another enzyme requires the cofactor; (C) Gene encoding a transporter protein for a compound related to the metabolite sensed by the riboswitch; (E) Gene that codes for a transcriptional regulator. Metabolic route active in the presence (yellow) or absence (green) of the metabolite sensed (red circle) by the antisense-acting riboswitch. Representative elements are colored as follows: Target genes, blue; antisense-acting riboswitch, violet. Enzymes are indicated by rectangles; compounds are indicated by ovals. Created with BioRender.com.

<https://doi.org/10.1371/journal.pone.0281744.g002>

dehydrogenase (NADP+)/methylnetetrahydrofolate cyclohydrolase enzyme, FOLD. This last compound is the precursor for the synthesis of 5-methylTHF catalyzed by methylenETHF reductase enzyme, MetF, with the participation of NADPH and flavoproteins (Fig 2B). Additionally, through another route, 5-methylTHF is directly synthesized from THF by the methionine synthase enzyme MetH1, whose activity is cobalamin-dependent (Fig 2B).

The cobalamin antisense-acting riboswitch identified regulates the *clj-CLJU_c22210* gene, which codes for the bifunctional AICARFT/IMPCHase, which allows the selection of the route that will be used for the synthesis of 5-methylTHF, according to the bacterial cobalamin availability. We propose that when the intracellular concentration of cobalamin is high, 5-methylTHF can be directly synthesized from THF by the cobalamin-dependent enzyme MetH1 (Fig 2B), while when there are changes in the intracellular concentration of cobalamin, the riboswitch-dependent regulation mechanism will favor the activity of the bifunctional AICARFT/IMPCHase enzyme. By these means, the regulation of the *clj-CLJU_c22210* gene by the cobalamin antisense-acting riboswitch guarantees 5-methylTHF biosynthesis and re-assures the correct cobalamin utilization due to its metabolically expensive biosynthesis and uptake. In addition, the metabolic intermediates of this route have an essential role in the biosynthesis of vital molecules such as purines, DNA, CoA, and serine (Fig 2B) [32,33].

Due to the nature of the genes regulated in this group, which are characterized as encoding enzymes that require a cofactor to be active, the family of cobalamin antisense-acting riboswitches (Cobalamin, AdoCbl, and AdoCbl-variant) is the most important member of this group, since this family represents 38% of the total number of riboswitches in this category. In general, the families of enzymes with isomerase and methyltransferase activities are the ones that most frequently require cobalamin, or some of its derivatives, as cofactors. Another important family of antisense-acting riboswitches in this group is the TPP family. The TPP cofactor consists of a pyrimidine ring, a thiazole ring, and a pyrophosphate functional group, which are connected to each other with the thiazole ring being the portion commonly involved in enzymatic reactions, as it is the most reactive part of the molecule. We also found examples of molybdenum cofactor (MOCO) antisense-acting riboswitches that regulate genes involved in the metabolism of molybdenum and tungsten cofactors (S3 Table) [34].

c) Genes encoding transporter proteins for compounds related to the metabolite sensed by the riboswitch. An antisense-acting yybP-ykoY [26] riboswitch, whose binding metabolite is manganese (Mn^{2+}), regulates the cthm-CFE_2050 gene that encodes an Mn^{2+} transporter protein in the *Carboxydocella therrautotrophica* 019 genome (Fig 2C). Mn^{2+} is an essential ion in bacterial metabolism that participates in several metabolic processes, acting mainly as a cofactor in various enzymes and as a protective agent against reactive oxygen species (ROS) through the transient substitution of Mn^{2+} for iron (Fe) molecules in the active sites of some enzymes to prevent the oxidative damage of proteins [35]. On the other hand, excess Mn^{2+} can be toxic since its accumulation in bacteria is directly related to a deficiency of other ions, which can then provoke the incorrect metalation of transcription factors and critical enzymes, affecting cellular development, virulence, and ROS sensitivity [26,36–39].

The gene identified as subject to regulation by this antisense-acting riboswitch is extensively distributed throughout the phylum Firmicutes and has an essential role in *C. therrautotrophica* 019 metabolism. This bacterium has been defined as carboxydophilic, but its metabolic capacities remain mostly unknown. Nevertheless, as the ability to transport Mn^{2+} has been observed in other species, it is known that these species can proliferate in hostile environments, including over mineral substrates compounded by Mn^{2+} , such as diatomite and glauconite, in the absence of a carbon source [40,41]. The antisense-acting riboswitch identified at the 3' end of the cthm-CFE_2050 gene would prevent the excessive internal accumulation of the Mn^{2+} that would otherwise have toxic effects.

Many genes of this third group are members of the ABC-type transport systems or code for different transporter proteins, such as the Co/Zn/Cd cation transporters, the Na^+/H^+ -dicarboxylate symporters, or the Ca^{2+}/Na^+ antiporters. The most common antisense-acting riboswitches of this group are the cobalamin, c-di-GMP-I, T-box, SAM, ykoK, TPP, and MOCO riboswitches (see S3 Table).

d) Genes coding cell signaling proteins. An AAC-AAD antisense-acting riboswitch, which senses and responds to aminoglycoside concentrations [42], was found to regulate the paeu-BN889_07203 gene that codes hydroxyacylglutathione hydrolase enzyme synthesis in *Pseudomonas aeruginosa* PA38182 (Fig 2D). This enzyme, better known as glyoxalase II (GlxII), participates in the methylglyoxal (MG) metabolism pathway through the glyoxalase (Glx) system, which is an enzymatic complex for critical detoxification. It comprises two metalloenzymes in charge of breaking down cytotoxic MG, metabolizing it into its corresponding α -hydroxy acids.

The first enzyme of this system is glyoxalase I (GlxI) or S-D-lactoyl-glutathione methylglyoxal lyase, which is an isomerizing divalent reductase that can be Zn^{2+} -dependent or non- Zn^{2+} -dependent (activated by Ni^{2+}/Co^{2+}). GlxI accepts MG as a glutathione hemiacetal (GSH) and catalyzes proton transfer, transforming GSH into its corresponding thioester,

lactoylglutathione. The second enzyme, GlxII, regenerates the thiol cosubstrate GSH and the α -hydroxy acid lactate in the case of MG detoxification. MG is highly toxic to cell reactive electrophilic species (RES) producer molecules, and they produce cellular stress through reactions with nucleophilic macromolecules. MG also acts as a transcriptional activator and participates in the kinase activation cascade, having a primary role in regulation processes or stress adjustment conditions, making it an essential signaling molecule [43,44].

During the *P. aeruginosa* pathogenic process, the bacteria are first attacked by the innate immunologic system, primarily via phagocytes. Within the phagocytes, ROS are generated in an NADPH oxidase-dependent manner as bactericidal substances. Thus, bacterial cells are exposed to oxidative stress, inducing damage due to interactions with cellular elements such as lipids, DNA, and proteins. This process leads to lipid peroxidation, DNA mutation, DNA-protein cross-linking, protein oxidation, and breakage. *P. aeruginosa* has evolved mechanisms to protect itself and survive under these hostile conditions. One of them is the production of GSH, the most abundant antioxidant molecule within the cells, which plays a relevant role in ROS removal, acting as an electron donor. GSH is also involved in cellular homeostasis preservation, sulfur transport regulation, metabolite combinations, xenobiotic detoxification, antibiotic resistance, enzymatic regulation, and stress response gene expression [45,46].

Compounds that produce cellular oxidative stress and their effects on different bacterial developmental stages have been widely studied. These compounds increase the aminoglycoside sensitivity of *P. aeruginosa* by disrupting membrane permeability, mainly during biofilm production, which is stimulated by sublethal antibiotic concentrations, such as tobramycin and gentamycin aminoglycosides. In addition, oxidative stress acts as a signal for the expression of efflux systems, which promotes bacterial antibiotic resistance [47–49].

The antisense-acting riboswitch identified at the 3' end of the para-BN889_07203 gene would enable the gene expression of the GlxII enzyme when the aminoglycoside is sensed in the environment. On the other hand, when no aminoglycoside is present in its medium, the expression of the enzyme is repressed, as the bacteria are not subject to oxidative stress from the antibiotic presence.

This fifth group of genes codes for proteins that are mainly involved in different stress responses and developmental processes, such as quorum-sensing autoinducer synthesis, and for different kinds of signal transduction histidine kinases. The principal antisense-acting riboswitches of this group are those that sense the cyclic dinucleotide second messengers c-di-GMP-I and cyclic di-AMP (YdaO/YuaA) riboswitches (S3 Table) [50].

e) Genes coding transcriptional regulators. We identified an antisense-acting c-di-GMP-I riboswitch [51] regulating the paen-P40081_02330 gene coding a transcriptional regulator of the XRE family in the *Paenibacillus* sp. FSL P4 genome (Fig 2E). This transcriptional regulator, whose nearest homolog is SinR, a transcriptional repressor known as the master regulator for biofilm formation in *Bacillus subtilis*, controls the expression of genes involved in exopolysaccharide, matrix protein production, and motility.

The synthesis of signaling molecules is one strategy bacteria employ to sense alterations in their environment in order to rapidly adjust according to those changes. One of these molecules is the bis-(3'-5')-cyclic dimeric GMP (c-di-GMP), which is able to induce the transition from a unicellular motile state to a multicellular sessile state through phosphorelay systems that control biofilm and spore formation (Fig 2E) [52–55].

For many bacteria, the quorum-sensing intercellular communication mechanism is essential for gene expression pattern coordination at the population level. Furthermore, for motility, biofilm matrix production, and spore formation processes, quorum sensing is critical, as it is one of the main strategies to sense and respond to environmental changes (Fig 2E) [56–59].

The c-di-GMP-I antisense-acting riboswitch regulation of a transcription factor coded by the paen-P40081_02330 gene that we propose is based on the detection of this molecule, triggering the expression of the genes involved in biofilm and spore formation, as well as quorum-sensing processes, and having a general effect on their expression when it is not sensed, thanks to its close relationship with the master regulator of these processes, Spo0A (Fig 2E).

As mentioned above, in the literature, there are only two examples of the regulation of gene expression mediated by riboswitches that act in an opposite direction, one of which is the cobalamin antisense-acting riboswitch identified at the 3' end of the gene encoding the Pocr transcriptional regulator of the enzymes involved in the metabolism of propanediol in *L. monocytogenes*, the enzymes of which require cobalamin as a cofactor to be active. In our study, we identified cobalamin antisense-acting riboswitches in the orthologous *pocR* genes of other *Listeria*s and in the Enterococcaceae *Vagococcus carniphilus*. The regulation of genes encoding transcription factors (TFs) makes it possible to extend the regulatory outcome of the riboswitches that act in an opposite direction to multiple functionally related genes in different operons, which have in common the operator site of the corresponding TFs. In our study, the c-di-GMP-I, cobalamin, and TPP antisense-acting riboswitches were the most frequent riboswitches of this group.

f) Genes coding transposases. Unlike the aforementioned examples, where regulation of gene expression mediated by antisense-acting riboswitches represents a selective advantage, we propose that in the case of transposon-encoding genes, the existence of antisense-acting riboswitches corresponds to the random insertion of the transposable element in chromosome regions where the riboswitches were previously regulating other genes in a canonical manner. Consistent with this hypothesis, Fig 3A shows the genomic context of the *bthU*-YBT1518_08170 gene in phylogenetically closely related organisms. In the genomes of all these organisms, except for that of the *Bacillus thuringiensis* YBT-1518, the T-box riboswitch regulates the downstream gene coding for a tryptophan transporter in a canonical manner.

In accordance, we evaluated the frequency with which the gene located upstream of transposase has an orthologous gene regulated by the same antisense-acting riboswitch. S4 Table shows that our hypothesis is consistent with their genomic context by over 50% of the cases of genes that code for transposases.

g) Pseudogenes. Another trend that we observed in our analysis was the presence of antisense-acting riboswitches next to ORFs annotated as pseudogenes, such as the *bacp*-SB24_06455 gene of *Bacillus* sp. Pc3 (Fig 3B). In a similar manner, as mentioned above, we performed a genome context analysis of the antisense-regulated gene and compared it to its nearest related organism genes. Based on the observed synteny conservation in all the genes, except in the *bacp*-SB24_06455 gene, we propose that the existence of an antisense-acting riboswitch does not correspond to a selective advantage for the organisms, but rather to a fortuitous event that might be related to the origin of the pseudogene (S5 Table).

h) Uncharacterized ORFs with clear orthologs in multiple genomes. The next category of genes regulated by antisense-acting riboswitches identified in our study corresponds to ORFs annotated as "uncharacterized" but with *bona fide* orthologs in other genomes. We included in this group genes larger than 300 nucleotides only (see S3 Table). An example of a gene that belongs to this category is *colw*-A3Q33_07785 of *Colwellia* sp. PAMC 21821, whose function is still unknown (Fig 3C).

i) Hypothetical small genes without orthologs. Genes identified as potentially regulated by an antisense-acting riboswitch included in this category are characterized by having small sizes of less than three hundred bases and lacking apparent orthologs in the set of sequenced genomes, accounting for 34% of the elements. These characteristics allow us to consider that these "hypothetical genes" correspond to probable annotation errors of open reading frames

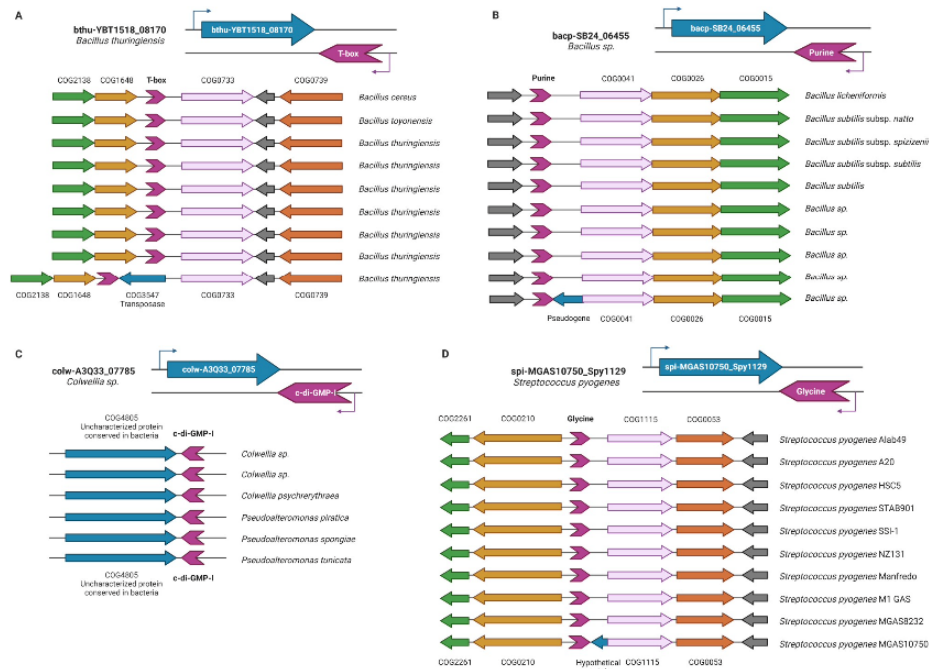


Fig 3. Representative syntenic context of genes regulated by antisense-acting riboswitches. The upper panel on each section shows in detail an example of the different types of genes regulated by the antisense-acting riboswitches, while the lower panel shows the context in closely related organisms compared to the example. (A) Gene coding for a transposase: COG2138-Uncharacterized conserved protein, COG1648-Siroheme synthase, COG0733-Na⁺-dependent transporters of the SNF family, COG0739-Membrane proteins related to metalloendopeptidases; (B) Pseudogene: COG0041-Phosphoribosylcarboxyaminoimidazole (NCAIR) mutase, COG0026-Phosphoribosylaminoimidazole carboxylase (NCAIR synthetase), COG0015-Adenylosuccinate lyase; (C) Gene coding for an uncharacterized ORF with clear orthologs in multiple genomes; (D) Hypothetical small gene without orthologs: COG2261-Predicted membrane protein, COG0210-Superfamily I DNA and RNA helicases, CO1115-Na⁺/alanine symporter, COG0053-Predicted Co/Zn/Cd cation transporters. Representative elements are colored as follows: Riboswitch spuriously regulated genes, blue; antisense-acting riboswitch, violet; *bona fide* regulated gene, pink. Genes are colored according to the COG they belong to. Created with BioRender.com.

<https://doi.org/10.1371/journal.pone.0281744.g003>

that, in most cases, are incorrectly annotated as genes. As in our previous group, the comparative analysis of syntenic regions in phylogenetically closely related organisms confirms our hypothesis that most of those riboswitches (72% of our studied cases) regulate the neighboring upstream gene of the hypothetical gene, according to the consistency of its genomic context. This is the case for the hypothetical small gene *spi-MGAS10750_Spy1129* in the *Streptococcus pyogenes* MGAS10750 genome with a Glycine riboswitch at its 3' end and a D-alanine glycine permease gene, whose orthologous counterparts are regulated by Glycine riboswitches in phylogenetically related organisms (Fig 3D). In S6 Table, we list all these hypothetical genes as well as the most likely regulated gene neighbor.

j) Other types of genes. Finally, the last group in our analysis is composed of antisense-acting riboswitches, identified next to genes that could not be classified into any of the previously described groups. The genes in this category represent approximately 9% of the antisense

riboswitches identified in our study, and the biochemical implications of their regulation are not obvious at first glance.

Genes regulated by tandem antisense-acting riboswitches

Tandem riboswitches have been found to regulate gene expression in several organisms [60]. They might include either two or more copies of the same kind of riboswitch to extend the range of regulation by the sensing signal, or copies of different riboswitch classes next to those genes that require a more complex regulation which obeys to more than one intracellular signal [60]. In our study, we searched for tandemly arranged antisense-acting riboswitches and found seven genes potentially regulated by elements with this characteristic (S7 Table); four of them correspond to transposase genes, two of them to ABC-transporter operons, and the last one to a small hypothetical gene of 126 nucleotides in length. As described for our classification of small hypothetical genes, these commonly correspond to annotation errors. In accordance with this hypothesis, the upstream genes to a small hypothetical gene regulated by an antisense-acting riboswitch tend to be regulated by the same riboswitch family in orthologous genes which do not have such small hypothetical gene. In the *Ralstonia insidiosa* FC1138 sequence genome, we identified a small gene, rin-ACS15_3586, annotated as hypothetical, with a tandemly arranged couple of glycine antisense-acting riboswitches. Upstream of this gene, we found the *gcvT-gcvH-gcvP* operon, coding for the glycine cleavage system T and H proteins, and the HNH endonuclease family protein, respectively, which are genes commonly regulated by glycine riboswitches in other organisms.

About a decade ago, when only two examples of antisense-acting riboswitches had been described, the number of publicly available and fully sequenced genomic sequences from prokaryotic organisms was barely one hundred. Today, we have tens of thousands of newly sequenced genomes. The question, then, is why no other examples of this kind of noncanonical riboswitch have been reported. Are there no more examples of antisense-acting riboswitches in orthologous genes or other kinds of regulatory systems involving enzymes that require cofactors analogous to those mentioned above? What new metabolic conditions might require regulation based on antisense-acting riboswitches that could be identified in a global study using currently available genomic sequences? These are the main questions that motivated our study. As a result of our analysis, we identified 292 antisense-acting riboswitches whose regulatory outcomes seem biologically significant. Analysis of these potentially antisense action riboswitches allowed us to classify them into ten different groups according to the nature of the genes they regulate or a common characteristic shared among them. Our results are consistent with our biochemical understanding of the requirements of organisms under different metabolic conditions and how these requirements are satisfied by the regulatory responses of genes subject to antisense-acting riboswitch responses. Additionally, it is important to note that our method was able to identify the two antisense-acting riboswitches previously described. Our findings concerning antisense-acting riboswitches provide evidence of the evolutionary plasticity that organisms have developed to obtain opposite regulatory responses using the same regulatory elements, but only modifying their location (5' or 3' UTR) and transcription orientation (sense or antisense) according to their target genes.

Supporting information

S1 Table. Riboswitch families.
(XLSX)

S2 Table. Antisense-acting riboswitches.

(XLSX)

S3 Table. Antisense regulated genes.

(XLSX)

S4 Table. Transposases.

(XLSX)

S5 Table. Pseudogenes.

(XLSX)

S6 Table. Hypothetical proteins.

(XLSX)

S7 Table. Genes regulated by tandem antisense-acting riboswitches.

(XLSX)

S1 File.

(ZIP)

Acknowledgments

We sincerely thank Walter Santos, Ricardo Ciria, Arturo Ocadiz, and Juan Manuel Hurtado for computer support and Shirley Ainsworth for bibliographical assistance.

Author Contributions

Conceptualization: Mariela Serrano-Gutiérrez, Enrique Merino.

Data curation: Mariela Serrano-Gutiérrez, Enrique Merino.

Formal analysis: Mariela Serrano-Gutiérrez, Enrique Merino.

Funding acquisition: Enrique Merino.

Investigation: Mariela Serrano-Gutiérrez, Enrique Merino.

Methodology: Mariela Serrano-Gutiérrez, Enrique Merino.

Project administration: Mariela Serrano-Gutiérrez, Enrique Merino.

Resources: Enrique Merino.

Software: Mariela Serrano-Gutiérrez, Enrique Merino.

Supervision: Mariela Serrano-Gutiérrez, Enrique Merino.

Validation: Mariela Serrano-Gutiérrez, Enrique Merino.

Visualization: Mariela Serrano-Gutiérrez, Enrique Merino.

Writing – original draft: Mariela Serrano-Gutiérrez, Enrique Merino.

Writing – review & editing: Mariela Serrano-Gutiérrez, Enrique Merino.

References

1. Antunes D, Jorge NAN, Caffarena ER, Passetti F. Using RNA sequence and structure for the prediction of riboswitch aptamer: A comprehensive review of available software and tools. *Frontiers in Genetics*. Frontiers Media S.A.; 2018. <https://doi.org/10.3389/fgene.2017.00231> PMID: 29403526

2. Barrick JE, Breaker RR. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* 2007; 8. <https://doi.org/10.1186/gb-2007-8-11-r239> PMID: 17997835
3. Breaker RR. Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol.* 2012; 4. <https://doi.org/10.1101/cshperspect.a003566> PMID: 21106649
4. Sinumvayo JP, Zhao C, Tuyishime P. Recent advances and future trends of riboswitches: attractive regulatory tools. *World Journal of Microbiology and Biotechnology.* Springer Netherlands; 2018. <https://doi.org/10.1007/s11274-018-2554-0> PMID: 30413889
5. Winkler WC, Breaker RR. REGULATION OF BACTERIAL GENE EXPRESSION BY RIBOSWITCHES. *Annu Rev Microbiol.* 2005; 59: 487–517. <https://doi.org/10.1146/annurev.micro.59.030804.121336> PMID: 16153177
6. Breaker RR. Prospects for Riboswitch Discovery and Analysis. *Molecular Cell.* 2011. pp. 867–879. <https://doi.org/10.1016/j.molcel.2011.08.024> PMID: 21925376
7. Garst AD, Edwards AL, Batey RT. Riboswitches: Structures and mechanisms. *Cold Spring Harbor Perspectives in Biology.* Cold Spring Harbor Laboratory Press; 2011. pp. 1–13. <https://doi.org/10.1101/cshperspect.a003533> PMID: 20943759
8. Mccown PJ, Corbino KA, Stav S, Sherlock ME, Breaker RR. Riboswitch diversity and distribution. 2017. <https://doi.org/10.12611/ma.061234>
9. Sherwood A V., Henkin TM. Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annu Rev Microbiol.* 2016; 70: 361–374. <https://doi.org/10.1146/annurev-micro-091014-104306> PMID: 27607554
10. Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr Opin Microbiol.* 2007; 10: 102–109. <https://doi.org/10.1016/j.mib.2007.03.012> PMID: 17387036
11. Sesto N, Wurtzel O, Archambaud C, Sorek R, Cossart P. The excludon: A new concept in bacterial antisense RNA-mediated gene regulation. *Nat Rev Microbiol.* 2013; 11: 75–82. <https://doi.org/10.1038/nrmicro2934> PMID: 23268228
12. Pavlova N, Kaloudas D, Penchovsky R. Riboswitch distribution, structure, and function in bacteria. *Gene.* 2019; 708: 38–48. <https://doi.org/10.1016/j.gene.2019.05.036> PMID: 31128223
13. Georg J, Hess WR. cis -Antisense RNA, Another Level of Gene Regulation in Bacteria. *Microbiol Mol Biol Rev.* 2011; 75: 286–300. <https://doi.org/10.1128/MMBR.00032-10> PMID: 21646430
14. Breaker RR. Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol.* 2012; 4. <https://doi.org/10.1101/cshperspect.a003566> PMID: 21106649
15. Grundy FJ, Henkin TM. The S box regulon: A new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Mol Microbiol.* 1998; 30: 737–749. <https://doi.org/10.1046/j.1365-2958.1998.01105.x> PMID: 10094622
16. André G, Even S, Putzer H, Burguière P, Croux C, Danchin A, et al. S-box and T-box riboswitches and antisense RNA control a sulfur metabolic operon of *Clostridium acetobutylicum*. *Nucleic Acids Res.* 2008; 36: 5955–5969. <https://doi.org/10.1093/nar/gkn601> PMID: 18812398
17. Mellin JR, Tiensuu T, Bécavin C, Gouin E, Johansson J, Cossart P. A riboswitch-regulated antisense RNA in *Listeria monocytogenes*. *Proc Natl Acad Sci U S A.* 2013; 110: 13132–13137. <https://doi.org/10.1073/pnas.1304795110> PMID: 23878253
18. Cui X, Lu Z, Wang S, Jing-Yan Wang J, Gao X. CMsearch: Simultaneous exploration of protein sequence space and structure space improves not only protein homology detection but also protein structure prediction. *Bioinformatics.* 2016; 32: i332–i340. <https://doi.org/10.1093/bioinformatics/btw271> PMID: 27307635
19. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: An RNA family database. *Nucleic Acids Res.* 2003; 31: 439–441. <https://doi.org/10.1093/nar/gkg006> PMID: 12520045
20. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* 2021; 49: D545–D551. <https://doi.org/10.1093/nar/gkaa970> PMID: 33125081
21. Martínez-Guerrero CE, Ciria R, Abreu-Goodger C, Moreno-Hagelsieb G, Merino E. GeConT 2: gene context analysis for orthologous proteins, conserved domains and metabolic pathways. *Nucleic Acids Res.* 2008; 36. <https://doi.org/10.1093/nar/gkn330> PMID: 18511460
22. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol.* 2011; 7. <https://doi.org/10.1371/journal.pcbi.1002195> PMID: 22039361
23. Taboada B, Verde C, Merino E. High accuracy operon prediction method based on STRING database scores. *Nucleic Acids Res.* 2010; 38. <https://doi.org/10.1093/nar/gkq254> PMID: 20385580

24. Tatusov RL, Natale DA, Garkavtsev I V., Tatusova TA, Shankavaram UT, Rao BS, et al. The COG database: New developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 2001; 29: 22–28. <https://doi.org/10.1093/nar/29.1.22> PMID: 11125040
25. Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, et al. Rfam 13.0: Shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* 2018; 46: D335–D342. <https://doi.org/10.1093/nar/gkx1038> PMID: 29112718
26. Dambach M, Sandoval M, Updegrove TB, Anantharaman V, Aravind L, Waters LS, et al. The Ubiquitous yybP-ykoY Riboswitch Is a Manganese-Responsive Regulatory Element. *Mol Cell.* 2015; 57: 1099–1109. <https://doi.org/10.1016/j.molcel.2015.01.035> PMID: 25794618
27. Caldara M, Dupont G, Leroy F, Goldbeter A, De Vuyst L, Cunin R. Arginine biosynthesis in *Escherichia coli*: Experimental perturbation and mathematical modeling. *J Biol Chem.* 2008; 283: 6347–6358. <https://doi.org/10.1074/jbc.M705884200> PMID: 18165237
28. Nelson J, Atilho R, Sherlock M, Stockbridge R, Breaker R. Metabolism of Free Guanidine in Bacteria is Regulated by a Widespread Riboswitch Class. *Mol Cell.* 2017; 65: 220–230. <https://doi.org/10.1016/j.molcel.2016.11.019> PMID: 27989440
29. Reiss C, Xiong Y, Strobel S. Structural basis for ligand binding to the guanidine-I riboswitch. *Structure.* 2017; 25: 195–202. <https://doi.org/10.1016/j.str.2016.11.020> PMID: 28017522
30. Sinn M, Hauth F, Lenkeit F, Weinberg Z, Hartig JS. Widespread bacterial utilization of guanidine as nitrogen source. *Mol Microbiol.* 2021; 116: 200–210. <https://doi.org/10.1111/mmi.14702> PMID: 33590553
31. Köpke M, Held C, Hujer S, Liesegang H, Wiezer A, Wollherr A, et al. *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc Natl Acad Sci U S A.* 2010; 107: 13087–13092. <https://doi.org/10.1073/pnas.1004716107> PMID: 20616070
32. Romine MF, Rodionov DA, Maezato Y, Anderson LN, Nandhikonda P, Rodionova IA, et al. Elucidation of roles for vitamin B12 in regulation of folate, ubiquinone, and methionine metabolism. *Proc Natl Acad Sci U S A.* 2017; 114: E1205–E1214. <https://doi.org/10.1073/pnas.1612360114> PMID: 28137868
33. Yi J, Huang H, Liang J, Wang R, Liu Z, Li F, et al. A Heterodimeric Reduced-Ferredoxin-Dependent Methylene-tetrahydrofolate Reductase from Syngas-Fermenting *Clostridium ljungdahlii*. *Microbiol Spectr.* 2021; 9: e00958–21. <https://doi.org/10.1128/Spectrum.00958-21> PMID: 34643446
34. Regulski EE, Moy RH, Weinberg Z, Barrick JE, Yao Z, Ruzzo WL, et al. A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Mol Microbiol.* 2008; 68: 918–932. <https://doi.org/10.1111/j.1365-2958.2008.06208.x> PMID: 18363797
35. Waters LS. Bacterial manganese sensing and homeostasis. *Curr Opin Chem Biol.* 2020; 55: 96–102. <https://doi.org/10.1016/j.cbpa.2020.01.003> PMID: 32086169
36. Kehres DG, Maguire ME. Emerging themes in manganese transport, biochemistry and pathogenesis in bacteria. *FEMS Microbiol Rev.* 2003; 27: 263–290. [https://doi.org/10.1016/S0168-6445\(03\)00052-4](https://doi.org/10.1016/S0168-6445(03)00052-4) PMID: 12829271
37. Richer E, Courville P, Bergevin I, Cellier MFM. Horizontal Gene Transfer of “Prototype” Nmp in Bacteria. *J Mol Evol.* 2003; 57: 363–376. <https://doi.org/10.1007/s00239-003-2472-z> PMID: 14708570
38. Waters LS, Sandoval M, Storz G. The *Escherichia coli* MntR miniregulon includes genes encoding a small protein and an efflux pump required for manganese homeostasis. *J Bacteriol.* 2011; 193: 5887–5897. <https://doi.org/10.1128/JB.05872-11> PMID: 21908668
39. Zeinert R, Martinez E, Schmitz J, Senn K, Usman B, Anantharaman V, et al. Structure–function analysis of manganese exporter proteins across bacteria. *J Biol Chem.* 2018; 293: 5715–5730. <https://doi.org/10.1074/jbc.M117.790717> PMID: 29440394
40. Slobodkina GB, Panteleeva AN, Sokolova TG, Bonch-Osmolovskaya EA, Slobodkin AI. *Carboxydocella manganica* sp. nov., a thermophilic, dissimilatory Mn(IV)- and Fe(III)-reducing bacterium from a Kamchatka hot spring. *Int J Syst Evol Microbiol.* 2012; 62: 890–894. <https://doi.org/10.1099/ijms.0.027623-0> PMID: 21642487
41. Toshchakov S V., Lebedinsky A V., Sokolova TG, Zavarzina DG, Korzhnikov AA, Teplyuk A V., et al. Genomic insights into energy metabolism of *Carboxydocella thermotrophica* coupling hydrogenogenic co oxidation with the reduction of Fe(III) minerals. *Frontiers in Microbiology.* 2018. <https://doi.org/10.3389/fmicb.2018.01759> PMID: 30123201
42. He W, Zhang X, Zhang J, Jia X, Zhang J, Sun W, et al. Riboswitch control of induction of aminoglycoside resistance acetyl and adenylyl-transferases. *RNA Biol.* 2013; 10: 1266–1273. <https://doi.org/10.4161/ma.25757> PMID: 23880830
43. Sukdeo N, Clugston SL, Daub E, Honek JF. Distinct classes of glyoxalase I: Metal specificity of the *Yersinia pestis*, *Pseudomonas aeruginosa* and *Neisseria meningitidis* enzymes. *Biochem J.* 2004; 384: 111–117. <https://doi.org/10.1042/BJ20041006> PMID: 15270717

44. Sukdeo N, Honek JF. *Pseudomonas aeruginosa* contains multiple glyoxalase I-encoding genes from both metal activation classes. *Biochim Biophys Acta—Proteins Proteomics*. 2007; 1774: 756–763. <https://doi.org/10.1016/j.bbapap.2007.04.005> PMID: 17513180
45. Lee C, Park C. Bacterial responses to glyoxal and methylglyoxal: Reactive electrophilic species. *Int J Mol Sci*. 2017; 18. <https://doi.org/10.3390/ijms18010169> PMID: 28106725
46. Wongsaraj L, Saninjak K, Romsang A, Duang-nkern J, Trinachartvanit W, Vattanaviboon P, et al. *Pseudomonas aeruginosa* glutathione biosynthesis genes play multiple roles in stress protection, bacterial virulence and biofilm formation. *PLoS One*. 2018; 13: 1–18. <https://doi.org/10.1371/journal.pone.0205815> PMID: 30325949
47. da Cruz Nizer WS, Inkovskiy V, Versey Z, Stempel N, Cassol E, Overhage J. Oxidative stress response in *pseudomonas aeruginosa*. *Pathogens*. 2021; 10: 1–32. <https://doi.org/10.3390/pathogens10091187> PMID: 34578219
48. Lochab V, Jones TH, Dusane DH, Peters CW, Stoodley P, Wozniak DJ, et al. Ultrastructure imaging of *Pseudomonas aeruginosa* lawn biofilms and eradication of the tobramycin-resistant variants under in vitro electrochemical treatment. *Sci Rep*. 2020; 10: 1–12. <https://doi.org/10.1038/s41598-020-66823-y> PMID: 32555250
49. Pan X, Dong Y, Fan Z, Liu C, Xia B, Shi J, et al. In vivo host environment alters *pseudomonas aeruginosa* susceptibility to aminoglycoside antibiotics. *Front Cell Infect Microbiol*. 2017; 7: 1–12. <https://doi.org/10.3389/fcimb.2017.00083> PMID: 28352614
50. Jenal U, Reinders A, Lori C. Cyclic di-GMP: Second messenger extraordinaire. *Nat Rev Microbiol*. 2017; 15: 271–284. <https://doi.org/10.1038/nrmicro.2016.190> PMID: 28163311
51. Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, et al. Riboswitches in eubacteria sense the second messenger cyclic Di-GMP. *Science (80-)*. 2008; 321: 411–413. <https://doi.org/10.1126/science.1159519> PMID: 18635805
52. Chen Y, Chai Y, Guo J hua, Losick R. Evidence for cyclic Di-GMP-mediated signaling in *Bacillus subtilis*. *J Bacteriol*. 2012; 194: 5080–5090. <https://doi.org/10.1128/JB.01092-12> PMID: 22821967
53. Reyes-Darias JA, Krell T. Riboswitches as Potential Targets for the Development of Anti-Biofilm Drugs. *Curr Top Med Chem*. 2017; 17: 1945–1953. <https://doi.org/10.2174/1568026617666170407163517> PMID: 28403796
54. Sornchuer P, Saninjak K, Prathaphan P, Tiengtip R, Wattanaphansak S. Antimicrobial Susceptibility Profile and Whole-Genome Analysis of a Strong Biofilm-Forming *Bacillus* Sp. B87 Strain Isolated from Food. *Microorganisms*. 2022; 10. <https://doi.org/10.3390/microorganisms10020252> PMID: 35208707
55. Tamayo R. Cyclic diguanylate riboswitches control bacterial pathogenesis mechanisms. *PLoS Pathog*. 2019; 15: 1–7. <https://doi.org/10.1371/journal.ppat.1007529> PMID: 30730989
56. Rahman F Bin, Sarkar B, Moni R, Rahman MS. Molecular genetics of surfactin and its effects on different sub-populations of *Bacillus subtilis*. *Biotechnol Reports*. 2021; 32: e00686. <https://doi.org/10.1016/j.btre.2021.e00686> PMID: 34786355
57. Voichek M, Maaß S, Kroniger T, Becher D, Sorek R. Peptide-based quorum sensing systems in *Paenibacillus polymyxa*. *Life Sci Alliance*. 2020; 3: 1–12. <https://doi.org/10.26508/lsa.202000847> PMID: 32764104
58. Weiss CA, Hoberg JA, Liu K, Tu BP, Winkler WC. Single-Cell Microscopy Reveals That Levels of Cyclic di-GMP Vary among *Bacillus subtilis* Subpopulations. 2019; 1–17.
59. Newman JA, Rodrigues C, Lewis RJ. Molecular basis of the activity of SinR Protein, the master regulator of biofilm formation in *bacillus subtilis*. *J Biol Chem*. 2013; 288: 10766–10778. <https://doi.org/10.1074/jbc.M113.455592> PMID: 23430750
60. Sherlock ME, Higgs G, Yu D, Widner DL, White NA, Sudarsan N, et al. Architectures and complex functions of tandem riboswitches. *RNA Biol*. 2022; 19: 1059–1076. <https://doi.org/10.1080/15476286.2022.2119017> PMID: 36093908