



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

**Identificación de sRNAs en *Geobacter sulfurreducens* durante la
reducción de fumarato y Fe (III)**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M. en C. María Getzabeth González Gutiérrez

TUTOR PRINCIPAL

Dra. Katy Juárez López
Instituto de Biotecnología

MIEMBROS DEL COMITÉ TUTOR

Dr. José Luis Puente García
Instituto de Biotecnología

Dr. Miguel Ángel Cevallos Gaos
Centro de Ciencias Genómicas

Cuernavaca, Morelos. Junio, 2018.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Jurado de Examen

Dr. Enrique Merino Pérez
Instituto de Biotecnología, UNAM

Dr. Tomás David López Díaz
Instituto de Biotecnología, UNAM

Dr. Julio Collado Vides
Centro de Ciencias Genómicas, UNAM

Dr. Víctor Humberto Bustamante Santillán
Instituto de Biotecnología, UNAM

Dr. Armando Hernández Mendoza
Centro de Investigación en Dinámica Celular, UAEM

Este trabajo fue realizado bajo la tutoría de la Dra. Katy Juárez en el Departamento de Ingeniería Celular y Biocatálisis del Instituto de Biotecnología de la UNAM. Este proyecto fue financiado por los proyectos PAPIIT-DGAPA IN210017, Office of Science US Department of Energy (DOE) No. DEFC02-02ER63446 (PI Derek Lovley), Consejo de Ciencia y Tecnología (CONACYT) por medio de una beca para estudios de doctorado (Núm. de Becario 21942) y a través del proyecto de ciencia básica Núm. 179684. Así mismo, se conto con el apoyo económico por parte del Programa de Apoyo a Estudios de Posgrado (PAEP) para la asistencia a congresos y cursos nacionales e internacionales.

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México (UNAM) por todo lo que he recibido de esta casa de estudios. Por la formación que me ha dado más allá de la parte académica.

A la Dra. Katy Juárez por su confianza, paciencia, apoyo y libertad para llevar este proyecto. Gracias Katy, eres una gran científica pero sobre todo, eres una súper persona.

A los Dres. José Luis Puente y Miguel Ángel Cevallos por siempre darme buenos consejos y contribuir a mejorar mi trabajo y mi formación.

A los miembros de mi Jurado Revisor de Tesis, por sus sugerencias y correcciones que han enriquecido este trabajo.

A todos los Katitos que han pasado y siguen en el Lab Anita, Beto, Paloma, Fany, Emmanuel (perdón por todos lo que me faltaron) por su compañerismo, buena vibra y consejos.

Al Dr. Enrique Morett, Ing. Leticia Olvera y Biol. Maricela Olvera, por compartir su experiencia y conocimiento. Al Dr. Ricardo Grande de la Unidad de Secuenciación Masiva por su asesoría y disposición para el uso de sus instalaciones.

A la Dra. Leticia Vega, M. en C. Verónica Jiménez y M. en C. Aurora Labastida. Chicas: son las mujeres más inteligentes y buena onda que he conocido. Gracias por compartir conmigo toda su sabiduría, contribuir enormemente a este trabajo y ser tan lindas personas.

A toda la gente bonita del Lab 9. Humbe, Joel, Rodrigo, Emma y todas las personas que en algún momento coincidimos en ese espacio que es como el segundo hogar y que hacen que todo sea más llevadero y divertido.

Al personal de la Unidad de Posgrado del IBT: Toño, Gloria y Jalil: son unos soles y sin ustedes nos volveríamos locos. Gracias por hacer su trabajo de una forma tan excelente.

A mis papas, por ser siempre la más buena onda y amorosos del mundo y estar conmigo siempre. A mis hermanit@s y sobrinitos por siempre darme sus buenas vibras ya apoyo.

Finalmente quiero darles las gracias a mis personas favoritas en este mundo y a mi gordito: Adan, Ana Victoria y Chicle. Los amos con todo mi corazón y gracias por toda la felicidad que me dan día a día.



ÍNDICE

RESUMEN	10
ABSTRACT	11
CAPÍTULO I	12
Introducción	12
1.1 Características generales de las bacterias pertenecientes al género <i>Geobacter</i> .	12
1.2 Transferencia extracelular de electrones a óxidos de Fe(III) en <i>Geobacter sulfurreducens</i> .	13
1.3 Regulación genética de la transferencia de electrones en <i>G. sulfurreducens</i> .	14
1.4 sRNAs bacterianos y sus mecanismos de regulación.	14
CAPÍTULO II	19
Hipótesis y Objetivos	19
2.1 Hipótesis.	19
2.2 Objetivos.	19
Objetivo General.	19
Objetivos Específicos.	19
CAPÍTULO III	20
Mapeo Global de Sitios de Inicio de la Transcripción.	20
3.1 Materiales y Métodos.	20
3.1.1 Cepas Bacterianas y condiciones de cultivo.	20
3.1.2 Extracción y preparación de RNA.	20
3.1.3 Construcción de bibliotecas para mapeo de sitios de inicio de la transcripción.	20
3.1.4 Análisis bioinformático.	21
3.2 Resultados y Discusión.	22
3.2.1 Descripción general de los TSSs identificados.	22
3.2.2 Categorización de los TSS identificados.	25
3.2.3 Búsqueda de regiones promotoras.	27
CAPÍTULO IV	29
Análisis de expresión diferencial utilizado Fe(III) y fumarato como aceptores de electrones.	29
4.1 Materiales y Métodos	29
4.1.1 Cepas Bacterianas y condiciones de cultivo	29
4.1.2 Construcción de bibliotecas para análisis de RNAseq.	29
4.1.3 Análisis Bioinformático.	29
4.1.4 Validación de los datos de RNAseq por RT-PCR.	30
4.2 Resultados y Discusión.	31
4.2.1 Estadísticas generales de los datos de RNAseq.	31
4.2.2 Expresión diferencial utilizando Fumarato y Fe(III) como aceptor final de electrones.	32
4.2.3 Expresión de genes relacionados con proteínas de unión y transporte.	34
4.2.4 Expresión de genes asociados a energía y metabolismo.	35
4.2.5 Expresión diferencial en genes implicados en regulación y sistemas de dos componentes.	36
4.2.6. Expresión de genes relacionados con mecanismos de regulación genética a través de sRNAs.	37
CAPITULO V	39
Identificación y validación de sRNAs en regiones intergénicas	39
5.1 Materiales y Métodos.	39
5.1.1 Identificación de sRNAs en regiones intergénicas.	39
5.1.2 Validación mediante Northern blot.	39
5.1.3 Validación mediante 5' RACE (Rapid amplification of cDNA 5' ends).	40
5.1.4. Ensayos de RT-qPCR y de Co-transcripción.	41

5.2 Resultados y Discusión	41
5.2.1 Identificación de sRNAs en regiones intergénicas.	41
5.2.1 Validación y caracterización mediante northern blot, RT-qPCR y 5'RACE.	45
5.2.2 Análisis de expresión diferencial de los sRNAs identificados.	48
5.2.3 Ensayos de co-transcripción.	51
5.2.4 Identificación de homólogos de los sRNAs identificados y conservación de su secuencia.	52
5.2.5. Identificación de posibles genes blanco.	58
<u>CAPÍTULO VI</u>	<u>60</u>
Conclusiones generales y Perspectivas	60
6.1. Conclusiones generales.	60
6.2. Perspectivas.	61
<u>REFERENCIAS</u>	<u>62</u>
<u>ANEXOS</u>	<u>73</u>

ÍNDICE DE FIGURAS

FIGURA 1. MODELO PROPUESTO DE LA TRANSFERENCIA DE ELECTRONES A LOS ÓXIDOS DE Fe(III) EN <i>G. SULFURREDUCTENS</i>	13
FIGURA 2. MECANISMOS DE REGULACIÓN DE UN RIBOSWITCH.....	15
FIGURA 3. DIFERENTES TIPOS DE sRNAs BASADOS EN SU MECANISMO DE ACCIÓN.....	16
FIGURA 4. MECANISMOS DE REGULACIÓN DE UN RNA ANTI-SENSE.....	17
FIGURA 5. SITIOS DE INICIO DE LA TRANSCRIPCIÓN (TSSs) DETECTADOS EN LAS DOS CONDICIONES DE CRECIMIENTO.....	23
FIGURA 6. AGRUPACIÓN DE LOS SITOS DE INICIO DE LA TRANSCRIPCIÓN.....	24
FIGURA 7. NÚMERO DE COINCIDENCIAS POR POSICIÓN CON RESPECTO A LOS DATOS REPORTADOS POR QUI ET AL (2010).....	25
FIGURA 8. CATEGORIZACIÓN DE TSSs CON RESPECTO A SU LOCALIZACIÓN GENÓMICA.....	26
FIGURA 9. LOCALIZACIÓN GENÓMICA DEL GEN PPCD.....	28
FIGURA 10. DIAGRAMA DE VENN DE LA INTERSECCIÓN DE GENES DIFERENCIALMENTE EXPRESADOS DETECTADOS POR TRES DIFERENTES MÉTODOS.....	32
FIGURA 11. CATEGORÍAS FUNCIONALES DE LOS GENES CON EXPRESIÓN DIFERENCIA.....	33
FIGURA 12. CARACTERIZACIÓN DEL sRNA GSUSR27.....	45
FIGURA 13. CARACTERIZACIÓN DEL sRNA GSUSR41.....	46
FIGURA 14. CARACTERIZACIÓN DE LOS TRANSCRITOS GSUSR17, GSUSR40, GSUSR08 Y GSUSR36.....	47
FIGURA 15. CARACTERIZACIÓN DEL sRNA GSUSR06.....	48
FIGURA 16. CARACTERIZACIÓN DE LOS sRNA GSUS45, GSUS32 Y GSUSR34.....	50
FIGURA 17. ENSAYOS DE CO-TRANSCRIPCIÓN DE LOS sRNA GSUSR06, GSUSR17, GSUSR40 Y GSUSR41.....	51
FIGURA 18. ALINEAMIENTO DE SECUENCIAS MÚLTIPLES DEL TRANSCRITO GSUSR27	53
FIGURA 19. ALINEAMIENTO DE SECUENCIAS MÚLTIPLES DEL TRANSCRITO GSUSR17.....	58

ÍNDICE DE TABLAS

TABLA 1. NÚMERO DE LECTURAS Y POSICIONES ANTES Y DESPUÉS DEL FILTRO	23
TABLA 2. ESTADÍSTICAS GENERALES DE LAS 6 BIBLIOTECAS ANALIZADAS.	31
TABLA 3. VALORES DE CORRELACIÓN OBTENIDOS DE LAS 6 BIBLIOTECAS SECUENCIADAS	31
TABLA 4. LISTA DE CITOCROMOS DIFERENCIALMENTE EXPRESADOS.....	35
TABLA 5. LISTA DE sRNAs IDENTIFICADOS EN REGIONES INTERGÉNICAS.	43
TABLA 6. LISTA DE MICROORGANISMOS DONDE SE OBSERVÓ CONSERVACIÓN DE SECUENCIAS CON LOS sRNAs IDENTIFICADOS.....	54
TABLA 7. LISTA DE PREDICCIONES DE POSIBLES BLANCOS DE LOS sRNAs IDENTIFICADOS.....	59

ABREVIATURAS

5'RACE: Rapid Amplification of 5' cDNA ends

5'UTR: 5' untranslated region

asRNA: *anti sense* RNA

RNA: Ribonucleic Acid (ácido ribonucleico)

nt: Nucleótido

DNA: Deoxyribonucleic Acid (ácido desoxirribonucleico)

PCR : Polymerase Chain Reaction

q-PCR: Quantitative PCR

RT-PCR : Reverse transcription PCR

RNAseq: RNA sequencing

sRNA: Small RNA (ARN pequeño)

TSSs: Transcriptional Start Sites

RESUMEN

Geobacter sulfurreducens es una bacteria anaeróbica del suelo con un papel relevante en los ciclos biogeoquímicos del Fe y Mn. Además, posee características metabólicas y fisiológicas que han servido para el desarrollo de aplicaciones biotecnológicas, principalmente en la producción de bioelectricidad y biorremediación de suelos contaminados con metales pesados. Diversos trabajos se han enfocado a estudiar aspectos de regulación genética implicados en estos procesos; sin embargo, estudios dirigidos a la identificación de small RNAs (sRNA) en este organismo aún no han sido reportados. El principal objetivo de este trabajo fue la identificación de sRNA en *G. sulfurreducens* durante la reducción de Fe(III) y fumarato utilizando herramientas de secuenciación masiva. En una primera aproximación, se realizó un mapeo global de sitios de inicio de la transcripción mediante el cual fueron identificados 2479 y 1247 sitios durante la reducción de Fe(III) y fumarato, respectivamente. Posteriormente, se realizó un análisis del transcriptoma mediante RNA-seq. Con los datos obtenidos se buscaron transcritos localizados en regiones intergénicas que pudieran corresponder a sRNA. Con esta aproximación fueron detectados 45 transcritos, los cuales fueron considerados como candidatos a sRNA. Posteriormente fueron seleccionados 12 candidatos para su validación mediante ensayos de q-PCR, 5'RACE y northern blot. Se observó que de los 45 candidatos a sRNAs, 5 presentaron expresión diferencial durante la reducción de Fe(III) y durante la reducción de fumarato. Además, se observó que en la mayoría de los casos, los sRNAs identificados no presentan secuencias homólogas mas allá del género *Geobacter*. Los resultados obtenidos en este trabajo presentan evidencia de la expresión de sRNA durante la reducción de fumarato y Fe(III) en *G. Sulfurreducens*, lo cual representa una contribución significativa para estudios futuros de RNA reguladores en otras especies de *Geobacter*.

ABSTRACT

Geobacter sulfurreducens is an anaerobic soil bacterium with a relevant role in the biogeochemical cycles of Fe and Mn. In addition, *G. sulfurreducens* has metabolic and physiological characteristics that have served for the development of biotechnological applications, mainly in the production of bioelectricity and bioremediation of soils contaminated with heavy metals. Several studies have focused on the genetic regulation involved in these processes; however, studies aimed at the identification of sRNAs in this organism have not yet been reported. The main goal of this work was the identification of sRNA in *G. sulfurreducens* during the reduction of Fe (III) and fumarate using massive sequencing tools. In a first approximation, a global mapping of transcription initiation sites was carried out through which 2479 and 1247 sites were identified during the reduction of Fe (III) and fumarate, respectively. Subsequently, a transcriptome analysis was performed by RNA-seq. With the obtained data, we searched for transcripts located in intergenic regions that could correspond to sRNA. With this approach, 45 transcripts were detected, which were considered as candidates for sRNA. Subsequently, 12 candidates were selected for validation using q-PCR, 5'RACE and northern blot assays. It was observed that of the 45 candidates for sRNA, 5 presented differential expression during the reduction of Fe (III) and 1 during the reduction of fumarate. Furthermore, it was observed that most of the identified sequences are rarely conserved beyond the genus *Geobacter*. The results obtained in this work present evidence of the expression of sRNA during the reduction of fumarate and Fe (III) in *G. sulfurreducens* as well as provide a significant contribution for future studies of regulatory RNAs in other species of *Geobacter*.

CAPÍTULO I

Introducción

1.1 Características generales de las bacterias pertenecientes al género *Geobacter*.

Los microorganismos pertenecientes al género *Geobacter* son Delta-Proteobacterias anaeróbicas que se encuentran ampliamente distribuidas en una gran diversidad de ambientes sedimentarios (Coates et al. 1996). Las bacterias del género *Geobacter* son las más abundantes en ambientes donde se lleva a cabo la reducción de Fe(III) y Mn(IV), participando activamente en los ciclos biogeoquímicos de estos elementos (Melton et al. 2014; Lovley et al. 2004).

Además de llevar a cabo la reducción de óxidos de Fe(III) y Mn(IV), las especies de *Geobacter* son capaces de reducir otros iones metálicos, como por ejemplo U(VI) soluble, TC(VII) y Co(III) (revisado en: Lovley et al. 2011). Estos iones metálicos son reducidos por las especies de *Geobacter* utilizando una amplia gama de donadores de electrones, como son acetato, ácidos grasos de cadena corta, hidrógeno, ácidos húmicos y algunos compuestos aromáticos (Caccavo et al. 1994). Esta capacidad de reducir metales, conjuntamente con la degradación de compuestos aromáticos ha hecho que las especies de *Geobacter* sean utilizadas en el desarrollo de estrategias de biorremediación de sitios contaminados con metales pesados o con hidrocarburos aromáticos (Lovley, 2003).

Otra característica interesante de las especies de *Geobacter* es su capacidad de transferir electrones extracelularmente hacia electrodos para generar energía eléctrica (Bond and Lovley 2003), lo que ha impulsado al desarrollo de dispositivos para la generación de bioelectricidad utilizando cultivos de *Geobacter* (Lovley 2006; Lovley 2012). En este sentido, *Geobacter sulfurreducens* es el organismo modelo de este género bacteriano para el estudio del metabolismo, fisiología, transferencia de electrones y regulación genética (Lovley et al., 2011); debido a que es cultivable, fue la primera especie de este género con el genoma secuenciado (Methé et al. 2003) y cuenta con un sistema de manipulación genética (Coppi et al. 2001).

1.2 Transferencia extracelular de electrones a óxidos de Fe(III) en *Geobacter sulfurreducens*.

En ambientes anaerobios del suelo, las especies de *Geobacter* son capaces de transferir los electrones producidos de la oxidación de compuestos orgánicos al ambiente extracelular donde se localizan los iones metálicos. *G. sulfurreducens* utiliza el acetato como donador de electrones preferentemente, el cual se oxida completamente por medio de reacciones del ciclo de Krebs (Galushko and Schink 2000).

Los electrones producidos son transferidos a través de la membrana interna vía las enzimas NADH deshidrogenasa y succinato-deshidrogenasa y posteriormente estos electrones se transfieren del periplasma a la membrana externa a través de citocromos tipo *c* y con la ayuda de pili, el cual es eléctricamente conductivo (Malvankar et al. 2011), se hace contacto con los óxidos de Fe(III) para su reducción como se indica en la Figura 1 (Mahadevan et al. 2006; Malvankar and Lovley 2012; Butler et al. 2010)

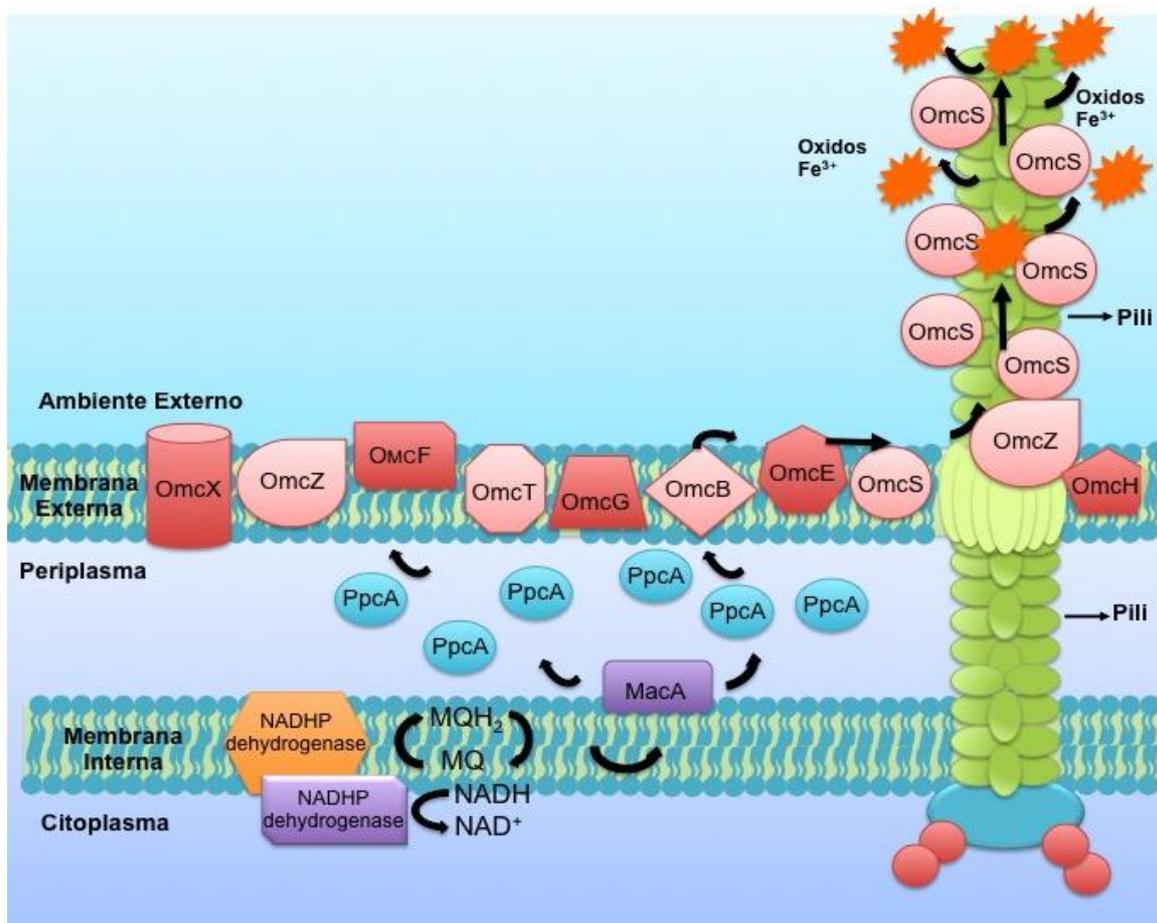


Figura 1. Modelo propuesto de la transferencia de electrones a los óxidos de Fe(III) en *G. sulfurreducens* Modificado de Lovley (2006).

G. sulfurreducens posee 89 citocromos codificados en su genoma (Butler et al. 2010), de los cuales, OmcS es el citocromo más abundante durante la reducción de Fe(III) insoluble (Ding et al. 2008) y se encuentra localizado tanto en la membrana externa como a lo largo de los pili (Leang et al. 2010). La capacidad de transferir electrones con su entorno extracelular es una de las características más interesantes del género *Geobacter* y es el mecanismo clave involucrado en las aplicaciones biotecnológicas en las que es utilizado (revisado en Lovley et al. 2011). La transferencia extracelular de electrones en *G. sulfurreducens* ha sido ampliamente estudiada evaluando la respuesta metabólica de diferentes donadores y aceptores de electrones. Sin embargo, para un mejor entendimiento del mecanismo de transferencia de electrones es necesario ampliar nuestro conocimiento acerca del mecanismo de regulación, así como de los elementos que participan en dicho mecanismo.

1.3 Regulación genética de la transferencia de electrones en *G. sulfurreducens*.

A la fecha, se han identificado en *G. sulfurreducens* algunos elementos de regulación genética, tales como factores sigma alternativos (Núñez et al., 2004), el regulador pilR (Juárez et al. 2009), sistemas de dos componentes (Hernández-Eligio et al. 2016), riboswitches (Tremblay et al. 2011; Kellenberger et al. 2015) y algunos reguladores globales como el caso de Fur (Embree, Qiu, Shieu, Nagarajan, O'Neil, et al. 2014). Además, se han realizado estudios dirigidos a caracterizar la organización estructural de genoma de *G. sulfurreducens*, en los cuales fue posible la identificación de algunos sitios de inicio de la transcripción (TSSs por sus siglas en inglés) en regiones intergénicas que posiblemente correspondan a sRNAs (Qiu et al., 2010). Sin embargo, no se han llevado a cabo estudios dirigidos específicamente a la identificación y caracterización de sRNAs presentes en el genoma de bacterias pertenecientes al género *Geobacter*.

1.4 sRNAs bacterianos y sus mecanismos de regulación.

En años recientes se ha demostrado que los sRNAs juegan un papel clave en los circuitos de regulación en los tres dominios de la vida (Michaux et al., 2014). Particularmente en las bacterias, los sRNAs están involucrados en la regulación genética de diversos procesos fisiológicos y metabólicos (Bobrovskyy and Vanderpool 2013), tales como *quorum sensing* (Bardill et al., 2011), virulencia (Kröger et al. 2012) y adquisición del Fe

(revisado en Oglesby-Sherrouse & Murphy, 2013), entre otros. Los sRNAs actúan en respuesta a señales ambientales y su mecanismo de acción se da a nivel transcripcional o post-transcripcional mediante la unión complementaria con un gen blanco, ya sea incrementando o reprimiendo su expresión (Waters and Storz 2009).

Este grupo de RNAs pueden ser clasificados dependiendo de su localización genómica en tres categorías: *cis*, *trans* y *antisentido* (Zorgani, Quentin, and Lartigue 2016). Los sRNAs tipo *cis* (Figura 2) usualmente se encuentran codificados en las regiones 5' UTR de los mRNA que regulan (Mellin and Cossart 2015). Dentro de este grupo se encuentran los riboswitches, los cuales se encargan de regular la expresión del gen que se encuentra río abajo (Mellin and Cossart 2015).

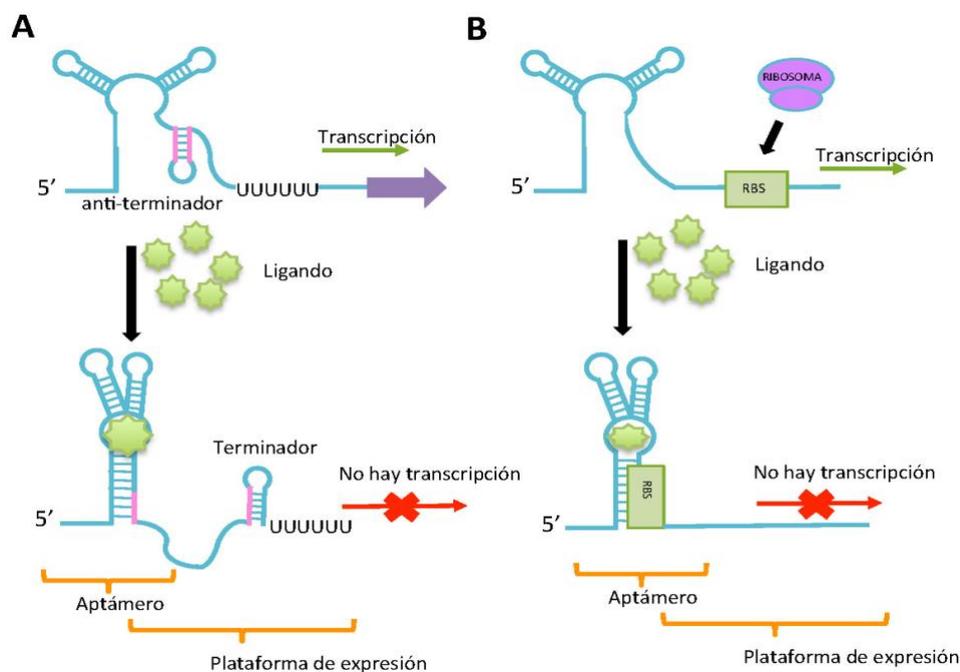


Figura 2. Mecanismos de regulación de un Riboswitch. A) Terminación de la transcripción: un ligando se une a la región del aptámero y forma una estructura de horquilla que abre el bucle anti-terminación y ayuda en la formación del bucle de terminación. B) Bloqueo del inicio de la traducción: la unión al ligando conduce a la formación de una horquilla que bloquea el RBS (sitio de unión a ribosoma) el cual no está disponible para iniciar la traducción.

El segundo grupo corresponde a sRNAs del tipo *trans* (Figura 3), los cuales se encuentran codificados en sitios alejados de su mRNA blanco. Estos sRNAs presentan secuencias que son parcialmente complementarias al gen blanco (Bobrovskyy and Vanderpool 2013) con una región de unión típicamente de 10 a 25 nucleótidos (Peer &

Margalit, 2011). Se ha reportado que este tipo de sRNAs requieren de la interacción con proteínas chaperonas (por ejemplo Hfq en *E. coli* y bacterias relacionadas) que promueven la unión con su mRNA blanco (Vogel and Luisi 2011).

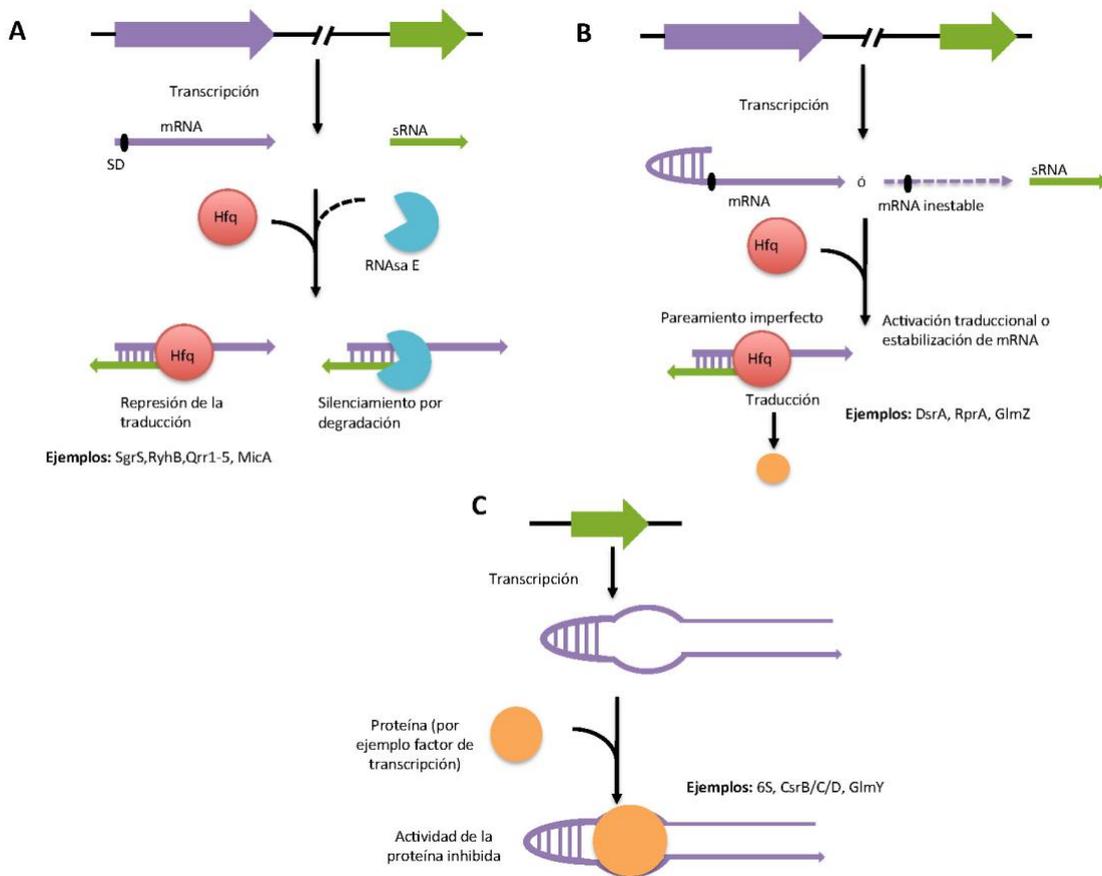


Figura 3. Diferentes tipos de sRNAs basados en su mecanismo de acción. A) Represión B) Activación de la traducción y C) Interacción con proteínas. Imagen modificada de (Liu and Camilli 2010).

Por último, los RNAs *antisentido* (asRNA) son aquellos que se transcriben en la hebra opuesta de su blanco, uniéndose con complementariedad total (Figura 4). Presentan un tamaño aproximado entre los 100 y 700 nucleótidos y son capaces de regular la expresión de su gen blanco a nivel transcripcional, traduccional o tener efectos en la estabilidad del mRNA blanco (Thomason and Storz 2010; Georg and Hess 2011).

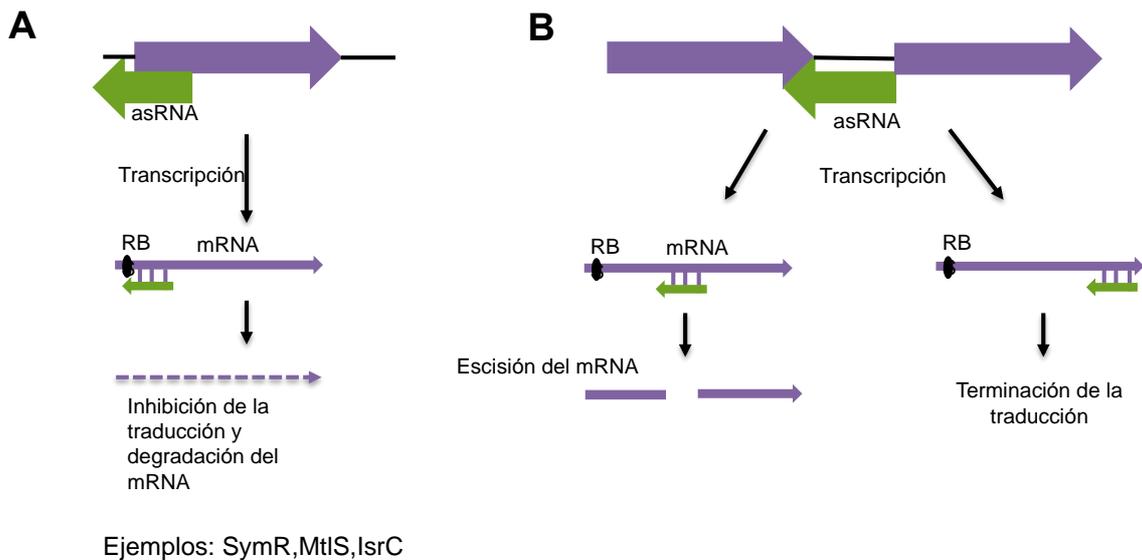


Figura 4. Mecanismos de regulación de un RNA anti-sense. A) asRNA codificado en la hebra opuesta a la región 5' UTR de su mRNA blanco que inhibe el sitio de unión a ribosoma (RBS). B) El asRNA se une a su mRNA separando dos genes de un operón ó propiciando la terminación de la transcripción. Imagen modificada de (Waters and Storz 2009).

Debido al creciente interés en el estudio de los sRNAs en años recientes se han desarrollado diferentes estrategias para su identificación. La identificación *de novo* de sRNAs se ha realizado principalmente mediante herramientas bioinformáticas basadas en la búsqueda de homólogos y patrones de conservación. Sin embargo, estas aproximaciones tienen una alta tasa de falsos positivos (Richter and Backofen 2012). Actualmente, técnicas de RNA-seq han demostrado ser herramientas poderosas para la identificación *de novo* de sRNAs (Sorek & Cossart, 2010; Sharma & Vogel, 2014), dando como resultado un incremento en la identificación de nuevos sRNAs en diversos grupos bacterianos.

El presente trabajo estuvo dirigido a la identificación de sRNAs en *G. sulfurreducens* utilizando dos condiciones de crecimiento: fumarato y Fe(II) como aceptor final de electrones. En una primera aproximación se llevó a cabo un mapeo global de sitios de inicio de la transcripción. Posteriormente, se realizó un análisis del transcriptoma mediante RNAseq, utilizando dos diferentes condiciones de cultivo para poder identificar transcritos en regiones no codificantes que pudieran corresponder a sRNAs, los cuales podrían estar involucrados en la transferencia de electrones. Con estas aproximaciones pudimos identificar 45 sRNA expresados durante las condiciones de crecimiento estudiadas. Nuestros resultados presentan evidencia de la expresión de sRNA durante la

reducción de fumarato y Fe(III) en *G. sulfurreducens*, lo cual representa una contribución significativa para estudios futuros de RNA reguladores en otras especies del género *Geobacter*.

CAPÍTULO II

Hipótesis y Objetivos

2.1 Hipótesis.

El análisis del transcriptoma de *Geobacter sulfurreducens* durante la reducción de Fe(III) soluble y fumarato permitirá la identificación de sRNAs codificados en regiones intergénicas y expresados diferencialmente durante estas dos condiciones de crecimiento.

2.2 Objetivos.

Objetivo General.

Identificación global y validación experimental de sRNAs expresados en *Geobacter sulfurreducens* durante la reducción de Fe(II) y fumarato.

Objetivos Específicos.

- Mapeo global de sitios de inicio de la transcripción durante la reducción de fumarato y Fe(III).
- Análisis del transcriptoma mediante la metodología de RNAseq durante la reducción de fumarato y Fe(III).
- Identificar candidatos a genes que codifican sRNA en regiones intergénicas mediante el análisis de los datos del transcriptoma.
- Validación de los candidatos a sRNA identificados.

CAPÍTULO III

Mapeo Global de Sitios de Inicio de la Transcripción.

3.1 Materiales y Métodos.

3.1.1 Cepas Bacterianas y condiciones de cultivo.

La cepa silvestre de *G. sulfurreducens* PCA (ATCC51573) (Caccavo et al. 1994) fue cultivada rutinariamente en condiciones anaeróbicas a 30° C utilizando acetato 10 mM como donador de electrones y fumarato (40mM) o Citrato Férrico (50mM) como aceptor final de electrones. El crecimiento celular fue determinado por densidad óptica en los cultivos con fumarato y en los cultivos con citrato férrico se midió la reducción de Fe(III) mediante su reacción con ferrozina (Lovley and Philips 1988). Los cultivos fueron cosechados en fase exponencial de crecimiento (Coppi et al. 2001); posteriormente se les añadió RNA *Later* (Ambion) para preservar la integridad del RNA y fueron almacenados a -80 °C.

3.1.2 Extracción y preparación de RNA.

El RNA total fue extraído utilizando las columnas de RNAeasy (QIAGEN Inc.). El RNA fue tratado para la eliminación del DNA residual utilizando DNaseI (Fermentas, Life technologies), utilizando el siguiente protocolo:

RNA	1 µg
10X reaction buffer with MgCl ₂	1 µl
DNase I, RNase-free	1 µl (1 u)
DEPC-treated Water	7 µl

La mezcla fue incubada a 37° C durante 30 min y posteriormente se le añadió 1 µl de 50 mM de EDTA e incubada a 60° C por 10 min para detener la reacción. Finalmente, el RNA fue limpiado y concentrado con las mini columnas del kit RNeasy (QIAGEN Inc.).

3.1.3 Construcción de bibliotecas para mapeo de sitios de inicio de la transcripción.

Con la finalidad de eliminar la mayor cantidad de transcritos que correspondieran a

mRNA degradados o procesados, se realizaron dos diferentes protocolos de eliminación de esta clase de RNA. En un primer paso, los rRNAs fueron removidos utilizando el kit de RiboMinus (Invitrogen), siguiendo las instrucciones del fabricante.

El segundo paso de enriquecimiento de mRNA se llevó a cabo siguiendo el protocolo reportado por Salgado et al. (2013), el cual consiste en utilizar un adaptador biotinilado de RNA que se liga a los mRNA con extremos 5' P, los cuales corresponden a mRNA degradados. Esta ligación se realizó utilizando la enzima T4 RNA ligase (New England Bio Labs) y una vez que el adaptador biotinilado se unió a los mRNA con extremos 5' P estos fueron recuperados utilizando perlas magnéticas del kit de RiboMinus (Invitrogen).

El mRNA enriquecido fue digerido con la enzima Tobacco Acid Pyrophosphatase (Epicenter, Madison, WI, USA) para convertirlo a mRNA 5' P y realizar la transcripción reversa a cDNA utilizando la enzima SuperScript II Reverse Transcriptase (INVITROGEN, Carlsbad, CA, USA). Una vez sintetizado el cDNA se realizó un PCR de 10 ciclos para generar DNA de doble cadena. El producto obtenido fue visualizado en geles de agarosa al 1% y los fragmentos correspondientes a un tamaño de 100-300 nt fueron purificados y entregados a la Unidad Universitaria de Secuenciación Masiva (UNAM) para su posterior tratamiento y secuenciación utilizando el equipo Illumina Genome Analyzer Iix (36 bp single-read run).

3.1.4 Análisis bioinformático.

A partir de los datos crudos, fueron removidas las lecturas con baja calidad (calidad media Phred por debajo de 20). Las lecturas obtenidas fueron mapeadas al genoma de referencia de *G. sulfurreducens* (NC_002939) utilizando el programa Bowtie2 (Langmead and Salzberg 2012) con los siguientes parámetros: 0 mismatches en la longitud de la semilla, longitud de semilla 29 nt, tomando el mejor alineamiento. así como las lecturas correspondientes a rRNA utilizando código en lenguaje Perl. Posteriormente, para cada experimento se calculó la media del número de lecturas por posición y sólo las posiciones que tuvieran un número de lecturas igual o mayor a la media fueron consideradas para los análisis posteriores.

Debido a que algunos TSSs presentaban lecturas en posiciones adyacentes, se formaron grupos "clusters" y fueron considerados como un solo TSS. Después de este agrupamiento, los TSSs fueron asignados a seis diferentes categorías dependiendo de su localización con respecto al inicio de los genes anotados en el genoma:

- a) UpSense: -150 pb a 50 pb
- b) UpSenseExtended: 150 pb río arriba del inicio del gen
- c) InSense: dentro de la región codificante del gen
- d) UpAntiSense: en la región río arriba del gen, pero en la hebra opuesta
- e) UpAntiSenseExtended: 150 pb río arriba del gen, pero en la hebra opuesta
- f) InAntisense: dentro de la región codificante, pero en la hebra opuesta
- e) Convergente: en medio de dos genes que convergen

3.2 Resultados y Discusión.

3.2.1 Descripción general de los TSSs identificados.

En la primera parte de esta tesis y como una primera aproximación para la identificación de sRNAs en *G. sulfurreducens*, se realizó un mapeo global de extremos 5' para la identificación sitios de inicio de la transcripción (TSSs) bajo dos condiciones de crecimiento utilizando secuenciación masiva.

Una vez que las lecturas fueron alineadas al genoma de referencia y filtradas, se observó que un alto número de posiciones presentó un bajo número de lecturas. Con la finalidad de tener un valor de corte para estas posiciones, se calculó la media de lecturas obtenidas por posición. Las posiciones que presentaron lecturas por debajo de la media fueron eliminadas. Una vez realizado este filtro, se pudo observar que el número de inicios detectados disminuyó drásticamente, mientras que el número de lecturas totales solo disminuyó aproximadamente un 12% (Tabla 1), lo cual sugiere que las posiciones eliminadas con un bajo número de lecturas pudieron corresponder a “ruido transcripcional”¹ (Wade and Grainger 2014).

¹ Wade & Grainger (2014) Definen el ruido transcripcional como la detección de sitios de inicio de transcripción en sitios no esperados (principalmente en posición anti sentido a regiones codificantes) que surgen de regiones promotoras espurias.

Tabla 1. Número de lecturas y posiciones antes y después del filtro para cada una de las librerías secuenciadas.

	Fumarato	Fe(III) soluble
Total de TSSs	28,568	49,298
Total de lecturas alineadas	2,065,503	1,793,119
TSSs después del filtro	1,929	3,463
Lecturas alineadas después del filtro	1,835,164	1,567,135
Porcentaje de lecturas conservadas después del filtro	89%	87%

Los TSSs identificados en ambas condiciones de crecimiento se encontraron distribuidos a lo largo del genoma y fueron asociados a 2196 genes de los 3711 que están anotados para *G. sulfurreducens* PCA (Figura 5).

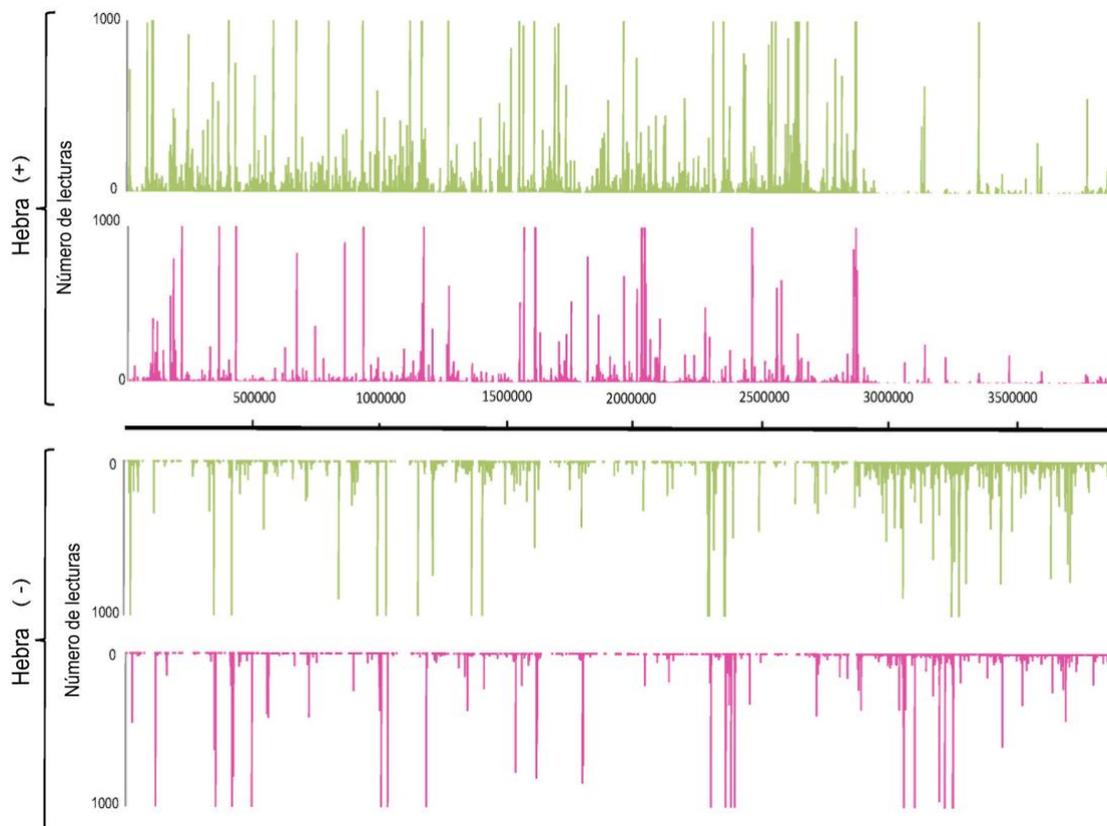


Figura 5. Sitios de Inicio de la Transcripción (TSSs) detectados en las dos condiciones de crecimiento. El eje de las X corresponde a las coordenadas genómicas de *G. sulfurreducens* y el eje de las Y al número de

lecturas por posición. Las barras rosas corresponden a los TSSs detectados durante la reducción de fumarato y las barras verdes a los detectados durante la reducción de Fe(III).

Se pudo observar la presencia de múltiples inicios de transcripción en posiciones adyacentes. Para estos casos, los inicios de transcripción detectados en posiciones continuas fueron agrupados en "clusters", los cuales fueron considerados como un solo TSS. Posterior a este agrupamiento, se obtuvieron 2479 y 1247 "clusters" durante la reducción de Fe(III) y fumarato respectivamente, de los cuales 389 fueron detectados en ambas condiciones de crecimiento (Figura 6A). Cabe destacar que los TSSs con mayor número de lecturas corresponden a TSSs que se encuentran presentes en ambas condiciones de cultivo (Figura 6B).

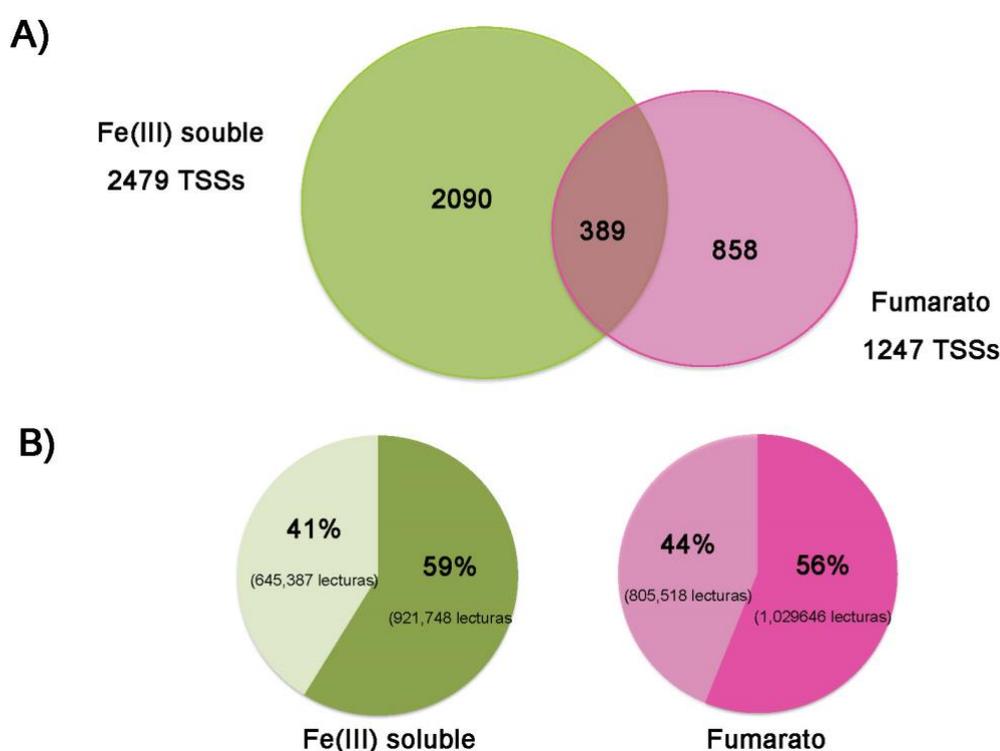


Figura 6. Agrupación de los sitios de inicio de la transcripción. A) Diagrama de Venn de los TSSs agrupados en clusters. B) Porcentaje de lecturas obtenidas en cada experimento. Los tonos claros corresponden a las lecturas de los TSSs detectados en una sola condición de cultivo y los colores fuertes a aquellos TSSs detectados en ambas condiciones de crecimiento.

En un trabajo previo, Qiu y colaboradores (2010) reportaron la identificación de 1374 TSSs durante la reducción de fumarato. Al buscar si existían coincidencias entre nuestros datos y los previamente reportados, encontramos que solo 26 TSSs coincidían exactamente en la misma posición. Sin embargo, cuando ampliamos la búsqueda de coincidencias en posiciones adyacentes encontramos que este número incrementaba a 603 posiciones (Figura 7), lo cual sugiere que el mapeo global de TSSs tiende a presentar

pequeñas variaciones en las posiciones cromosómicas que detecta, las cuales probablemente son ocasionados por las diferentes metodologías de secuenciación empleadas, las instalaciones donde se realizan los experimentos o inclusive por el tipo de cepa utilizada. Además de las coincidencias detectadas en los TSSs, buscamos si existían promotores asociados a estas posiciones y encontramos que promotores del tipo RpoN y RpoD, reportados en Qiu et al (2010), estaban asociados a 668 de los TSSs reportados en nuestro trabajo. En este sentido, se ha demostrado que un TSS puede considerarse con "evidencia fuerte" cuando su detección esta respaldada por al menos dos tipos de evidencia o experimentos independientes (Salgado et al. 2013; Weiss et al. 2013), por lo que los TSSs identificados en este trabajo y que también fueron detectados en el trabajo de Qiu et al (2010) pueden ser considerados como "evidencia fuerte" (aunque en ambos casos no se realizaron replicas), fueron detectados en experimentos independientes.

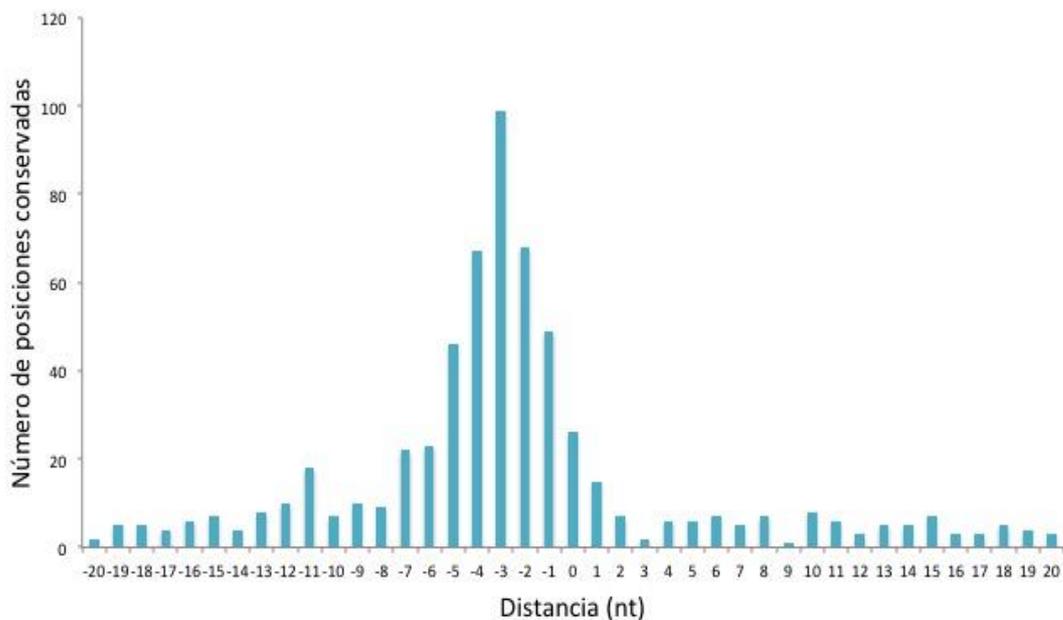


Figura 7. Número de coincidencias por posición con respecto a los datos reportados por Qui et al (2010). Se realizó una búsqueda de coincidencias en la posición exacta y en un rango de ± 1 a ± 10 nt con respecto a cada posición reportada.

3.2.2 Categorización de los TSS identificados.

Todos los TSSs identificados fueron asignados a una categoría en función de su localización genómica (Figura 8) como se menciona en el apartado 3.1.4 de materiales y métodos. Cuando se realizó la asignación de categorías a los TSSs presentes en los datos crudos, es decir antes de realizar el corte por debajo de la media, pudimos observar que

más del 50% de estos TSSs se localizaban en la región InSense. Posterior al corte, este porcentaje disminuyó considerablemente, lo cual sugiere que estas posiciones poseen bajo número de lecturas debido a que corresponden a transcritos espurios. Por otro lado, se observó que en las dos condiciones de crecimiento la mayoría de los TSSs fueron localizados en la región UpSense, que es la región donde se localizan lo TSSs canónicos. Esta misma tendencia estuvo presente en los 389 TSSs que coincidieron en ambas condiciones de crecimiento.

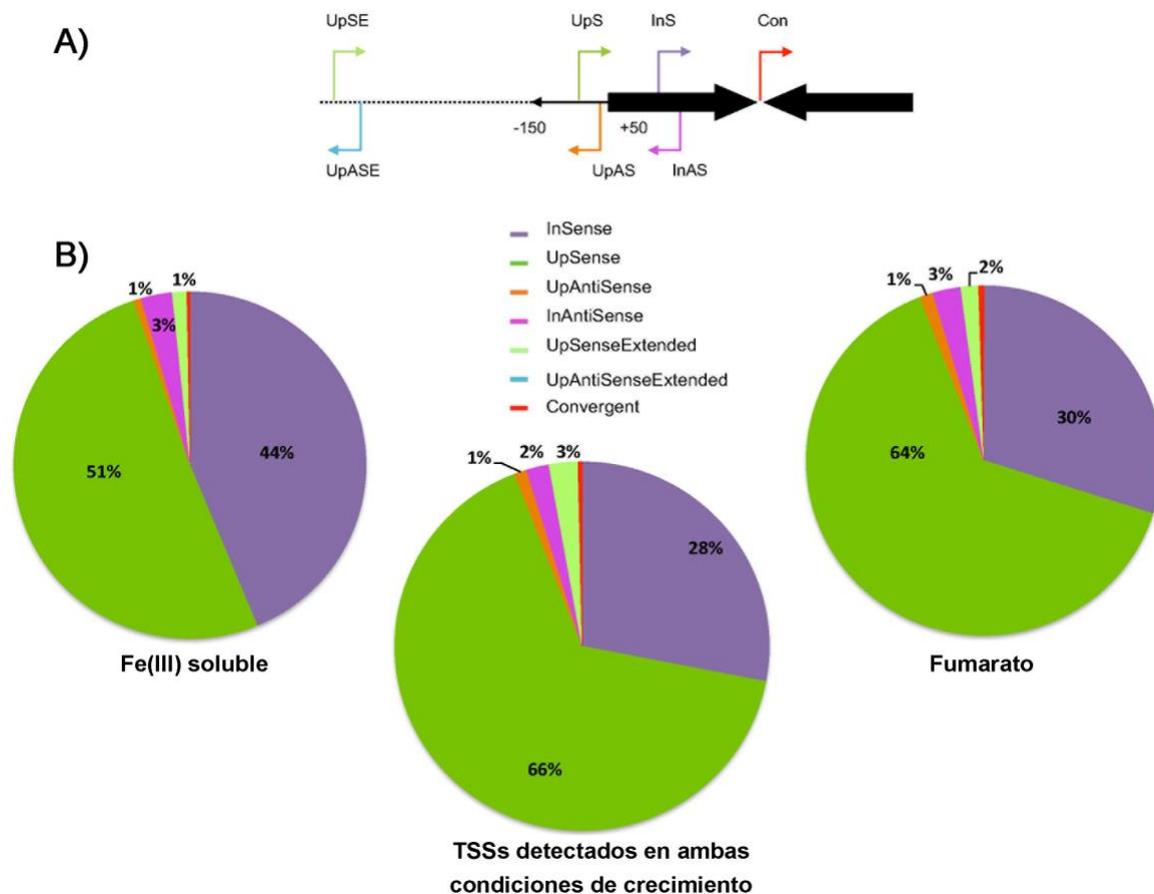


Figura 8. Categorización de TSSs con respecto a su localización genómica. A) Representación esquemática de las categorías utilizadas en este trabajo. B) Porcentajes de TSSs asignados a las diferentes categorías en las dos condiciones de crecimiento.

En años recientes se ha reportado que la transcripción antisentido es abundante en los transcriptomas bacterianos (Lasa and Toledo-Arana 2011; Raghavan, Sloan and Ochman 2012; Thomason et al. 2014). Sin embargo, el significado biológico de este tipo de transcritos es aún poco claro y algunos autores proponen que este tipo de transcripción es poco reproducible a través de diferentes experimentos (Raghavan, Sloan, and Ochman

2012; Lloréns-Rico et al. 2016). Mientras que para otros organismos como *Helicobacter pylori* se ha reportado que el 27% de sus transcritos corresponden a esta categoría (Sharma et al. 2010), nuestros datos indicaron que en *G. sulfurreducens* solo el 4% del total de TSSs corresponden a la categoría antisentido y cuando se consideran los TSSs detectados en ambas condiciones de crecimiento, esta cantidad disminuye 1% (Figura 8). Adicionalmente, se observó que 4 TSSs de esta categoría presentan regiones promotoras reportadas en trabajos anteriores (Qiu et al. 2010), y aunque el número es muy bajo, estos TSSs son fuertes candidatos para estudios posteriores de transcripción antisentido en este organismo modelo.

Los TSSs localizados en regiones intergénicas corresponden a las categorías Convergente y UpSenseExtendend. Para estas dos categorías fueron identificados 40 TSSs durante la reducción de Fe(III), 26 cuando se utilizó fumarato como aceptor de electrones y 16 en ambas condiciones de crecimiento. Estos TSSs representan una primera aproximación para la subsecuente búsqueda de candidatos a sRNA del tipo *trans*.

3.2.3 Búsqueda de regiones promotoras.

Se realizó una búsqueda de regiones promotoras asociadas a los TSSs identificados. Para este análisis se utilizaron regiones promotoras reportadas previamente, las cuales fueron identificadas mediante predicciones bioinformáticas (Yan et al. 2006) y por experimentos de ChIP-chip (Qiu et al. 2010). Con estos datos encontramos que 668 de los TSSs poseen al menos una región promotora para RpoN, RpoS ó RpoD.

En algunos casos fue posible asignar regiones promotoras a genes que presentaron múltiples TSSs, como fue el caso del gen *ppcD*, para el cual Yan y colaboradores (2006) habían reportado tres promotores putativos RpoS. Nuestros resultados indicaron la presencia de dos TSSs asociados a estas regiones promotoras. Se observó que uno de ellos fue detectado durante ambas condiciones de crecimiento y el otro exclusivamente durante la reducción de fumarato (Figura 9). Este resultado sugiere que la expresión de *ppcD* es controlada por diferentes TSSs dependiendo de la condición de cultivo.

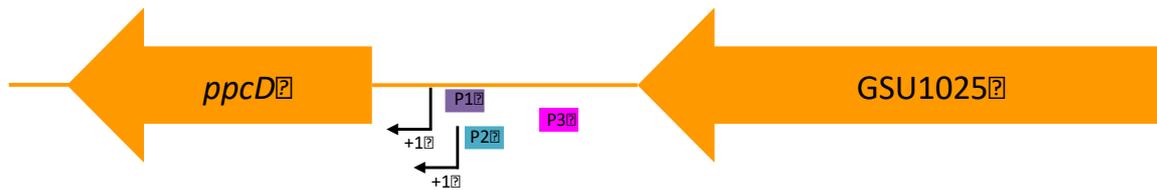


Figura 9. Localización genómica del gen *ppcD*: las regiones promotoras identificadas por Yan et al. 2006 corresponden a las líneas verde azul y rosa. Las flechas negras indican los dos sitios de inicio de la transcripción identificados por el mapeo global de TSSs, los cuales pudieran estar asociados a las regiones promotoras 1 y 2.

Los resultados obtenidos en este capítulo fueron publicados en el siguiente artículo:
 González, G., Labastida, A., Jiménez-Jacinto, V., Vega-Alvarado, L., Olvera, M., Morett, E. and Juárez, K. 2016. **Global transcriptional start site mapping in *Geobacter sulfurreducens* during growth with two different electrons acceptors.** *FEMS Microbiology Letters*, 363(17), fnw175. (Anexo)

CAPÍTULO IV

Análisis de expresión diferencial utilizado Fe(III) y fumarato como aceptores de electrones.

4.1 Materiales y Métodos

4.1.1 Cepas Bacterianas y condiciones de cultivo

La cepa silvestre de *G. sulfurreducens* PCA (ATCC51573) (Caccavo et al. 1994) fue reactivada de un glicerol en medio NBAF (fumarato como aceptor de electrones), posteriormente fue sub-cultivada en medio Citrato Férrico para adaptar las células a esta condición de cultivo. Finalmente, las células fueron cultivadas por triplicado para las dos condiciones de cultivo como se menciona en el apartado 3.1.1 del capítulo anterior.

4.1.2 Construcción de bibliotecas para análisis de RNAseq.

El RNA total fue extraído y tratado como se describe en la sección 3.1.2. Para este experimento se realizó la eliminación de rRNA utilizando el kit de RiboMinus (Invitrogen) siguiendo las instrucciones del fabricante. La concentración y calidad del RNA obtenido fue analizada utilizando el equipo Bioanalyzer 2100, Chip RNAnano (Aligent). Las muestras del RNA obtenido fueron enviadas a la USSM (Unidad Universitaria de Secuenciación Masiva, UNAM, México). Las bibliotecas fueron preparadas utilizando el kit TruSeq RNA Sample Prep Kit (ILLUMINA, San Diego, CA, USA). En total, se sintetizaron 6 librerías (3 por cada condición de cultivo). Las bibliotecas fueron secuenciadas con el equipo Illumina Genome Analyzer IIx con lecturas de tipo pair-end de 72 bp de longitud.

4.1.3 Análisis Bioinformático.

Las lecturas obtenidas fueron mapeadas al genoma de *G. sulfurreducens* utilizando el programa BWA (H. Li and Durbin 2009) con los siguientes parámetros: -l 35 -k 2 -t 2 -o 1. Las lecturas correspondientes a rRNA, duplicados ópticos y de PCR fueron eliminados utilizando códigos en lenguaje Perl. El análisis de la expresión diferencial fue llevado a acabo mediante tres diferentes métodos: DESeq (Anders and Huber 2010), EdgeR (Robinson, McCarthy, and Smyth 2010) y NOISeq (Tarazona et al. 2011). Sólo los genes

diferencialmente expresados detectados por los tres métodos anteriormente mencionados fueron considerados para análisis posteriores. La razón de emplear diferentes métodos de análisis fue poder identificar si dependiendo del tipo de estrategia pueden existir diferencias relevantes en los resultados. Si bien es difícil establecer que método tiene un mejor desempeño, esta comparación nos permitió determinar que métodos poseen mayor similitud.

Los genes con un nivel de cambio con valor de logaritmo base dos ≥ 2 fueron tomados como diferencialmente expresados. Para la asignación de clase funcional de los genes con expresión diferencial se utilizaron las siguientes bases de datos: RAST (http://www.theseed.org/wiki/Main_Page), PANTHER GO (<ftp://ftp.pantherdb.org>), KEGG (<http://www.genome.jp/kegg/pathway.html>) and JCVI (J. Craig Venter Institute). Los datos obtenidos fueron visualizados con la herramienta IGB Browser (Helt et al. 2009)

4.1.4 Validación de los datos de RNAseq por RT-PCR.

Fueron seleccionados 5 de los genes que presentaron los más altos cambios en los niveles de expresión para su validación por RT-qPCR. El RNA fue extraído y tratado como se describió anteriormente, el cDNA fue sintetizado utilizando el kit RevertAid H minus First Strand cDNA Synthesis (Thermo Scientific) siguiendo las instrucciones del fabricante. Para la reacción de PCR se utilizó el kit Maxima SYBR Green/ROX qPCR Master Mix (Thermo Scientific). La lista de los oligonucleótidos utilizados se encuentra en el Anexo I. El gen *RecA* (Holmes et al. 2008) fue utilizado como control interno para la normalización de los resultados. Los ensayos fueron realizados por triplicado en dos experimentos independientes y los datos fueron analizados por el método de $2^{-\Delta\Delta C_T}$ (Livak and Schmittgen 2001).

4.2 Resultados y Discusión.

4.2.1 Estadísticas generales de los datos de RNAseq.

Una vez que se realizó el procesamiento de los datos crudos se obtuvo el número final de lecturas que fueron utilizadas para los análisis subsecuentes (Tabla 3). No obstante que el RNA total utilizado fue tratado para la eliminación de rRNA, se observó que un gran porcentaje de las lecturas correspondieron a este tipo de RNA. Sin embargo, el número final de lecturas es el suficiente para poder detectar y cuantificar expresión diferencial en los transcriptomas bacterianos (Haas et al. 2012).

Tabla 2. Estadísticas generales de las 6 bibliotecas analizadas.

	Fe(III) Rep1	Fe(III) Rep2	Fe(III) Rep3	Fumarato_ Rep1	Fumarato _Rep2	Fumarato_ Rep3
Total de lecturas	35,258,610	32,186,726	41,825,970	37,741,156	48,213,778	38,861,666
Lecturas agrupadas en pares	27,801,434	24,067,012	32,558,430	29,507,202	31,384,550	33,339,626
Lecturas correspondientes a rRNA	22,922,360	20,163,080	28,265,220	20325046	25297034	26,047,730
Duplicados ópticos y de PCR	4,879,074	3,903,932	4,293,210	9182156	6,087,516	7,291,896
Lecturas correspondientes a mRNA	2,961,738	2,375,554	2,399,664	5,440,834	3,939,776	4,196,714

Para verificar la reproducibilidad de los experimentos, fueron calculados los valores de correlación utilizando el método de Spearman (R statistics tools). Los resultados indicaron que las tres réplicas biológicas analizadas para cada condición poseen altos valores de correlación entre ellas (Tabla 3).

Tabla 3. Valores de correlación obtenidos de las 6 bibliotecas secuenciadas

	Fe(III) Replica 1	Fe(III) Replica 2	Fe(III) Replica 3	Fumarato Replica 1	Fumarato Replica 2	Fumarato Replica 3
Fe(III) Replica 1	-	0.9361	0.9612	0.7818	0.7873	0.8436
Fe(III) Replica 2	0.9361	-	0.9830	0.6629	0.6782	0.7311
Fe(III) Replica 3	0.9612	0.9612	-	0.7027	0.7116	0.7650
Fumarato Replica 1	0.7818	0.6629	0.7027	-	0.9864	0.9085
Fumarato Replica 2	0.7873	0.6782	0.7116	0.9864	-	0.9119
Fumarato Replica 3	0.8436	0.7311	0.7650	0.9085	0.9119	-

4.2.2 Expresión diferencial utilizando Fumarato y Fe(III) como aceptor final de electrones.

Las dos condiciones de crecimiento empleadas en este trabajo, donde se utiliza acetato como donador de electrones y fumarato ó Fe (III) soluble como aceptor final de electrones, han sido utilizadas ampliamente en el estudio de *G. sulfurreducens* (Methe et al. 2005; Yang et al. 2010; Embree et al. 2014). El uso de Fe(III) soluble como aceptor final de electrones ha servido como modelo para el estudio del transporte de electrones extra celular en *Geobacter* (Ding et al. 2006; Ueki et al. 2017). Así mismo, el fumarato es utilizado como un aceptor de electrones alternativo que incrementa el crecimiento de *G. sulfurreducens*. El uso de estas dos condiciones de crecimiento ha permitido estudiar cambios en las condiciones metabólicas que permiten el estudio de la expresión diferencial relacionada con la utilización de diferentes aceptores de electrones en *G. sulfurreducens* (Esteve-Núñez et al. 2004).

El análisis de la expresión diferencial de los genes se llevó a cabo utilizando tres métodos independientes y solo los genes detectados por los tres métodos fueron considerados como diferencialmente expresados (Figura 10). Este análisis arrojó 447 genes expresados diferencialmente, 224 sobre expresados durante la reducción de Fe(III) y 223 durante la reducción de fumarato. Los genes seleccionados para la validación de la expresión mediante RT-qPCR mostraron valores de expresión consistentes con los obtenidos por RNAseq (Anexo).

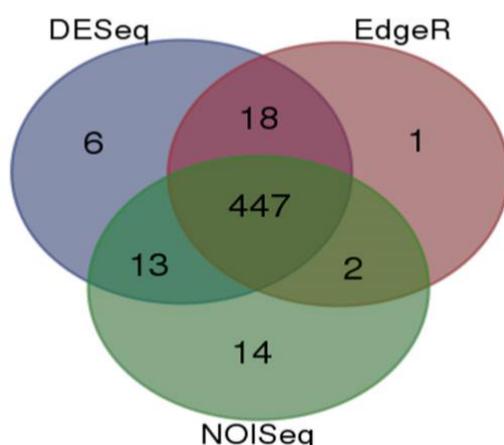


Figura 10. Diagrama de Venn de la intersección de genes diferencialmente expresados detectados por tres diferentes métodos.

En la Figura 11 se muestra los genes expresados diferencialmente, agrupados en categorías funcionales. Como se puede observar, la categoría con mayor número de genes sobre expresados durante la reducción de Fe(III) corresponde a proteínas hipotéticas; en total, 103 genes de esta categoría fueron sobre expresados en esta condición de crecimiento. De hecho, de los genes con mayor tasa de cambio en su expresión se encuentran en este grupo (GSU0919, GSU0597, GSU3410, GSU3409) con valores de nivel de cambio >5 con respecto a los valores de expresión detectados durante la reducción de fumarato. Así mismo, al realizar un análisis BLAST (Altschul et al. 1997) de las secuencias de estas proteínas no se encontraron secuencias homólogas más allá del género *Geobacter*. Los altos valores de expresión de este grupo de genes, así como la ausencia de homólogos sugiere que podrían estar en algún mecanismo relacionado con la reducción de Fe(III) llevada a cabo por microorganismos de este género.

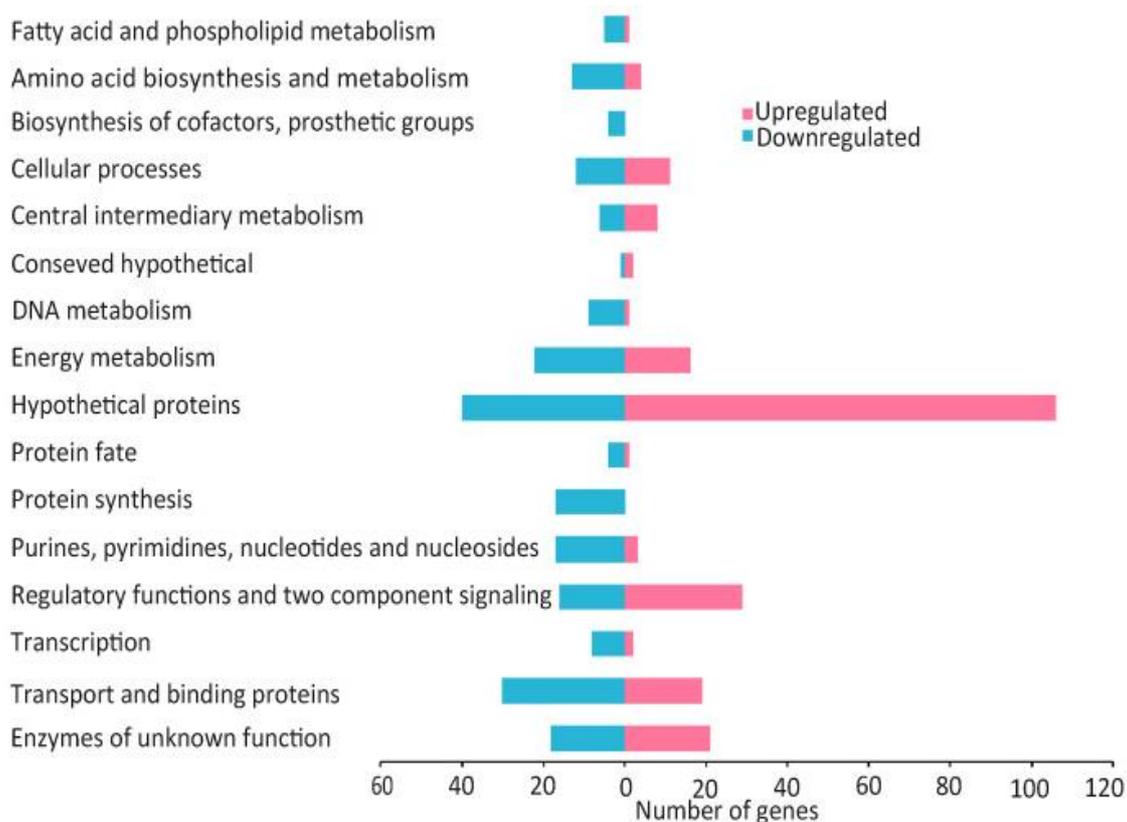


Figura 11. Categorías funcionales de los genes con expresión diferencial. Los cambios en la expresión son con respecto a la utilización de Fe(III) como aceptor final de electrones.

4.2.3 Expresión de genes relacionados con proteínas de unión y transporte.

El genoma de *G. sulfurreducens* tiene 100 genes anotados relacionados a proteínas de unión-transporte de los cuales 49 genes tuvieron expresión diferencial. En total, 30 de estos genes fueron sobre expresados durante la reducción de fumarato y 19 durante la reducción de Fe (III) soluble.

Se observó que 4 genes (GSU3269, GSU3268, GSU3270 y GSU1380) que codifican para proteínas de la familia Feo, relacionadas con el transporte de Fe, fueron sobre expresados durante la reducción de fumarato. Las proteínas de la familia Feo son las encargadas de la adquisición del Fe(II) del ambiente hacia las células (Cartron et al. 2006). En *G. sulfurreducens* se ha reportado que en condiciones de cultivo donde la concentración de Fe(II) es baja, el nivel de expresión de *FeoB* aumenta, mientras que su expresión se ve reprimida al aumentar los niveles de Fe(II) en el medio (O'Neil et al. 2008). Este comportamiento sugiere que la expresión de las 4 proteínas de la familia Feo puede deberse a la baja concentración de Fe(II) durante la reducción de fumarato; mientras que durante la reducción de Fe(III) los niveles de Fe(II) aumentan en el medio de cultivo ocasionando que los niveles de expresión de estos genes disminuyan.

Además, se observó que el gen GSU1307, el cual codifica para una ferritina, mostró sobre expresión durante la reducción de Fe(III). Debido a que las ferritinas están involucradas en el almacenamiento intracelular de Fe, los datos de expresión sugieren que ferritina podría estar siendo empleada por la célula cuando se encuentra expuesta a altas concentraciones de Fe(II).

Por otra parte, 4 genes (GSU1609, GSU1330 GSU1332 GSU1331) que codifican para proteínas de la familia RND (Resistance-Nodulation-Cell) presentaron sobre expresión durante la reducción de Fe(III). Se ha reportado que proteínas pertenecientes a esta familia están involucradas en la resistencia y flujo-transporte de metales pesados en bacterias (Nies 2003). Dado los altos niveles de expresión que presentaron estas proteínas durante la reducción de Fe(III), sería interesante conocer si estas proteínas participan de alguna manera en la reducción de Fe(III).

4.2.4 Expresión de genes asociados a energía y metabolismo.

Dentro de los elementos más importantes que participan en la transferencia extra celular de electrones realizada por las bacterias del género *Geobacter spp* se encuentran los citocromos. Una de las características mas relevantes del género es el gran número de citocromos que se encuentran codificados en sus genomas (Methé et al. 2003; Butler et al. 2010). De los 89 citocromos anotados en el genoma de *G. sulfurreducens* encontramos que 18 presentaron expresión diferencial, 9 fueron sobre expresados durante la reducción de fumarato y 11 durante la reducción de Fe(III) (Tabla 5). Para algunos de estos citocromos (OmcX,OmcA,OmcG y PccD) ya se ha identificado su papel en la transferencia de electrones durante la reducción de Fe(III) (revisado en Lovley et al. 2011, Butler et al. 2010).

Tabla 4. Lista de los citocromos que mostraron expresión diferencial en los experimentos de RNAseq. Los valores positivos corresponden a los genes sobre expresados durante la reducción de Fe(III) y los valores negativos durante la reducción de fumarato.

Gen ID	Nombre del gen	Producto del gen	Nivel de Cambio (Log2)
GSU2643		citocromo <i>c</i>	-2.13
GSU0593		citocromo <i>b</i>	4.76
GSU3274		citocromo <i>c</i>	-7.26
GSU1257		citocromo <i>c</i> y receptor periplásmico	-3.64
GSU0670	omcX	citocromo <i>c</i>	-2.89
GSU0274		citocromo <i>c</i>	-2.32
GSU1640		citocromo <i>bd</i>	-2.28
GSU3615		citocromo <i>c</i>	-2.05
GSU1024	ppcD	citocromo <i>c</i>	2.09
GSU1740		citocromo <i>c</i>	2.11
GSU2501		citocromo <i>c</i>	2.19
GSU2884	omcA	citocromo <i>c</i>	2.55
GSU2882	omcG	citocromo <i>c</i>	2.75
GSU0746		citocromo p460	3.06
GSU3214		citocromo <i>c</i>	3.29
GSU0594		citocromo <i>c</i>	4.44
GSU2642		citocromo <i>c</i>	-2.69
GSU0779		citocromo <i>c</i> , formato deshidrogenasa putativa	-4.14

Se observó que algunos genes que codifican para citocromos tipo *c* tuvieron una elevada sobre expresión durante la reducción de fumarato (GSU1257 y GSU3274). Específicamente estos dos citocromos presentan, a diferencia de los citocromos que son sobre expresados durante la reducción de Fe(III) soluble, solo un sitio de unión hemo. En

este sentido, se ha propuesto que durante condiciones limitantes de hierro (reducción de fumarato) la expresión de genes que contienen múltiples residuos hemo disminuye, mientras que los citocromos con menos "residuos" hemo son sobre expresados para participar en el transporte de electrones durante la escases de hierro (Embree et al. 2014).

En relación a genes involucrados en el metabolismo, observamos que durante la reducción de fumarato fueron sobre expresados dos de los cuatro genes que codifican para las enzimas periplasmáticas formato deshidrogenasas (GSU077 y GSU0778). Se ha descrito que en delta-protobacterias, las enzimas formato deshidrogenasas transfieren electrones hacia citocromos *c3* (Silva, Voordouw, and Leita 2013). Sin embargo, en *G. sulfurreducens* el patrón de expresión de los genes que codifican para estas enzimas solo se han estudiado en co-cultivos con *Pelobacter carbinolicus* (Rotaru et al. 2012) y su papel en la reducción de fumarato aún se desconoce.

Otros genes sobre expresados durante la reducción de Fe(III) soluble fueron dos genes que codifican para monóxido de carbono deshidrogenasas dependientes de níquel (GSU2097 y GSU2098). Aunque *G. sulfurreducens* se desconoce el papel de estas enzimas, se ha reportado que en otras bacterias participan en el metabolismo anaeróbico del CO (Oelgeschläger and Rother 2008). En este sentido, recientemente se reportó que *G. sulfurreducens* puede utilizar CO como donador de electrones y aunque no fue evaluada la actividad enzimática, los autores sugieren que las enzimas que codifican estos dos genes pudieran estar implicadas en la oxidación de CO (Geelhoed et al. 2016).

4.2.5 Expresión diferencial en genes implicados en regulación y sistemas de dos componentes.

Dentro de esta categoría se detectaron 46 genes expresados diferencialmente, de los cuales 29 estuvieron sobre expresados durante la reducción de Fe(III). Estos genes incluyen un regulador transcripcional de la familia ArsR (GSU2149), de la cual se ha reportado que está involucrada en la respuesta a estrés por iones metálicos (Summers 2009). De igual forma presentaron sobre expresión durante la reducción de Fe(III) genes que codifican para reguladores transcripcionales de la familia LuxR (GSU0682) y LysR (GSU1342).

El genoma de *G. sulfurreducens* codifica para 27 genes relacionados con la regulación a través de la molécula señal c-di-GMP (di-guanilato mono cíclico), de los cuales dos tuvieron sobre expresión durante la reducción de Fe(III) (GSU3350 y GSU3356) y sólo uno durante la reducción de fumarato (GSU3263).

Una característica muy interesante de *G. sulfurreducens* es que posee un gran número de sistemas de dos componentes anotados en su genoma: 95 caracterizados como reguladores de respuesta y 22 con ambos dominios (Karlin et al. 2006). Los datos obtenidos indican que 23 de estos genes mostraron expresión diferencial: 11 sobre expresados durante la reducción de fumarato y 12 en la reducción de Fe(III).

Así mismo, *G. sulfurreducens* posee un gran número de genes involucrados en mecanismos de regulación y de respuesta a cambios ambientales. En los datos de RNAseq se observó que un número considerable de estos genes presentaron expresión diferencial (Figura 11), los cuales en su mayoría mostraron sobre-expresión durante la reducción de Fe(III). Estos datos sugieren que el tipo de aceptor final de electrones es un aspecto del metabolismo de *G. sulfurreducens* que se encuentra altamente regulado.

4.2.6. Expresión de genes relacionados con mecanismos de regulación genética a través de sRNAs.

En los mecanismos de regulación genética llevados a cabo por sRNAs participan una amplia variedad de proteínas que incluyen desde chaperonas hasta enzimas con función RNAsas (Gottesman and Storz 2010; Van Assche et al. 2015). Una de las proteínas mas importantes que participan en la regulación negativa mediada por sRNA es la chaperona Hfq, la cual es la encargada de facilitar la unión entre el sRNA y el sitio de unión al ribosoma del mRNA blanco (Vogel and Luisi 2011). *G. sulfurreducens* posee el gen *hfq* que codifica para un homólogo de esta proteína, para el cual se detectó expresión en ambas condiciones de crecimiento, sin presentar expresión diferencial.

Otra proteína de unión a RNA que interactúa con sRNAs para llevar a cabo mecanismos de regulación post-transcripcional es CsrA (revisado en Romeo et al. 2013). En *G.*

sulfurreducens el gen GSU3041 codifica una proteína homóloga a CsrA, sin embargo, sus valores de expresión fueron muy bajos en ambas condiciones de crecimiento. Además de las proteínas anteriormente mencionadas, el genoma de *G. sulfurreducens* contiene otras 5 proteínas, que aunque no han sido caracterizadas, poseen motivos de unión a RNA (Finn et al. 2016). En cuanto al perfil de expresión de estas proteínas, solo se detectó expresión diferencial para tres de ellas (GSU0235, GSU0644 y GSU1587).

CAPITULO V

Identificación y validación de sRNAs en regiones intergénicas

5.1 Materiales y Métodos.

5.1.1 Identificación de sRNAs en regiones intergénicas.

Para la identificación de sRNAs fueron utilizados los datos obtenidos de los experimentos de RNAseq descritos en el capítulo anterior. En un primer paso, fueron seleccionadas todas las regiones intergénicas con un tamaño mayor a 160 pares de bases. Se decidió utilizar este tamaño de región intergénica para minimizar la identificación de transcritos que pudieran corresponder a los genes colindantes y no a la región intergénica. La identificación de señal transcripcional se realizó de manera manual utilizando la herramienta Integrated Genome Browser (Helt et al. 2009). Utilizando las secuencias de estos transcritos, se realizó un análisis de predicción de proteínas utilizando el software GeneMark (Besemer and Borodovsky 2005). Este análisis se llevo a cabo para descartar transcritos que pudieran corresponder a genes de proteínas no anotadas. Las secuencias restantes fueron analizadas para identificar posibles homólogos utilizando las plataformas BLAST (Altschul et al. 1997) y Rfam (Burge et al. 2012).

Los valores de expresión fueron calculados siguiendo el mismo procedimiento descrito en el capítulo anterior, considerando solo los sRNAs que mostraban expresión diferencial en los 3 métodos utilizados. Para los análisis de predicción de estructura secundaria se utilizó la herramienta RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) aplicando los valores estándar que ofrece el programa. Además, se llevó a cabo un análisis de predicción de genes blanco de los sRNAs identificados. Para esta predicción se empleó el programa TargetRNA2 (Kery et al. 2014).

5.1.2 Validación mediante Northern blot.

Este experimento se llevó a cabo para determinar el tamaño aproximado de los transcritos identificados. El RNA fue extraído de cultivos en fase exponencial y estacionaria

utilizando el protocolo descrito en los capítulos III y IV. Se tomaron 10µg del RNA total de cada muestra y fueron cargados en geles desnaturalizantes de poliacrilamida al 10% con una concentración de urea 8M (2 horas a 180 volts). En cada uno de los geles se utilizó como marcador de RNA el ssRNA Ladder (New England BioLabs Inc.) Posteriormente, el RNA desnaturalizado fue transferido a membranas Hybond-N+ (GE Healthcare) mediante electro-blotting y fijado a la membrana por UV-crosslinking. Las sondas de DNA utilizadas para la hibridación fueron generadas mediante PCR utilizando los oligonucleótidos enlistados en el Anexo II. Los productos obtenidos fueron marcados radioactivamente con [α -³²P]dCTP utilizando el kit Rediprime II DNA Labeling System (GE Healthcare Life Sciences), siguiendo las instrucciones del fabricante. Las membranas y las sondas fueron incubadas con 10mL del buffer Rapid-Hyb (GE Healthcare Life Sciences) con rotación a 65 °C durante toda la noche. Después de la hibridación, las membranas fueron lavadas dos veces con solución 2x SCC 0.1% SDS a 35 °C y fueron expuestas durante toda la noche a -80 °C y finalmente revelados.

5.1.3 Validación mediante 5' RACE (Rapid amplification of cDNA 5' ends).

La identificación del sitio de inicio de la transcripción de sRNA fue realizada mediante la metodología modificada de 5'RACE reportada previamente (Mendoza-Vargas et al. 2009). El RNA total fue extraído como se describió en apartados anteriores y posteriormente fue tratado con el enzima Tobacco Acid Pyrophosphatase (TAP Epicenter) para generar extremos 5'-monofosfato. Una vez realizado este paso se procedió a la ligación de un adaptador de RNA al extremo 5' (5'-GUUCAGAGUUCUACAGUC-CGACGAUC-3') utilizando la enzima T4 RNA ligasa 1 (ssRNA Ligase, New England Biolabs). Para la construcción de las librerías de cDNA se utilizó un random primer (5'-CAAGCAGAAGACGGCATAACGANNNNNN-3') para generar el cDNA mediante la reacción de transcripción reversa utilizando la enzima SuperScriptTM III (Invitrogen, Carlsbad, CA, USA). Posteriormente, se realizó una reacción de PCR utilizando oligonucleótidos específicos del gen de interés (enlistados en Anexo I) y los productos obtenidos fueron visualizados mediante electroforesis de geles de poliacrilamida al 8%, las bandas visualizadas fueron cortadas y purificadas. Los productos de PCR obtenidos fueron secuenciados por el método de Sanger en la Unidad de Secuenciación y Síntesis de Oligos del Instituto de Biotecnología de la UNAM. Las secuencias obtenidas fueron

alineadas al genoma de *G. sulfurreducens* y el sitio de inicio de la transcripción fue determinado como el primer nucleótido identificado inmediatamente después de la secuencia correspondiente al adaptador 5' utilizado.

5.1.4. Ensayos de RT-qPCR y de Co-transcripción.

El RNA total fue extraído y tratado con la enzima DNase como se describió previamente. El cDNA fue sintetizado utilizando el RevertAid H minus First Strand cDNA Synthesis Kit (Thermo Scientific). Este cDNA fue utilizado para dos experimentos diferentes: RT-qPCR y ensayos de co-transcripción. Para los ensayos de RT-qPCR fue empleado el Maxima SYBR Green/ROX qPCR Master Mix Kit (Thermo Scientific), siguiendo las instrucciones del fabricante. El gen *recA* fue utilizado como control interno para la normalización de los datos. Los ensayos fueron realizados por triplicado en dos experimentos independientes. Los resultados fueron analizados por el método de $2^{-\Delta\Delta C_T}$ (Livak and Schmittgen 2001). Los ensayos de co-transcripción consistieron en amplificar las regiones intergénicas correspondientes a los sRNAs junto con sus regiones colindantes. Los productos de PCR fueron visualizados en geles de agarosa al 2% y se utilizó el 100 bp DNA Ladder (New England BioLabs) como marcador de tamaño. La lista de los oligonucleótidos utilizados en estos ensayos se encuentra en el anexo I.

5.2 Resultados y Discusión

5.2.1 Identificación de sRNAs en regiones intergénicas.

En años recientes, la búsqueda de sRNAs localizados en regiones intergénicas (sRNA *trans-encoded*) mediante análisis de RNAseq ha sido ampliamente utilizada en un número creciente de microorganismos (Babski et al. 2016; Bronsard et al. 2017; Gaimster et al. 2016; Kwenda et al. 2016; McClure et al. 2014) . Utilizando los datos generados en los experimentos de RNAseq fue posible la identificación de sRNAs en regiones intergénicas en *G. sulfurreducens*. Como se mencionó en la sección de materiales y métodos de este capítulo, se utilizó como criterio seleccionar solo transcritos que se localizaran en regiones intergénicas con un tamaño mayor a >160 pares de bases y así evitar que estos pudieran corresponder a transcritos de genes colindantes. La búsqueda se realizó manualmente en las 736 regiones intergénicas con tamaño >160 pb que presenta el genoma de *G. sulfurreducens* PCA.

Siguiendo este procedimiento, pudimos identificar 75 candidatos a sRNAs a lo largo del genoma de *G. sulfurreducens*, expresados en al menos una de las condiciones de crecimiento probadas en este estudio. Posterior al análisis de búsqueda de ORFs, algunos de estos candidatos fueron descartados, dejando 45 sRNAs. Aunque el resto de los transcritos no fueron considerados para los análisis subsecuentes, estos transcritos pudieran ser estudiados posteriormente para determinar si realmente corresponden a transcritos codificantes y pudieran corresponder a genes no anotados en el genoma de *G. sulfurreducens*.

Con los 45 candidatos a sRNAs obtenidos de este análisis, se buscó si previamente se había identificado algún TSS localizado en estas regiones intergénicas (Tabla 5). Se encontró que en 6 casos había TSSs reportados por Qiu et al. (2010) y en 18 de estas regiones presentaban al menos un TSS identificados por el mapeo global de TSS realizado en este trabajo. Al igual que en el caso de la identificación de TSSs, la detección de transcritos por dos diferentes metodologías o experimentos independientes respaldan la identificación de estos sRNAs como transcritos verdaderos, lo cual es de gran importancia debido a la fecha no existen reportes de sRNAs validados experimentalmente en bacterias del género *Geobacter*.

Así mismo, se observó que los sRNAs identificados presentaron en promedio un tamaño de 241 nucleótidos, con un rango de 163 a 396 nucleótidos. La identificación de sRNAs de menor tamaño no fue posible debido a que la búsqueda de transcritos se realizó en regiones intergénicas con un tamaño de >160 nt; sin embargo, el tamaño de los sRNAs identificados se encuentra en el rango de los sRNAs caracterizados en otras bacterias (Wagner and Romby 2015). En la Tabla 5 se enlista los 45 sRNAs identificados, así como sus características generales.

Tabla 5. Lista de sRNAs identificados en regiones intergénicas. *Nomenclatura propuesta para los sRNAs identificados en este trabajo. &Tamaño del transcrito determinado por los experimentos de RNAseq. ^β TSSs identificados a través de mapeo global (Qiu et al. 2010; González et al. 2016). & Valor de logaritmo base dos del nivel de cambio (Log2 fold change).

Nombre*	Extremo Izquierdo	Extremo Derecho	Tamaño (nt)&	UpStream Gene	DownStream Gene	Valor de Expresión (Log2 NC&)	Sitio de Inicio de la transcripción previamente identificado ^β
gsusR01	87274	87446	172	GSU0071	GSU0073	0.90	Este trabajo
gsusR02	104069	104368	358	GSU0092	GSU0093	2.03	
gsusR03	133055	133269	214	GSU0117	GSU0118	1.31	
gsusR04	197815	198083	268	GSU0183	GSU0184	0.89	
gsusR05	49616	49949	333	GSUR003	GSU3471	-0.36	Este trabajo
gsusR06	201625	201904	286	GSU0188	GSU0189	3.15	
gsusR07	282674	282951	277	GSU0273	GSU0274	-1.03	
gsusR08	524637	524880	243	GSU0491	GSU0492	-2.45	Este trabajo
gsusR09	528510	528716	206	GSU0494	GSU0496	1.26	Este trabajo
gsusR10	602275	602600	325	GSU0573	GSU0574	1.50	
gsusR11	618931	619165	234	GSU0584	GSU0585	-0.46	
gsusR12	684497	684666	169	GSU0653	GSUR010	-0.81	Este trabajo
gsusR13	811717	812051	334	GSU3496	GSU3497	-0.10	Este trabajo
gsusR14	916580	916766	186	GSU0856	GSU0857	0.05	Qiu et al. 2010, este trabajo
gsusR15	999024	999252	228	GSU0929	GSU0930	1.09	
gsusR16	1014055	1014331	373	GSU0943	GSU0944	1.12	Qiu et al. 2010, este trabajo
gsusR17	1028665	1028938	280	GSU3521	GSU0955	0.03	Este trabajo
gsusR18	1034461	1034624	163	GSU3522	GSU0961	0.97	
gsusR19	1223942	1224128	186	GSU1139	GSUR017	-1.64	Este trabajo
gsusR20	1462928	1463133	205	GSU1334	GSU1336	0.67	
gsusR21	1468490	1468650	160	GSUR022	GSU1342	1.89	
gsusR22	1582345	1582628	283	GSU1444	GSU1445	-0.50	Este trabajo
gsusR23	2014960	2015291	331	GSU1844	GSU1845	-0.22	
gsusR24	1146941	1147190	249	GSU1060	GSU1061	-0.50	Qiu et al. 2010

gsusR25	1501662	1501853	191	GSU1372	GSU1374	1.23	Este trabajo
gsusR26	1728623	1728811	188	GSU1576	GSUR029	0.72	
gsusR27	1868689	1868866	176	GSU1704	GSU1705	0.80	Este trabajo
gsusR28	2136227	2136407	180	GSU1948	GSU1949	2.12	Qiu et al. 2010
gsusR29	2710567	2710932	365	GSU2471	GSU2472	-0.14	Este trabajo
gsusR30	2773965	2774256	291	GSU2515	GSU2516	-1.98	
gsusR31	2865179	2865410	231	GSU3598	GSU2599	1.59	
gsusR32	3087106	3087304	198	GSU2806	GSU2807	0.30	
gsusR33	3173046	3173442	396	GSU2887	GSU2888	0.83	
gsusR34	3772744	3773064	270	GSU3426	GSU3427	0.00	Este trabajo
gsusR35	3798878	3799080	202	GSU3450	GSU3451	3.21	Este trabajo
gsusR36	1712415	1712569	215	GSU1557	GSU1558	2.70	
gsusR37	2557663	2557846	183	GSU2337	GSU2338	1.84	
gsusR38	1292289	1292700	335	GSU1184	GSU1185	-0.10	Este trabajo
gsusR39	1468449	1468631	233	GSUR022	GSU1342	1.89	
gsusR40	1293188	1293777	198	GSU1185	GSU1188	0.78	
gsusR41	1281315	1281849	194	GSU1175	GSU1176	0.89	
gsusR42	1728927	1729225	191	GSUR029	GSU1577	-0.04	
gsusR43	1788673	1788893	211	GSU1631	GSU1632	-1.79	Este trabajo
gsusR44	3412180	3412611	177	GSU3108	GSU3109	-1.08	
gsusR45	2728179	2728471	161	GSU2484	GSU2485	1.36	Qiu et al. 2010

5.2.1 Validación y caracterización mediante northern blot, RT-qPCR y 5'RACE.

Para confirmar la expresión de los sRNAs identificados mediante RNAseq así como su tamaño, fueron seleccionados 12 sRNAs para realizar ensayos de northern blot. Además, para la validación de los patrones de expresión se realizaron ensayos de RT-qPCR y en algunos casos fue posible determinar el sitio de inicio de la transcripción de estos transcritos mediante la metodología de 5'RACE (Figura 12 y 13).

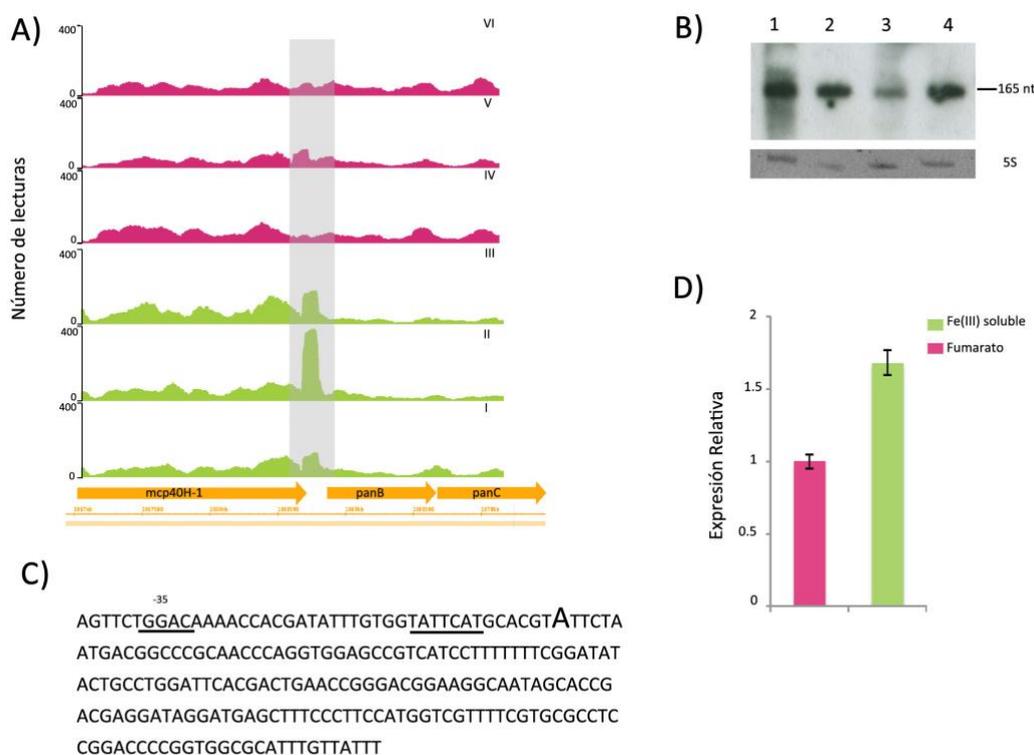


Figura 12. Caracterización del sRNA gsusR27. A) Patrón de expresión identificado para las dos condiciones de crecimientos estudiadas mediante RNAseq; I al III corresponden a triplicados durante la reducción de Fe(III) soluble y IV al VI durante la reducción de fumarato. La región gris corresponde al sRNA identificado B) Ensayo de northern blot con RNA extraído en fase exponencial y estacionaria durante la reducción de Fe(III) soluble (carril 1 y 2) y durante la reducción de fumarato (carril 3 y 4) 5S rRNA como control de carga. C) TSS identificado mediante ensayos de 5'RACE. D) Valores de expresión relativa obtenidos mediante RT-qPCR.

Los resultados de los ensayos de northern blot mostraron consistencia con los tamaños obtenidos mediante RNAseq. En general, se observó que el tamaño estimado de los transcritos por northern blot fueron ligeramente menores a los observados mediante RNAseq, a excepción de gsusR38, donde se observó una diferencia de aproximadamente 65 nt entre ambos ensayos. En la Figura 12 se observa la caracterización del sRNA gsusR27 mediante 4 diferentes metodologías. En este caso en particular, además del TSS identificado mediante

5'RACE, los resultados del mapeo global de TSSs detectaron un TSS exactamente en la misma posición identificada mediante 5'RACE, lo cual brinda una caracterización mas completa y robusta.

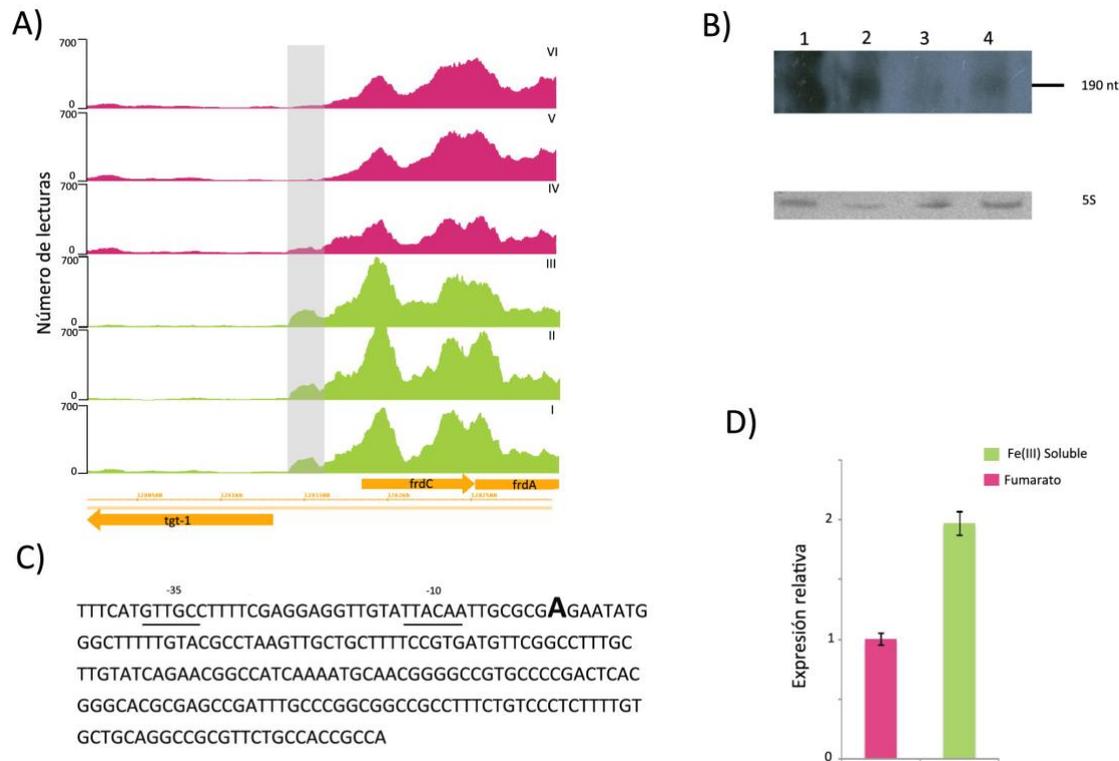


Figura 13. Caracterización del sRNA *gsusR41*. A) Patrón de expresión identificado para las dos condiciones de crecimientos estudiadas mediante RNAseq: I al III corresponden a triplicados durante la reducción de Fe(III) soluble y IV al VI durante la reducción de fumarato. B) Ensayo de northern blot con RNA extraído en fase exponencial y estacionaria durante la reducción de Fe(III) soluble (carril 1 y 2) y durante la reducción de fumarato (carril 3 y 4) 5S rRNA como control de carga. C) TSS identificado mediante ensayos de 5'RACE. D) Valores de expresión relativa obtenidos mediante RT-qPCR.

Debido a que en *E. coli* y *Salmonella* se sabe que solo un número reducido de sRNAs se expresan constitutivamente (Wagner and Romby 2015) y se ha observado que la mayoría de los sRNAs identificados y caracterizados en estos organismos modelo son activados bajo condiciones específicas de crecimiento (revisado en Corcoran et al. 2012), se decidió seleccionar sRNAs para detectar su expresión mediante ensayos de northern blot tanto en fase exponencial como estacionaria. Los resultados de estos ensayos indicaron que los sRNAs *gsusR08*, *gsusR17*, *gsusR27* y *gsusR41* (Figura 12B, 13B y 14) se expresan durante ambas fases de crecimiento en al menos una de las condiciones de cultivo estudiadas. Cabe destacar que *gsusR17* es un sRNA de 270 nt localizado en una región intergénica de 597 pb, lo cual es interesante ya que se ha planteado que los sRNAs tienden a localizarse en regiones intergénicas de un tamaño mayor al promedio (Tsai et al. 2015).

Por otro lado, los sRNAs *gsusR36* y *gsusR40* solo fueron detectados en fase exponencial durante la reducción de Fe(III) soluble. Para determinar si estos sRNAs se transcriben constitutivamente o son activados por algún tipo de estímulo es necesario hacer su detección en otras condiciones de crecimiento incluyendo diferentes tipos de donadores y aceptores de electrones.

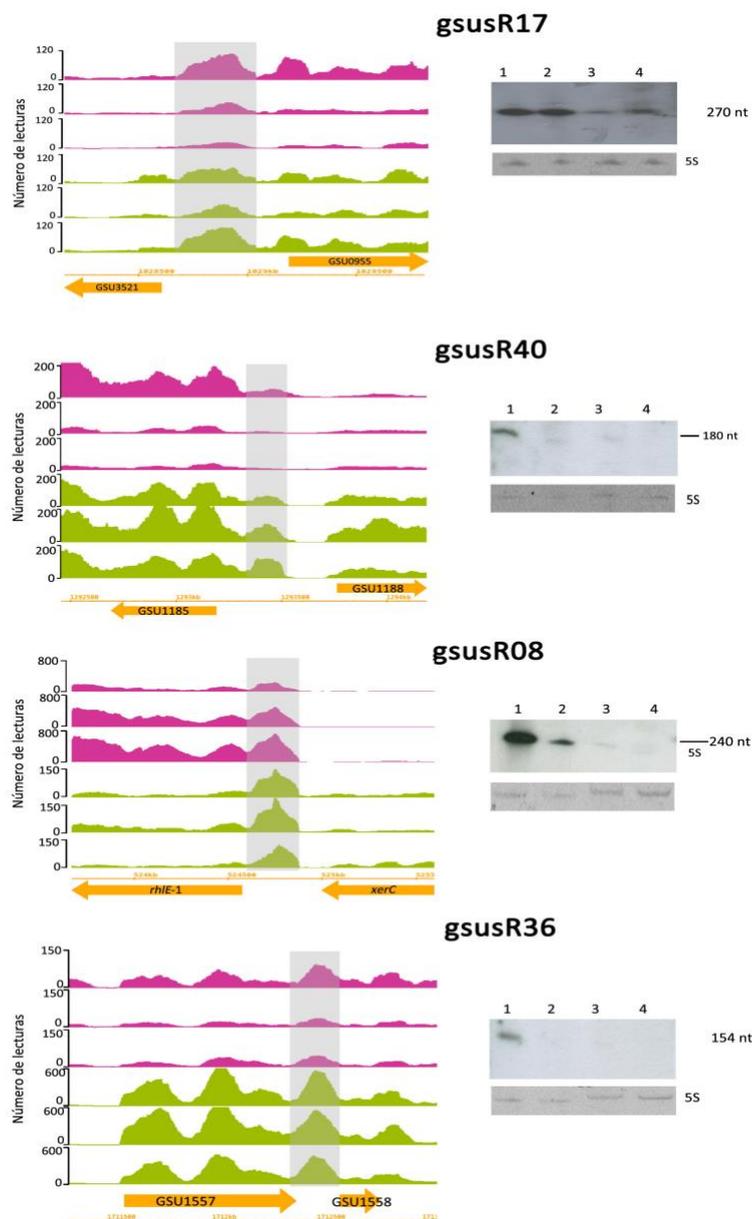


Figura 14. Caracterización de los transcritos *gsusR17*, *gsusR40*, *gsusR08* y *gsusR36*. Perfil transcripcional obtenido mediante RNaseq para las dos condiciones de crecimientos estudiadas: I al III corresponden a triplicados durante la reducción de Fe(III) soluble y IV al VI durante la reducción de fumarato. La detección de los transcritos por northern blot se realizó con RNA extraído en fase exponencial y estacionaria durante la reducción de Fe(III) soluble (carril 1 y 2) y durante la reducción de fumarato (carril 3y 4); el 5S rRNA fue utilizado como control de carga.

5.2.2 Análisis de expresión diferencial de los sRNAs identificados.

Con este análisis fue posible determinar si la expresión de los sRNAs identificados pudiera estar inducida por alguna de las condiciones de crecimiento estudiadas. De los 45 sRNAs identificados mediante RNAseq, 6 de estos mostraron valores de expresión diferencial significativa (Tabla 5), de los cuales 5 fueron sobre expresados durante la reducción de Fe(III) soluble. *gsusR02*, *gsusR06*, *gsusR28*, *gsusR35* y *gsusR36* fueron los sRNAs sobre expresados durante la reducción de Fe(III) soluble y *gsusR08* durante la reducción de fumarato. El sRNA *gsusR06* fue el que mostró el valor de expresión diferencial más alto durante la reducción de Fe(III) soluble (Figura 15). Este patrón de expresión fue validado tanto por los datos de RT-qPCR como por los ensayos de northern blot. Una característica interesante del sRNA *gsusR06* es el alto número de lecturas comparado con los otros sRNAs detectados, lo cual fue consistente en las 3 replicas biológicas (Figura 15A). Estos valores altos en su nivel de expresión pudieran sugerir un papel relevante de este sRNA durante esta condición de cultivo.

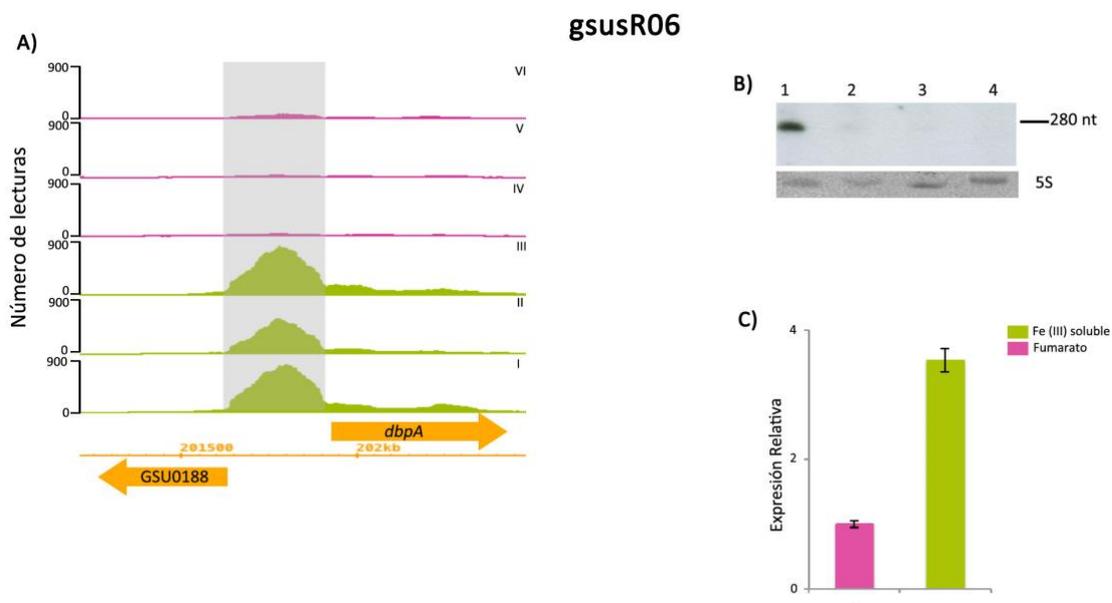


Figura 15. Caracterización del sRNA *gsusR06*. A) Patrón de expresión identificado para las dos condiciones de crecimientos estudiadas mediante RNA-seq: I al III corresponden a triplicados durante la reducción de Fe(III) soluble y IV al VI durante la reducción de fumarato. B) Ensayo de northern blot con RNA extraído en fase exponencial y estacionaria durante la reducción de Fe(III) soluble (carril 1 y 2) y durante la reducción de fumarato (carril 3 y 4) 5S rRNA como control de carga. C) Valores de expresión relativa obtenidos mediante RT-qPCR.

Se observó que algunos sRNAs presentaron valores de expresión diferencial cercanos al valor de corte (Log2 Fold change >2). Este fue el caso del sRNA *gsusR45*, el cual mostro un valor de cambio en su expresión de 1.3 durante la reducción de Fe(III), mientras que en el análisis de expresión relativa realizado mediante RT-qPCR arrojó un valor de 2.3 en su tasa de cambio. Las metodologías de RNA-seq han demostrado ser una estrategia eficaz para la identificación de sRNAs; sin embargo, puede presentar ciertos sesgos, como una baja profundidad en la secuenciación de ciertas regiones que no permiten la detección real en los cambios (Schroeder et al. 2015). Es por ello que debe considerarse muy importante que la identificación *de novo* de este tipo de transcritos se debe realizar por al menos dos metodologías independientes.

Como se mencionó anteriormente, 5 sRNAs presentaron valores de sobre expresión durante la reducción de Fe(III) soluble. Sería interesante profundizar en la hipótesis de que algunos de estos sRNAs pudieran estar involucrados en algún mecanismo de regulación durante la reducción de Fe(III). En *G. sulfurreducens*, los mecanismos de regulación genética durante la reducción de Fe(III) no son del todo conocidos, específicamente los relacionados con el metabolismo de Fe. En este sentido, Embree et al. (2014) estudiaron el papel del regulador transcripcional Fur durante la reducción de Fe(III) y fumarato. Los autores reportaron que Fur se encuentra sobre expresado durante la reducción de fumarato y coordina la respuesta transcripcional a las fluctuaciones en la concentración de Fe en el ambiente. Se sabe que en otros organismos la regulación llevada a cabo por Fur se da en coordinación con un sRNA, RyhB en el caso de *E. coli* (Oglesby-Sherrouse and Murphy 2013). Sin embargo, en *G. sulfurreducens* no se ha identificado al sRNA que participe en este circuito de regulación.

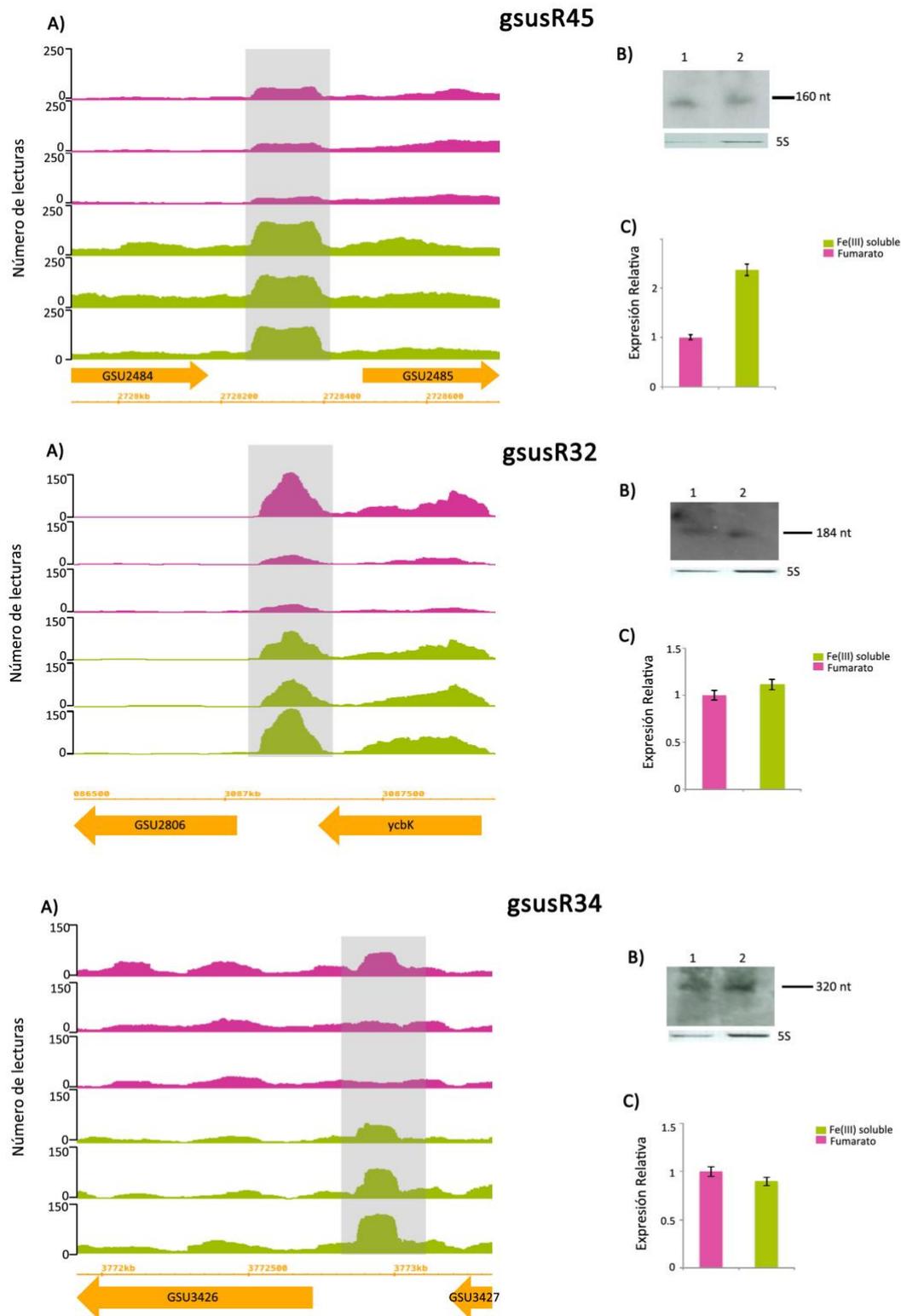


Figura 16. Caracterización de los sRNA gsus45, gsus32 y gsusR34. A) Perfil transcripcional obtenido mediante RNAseq para las dos condiciones de crecimientos estudiadas. Las gráficas de color verde corresponden a las 3 réplicas durante la reducción de Fe(III) y las rosas durante la reducción de fumarato B) Ensayo de northern blot con RNA durante la reducción de Fe(III) soluble y durante la reducción de fumarato (carril 1 y 2 respectivamente) y el 5S rRNA como control de carga. C) Valores de expresión relativa obtenidos mediante ensayos de RT-qPCR.

5.2.3 Ensayos de co-transcripción.

Algunos de los sRNAs identificados mediante RNAseq mostraron sobre posición de las lecturas detectadas. Con la finalidad de identificar si los transcritos de los sRNAs identificados en las regiones intergénicas son transcritos independientes o corresponden a transcritos de sus genes vecinos, fueron seleccionados 4 sRNAs para realizar ensayos de co-transcripción los cuales consistieron en amplificar fragmentos cortos y largos de las regiones intergénicas donde se localizaron los sRNAs, utilizando cDNA como templado (Figura 16). En este ensayo se pudo observar que, a excepción de *gsusR06*, los sRNAs evaluados en este ensayo se transcriben de manera independiente en ambas condiciones de crecimiento, ya que solo fue posible detectar el fragmento corto, el cual corresponde al sRNA identificado

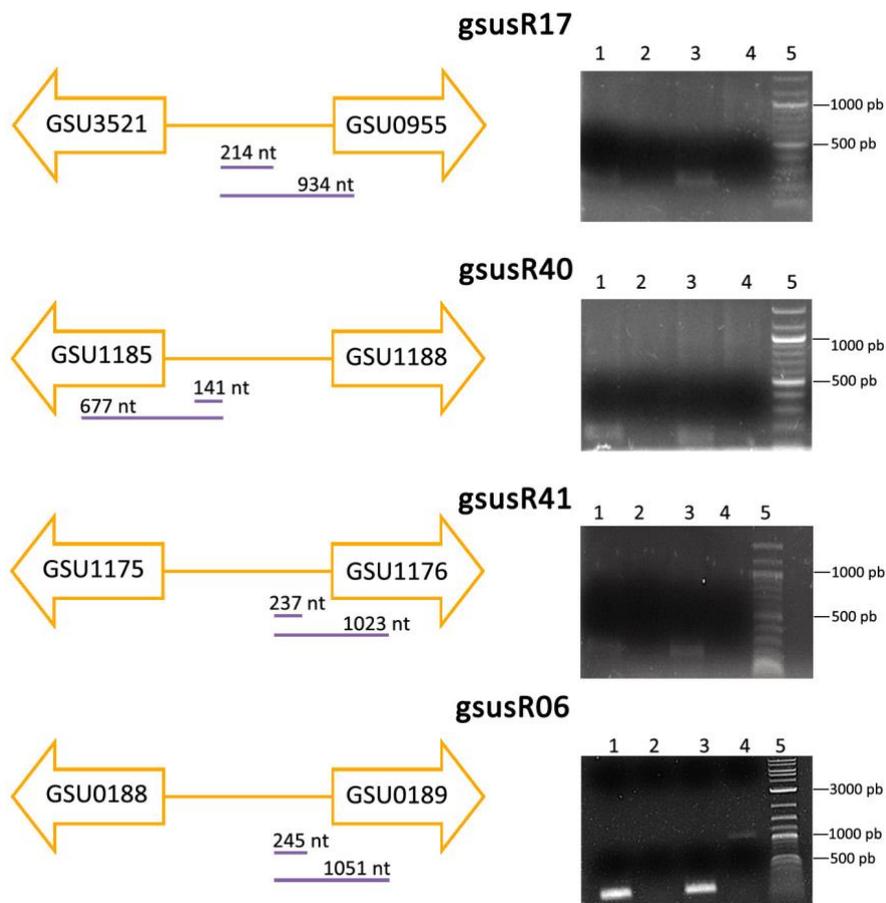


Figura 17. Ensayos de co-transcripción de los sRNAs *gsusR06*, *gsusR17*, *gsusR40* y *gsusR41*. Las flechas naranjas representan los genes colindantes de la región intergénica donde fue localizado el sRNA. Las líneas moradas representan el tamaño y localización de los fragmentos amplificados y su posición con respecto a los genes colindantes. Los fragmentos de PCR fueron visualizados en gels de agarosa al 2%. Carril 1 y 2 corresponde al fragmento corto y largo durante la reducción de Fe(III) soluble, carriles 3 y 4 fragmentos cortos y largos detectados durante la reducción de fumarato y carril 5 marcador de DNA.

Para el caso del sRNA gsusR06 se observó la detección del transcrito largo durante la reducción de fumarato, lo cual sugiere que durante esta condición de cultivo gsusR06 y el gen GSU0189 se co-transcriben. Se ha reportado que en algunos sRNAs la cotranscripción se presenta en función de alguna condición de crecimiento, como el sRNA R_{Ss0680a} de *R. Sphaeroides*, el cual se co-transcribe con el gen que se localiza río abajo en función de la exposición a oxígeno (Berghoff et al. 2009). Así mismo, *D. radiodurans* en condiciones de radiación ionizante se detectó la co-transcripción de 6 sRNAs (Tsai et al. 2015). En los trabajos anteriormente mencionados, se sugiere que a pesar de que estos sRNAs se cotranscriben, es muy probable que posteriormente sufran modificaciones post transcripcionales que puedan servir para dar estabilidad a la región 5' UTR e inducir la degradación del RNA o bien para brindar estabilidad durante la transcripción bajo alguna condición de crecimiento específica.

5.2.4 Identificación de homólogos de los sRNAs identificados y conservación de su secuencia.

Dado que los sRNAs identificados y validados en este trabajo son los primeros caracterizados para el género *Geobacter*, surgió la interrogante de conocer los niveles de conservación a nivel de secuencia entre las bacterias de este género. En este sentido, fueron utilizadas las secuencias de los 45 sRNAs identificados para realizar búsquedas en las bases de datos Rfam (Burge et al. 2012) y BSRD (Li et al. 2013). Los resultados obtenidos indicaron que un fragmento del sRNA gsusR27 mostró homología con un elemento de regulación llamado "pan RNA motif", el cual fue identificado en delta-protobacterias mediante estudios de genómica comparativa (Weinberg et al. 2010). Este motivo se describe como un RNA estructurado de un tamaño aproximado de 78 nt presente en la región 5' UTR de genes involucrados en la síntesis de pantotenato. Los resultados obtenidos por el mapeo de global de TSSs, 5'RACE y northern blot, indican que en *G. sulfurreducens* este transcrito posee un tamaño de 165 nt y se localiza río arriba del gen *panB* (Figura 12). Al realizar un análisis tipo BLAST (Altschul et al. 1997) utilizando la secuencia de gsusR27, observamos que esta secuencia solo se encontraba presente en tres cepas de *G. sulfurreducens* y en otros dos organismos del género *Geobacter* (Figura 18A). Sin embargo, se observó que en los otros dos organismos donde se detectó conservación de esta secuencia, el contexto genómico no mostró conservación. Con respecto a su estructura secundaria, el análisis de predicción mostró una estructura con elementos estructurales del tipo "hairpin loop" (Figura 18B), los cuales se sabe que son

identificación de sRNAs en diferentes especies de este género y se ha observado bajos niveles de conservación, inclusive en cepas de la misma especie (Gómez-Lozano et al. 2012). Este mismo patrón se ha observado en otros géneros bacterianos, por lo que se ha propuesto que estos bajos niveles de conservación pudiera ser reflejo de una diversificación metabólica que ayuda a una rápida adaptación en los diferentes ambientes y nichos que colonizan estas bacterias (Kopf et al. 2015; Voigt et al. 2014). En este sentido, sería interesante ampliar la identificación de sRNAs en otras especies del género *Geobacter* ya que se han especializado en colonizar nichos y ambientes con características ambientales donde otros organismos no han sido capaces de proliferar.

Tabla 6. Lista de microorganismos donde se observó conservación de secuencias con los sRNAs identificados.

sRNA	Clase	Orden	Género/Cepa
gsusR01	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR02	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR03	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15
gsusR04	δ -protobacteria	<i>Geobacter</i> , <i>Pelobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>P. propionicus</i> DSM 2370, <i>G. uraniireducens</i> Rf4
gsusR05	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1
gsusR06	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR07	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR08	δ -protobacteria	<i>Geobacter</i> , <i>Desulfuromonas</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>Desulfuromonas</i> sp DDH964
gsusR09	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>G. uraniireducens</i> Rf4, <i>G. daltonii</i> FRC-32, <i>G.</i>

			sp. M21, <i>G. bemidjensis</i> Bem.
gsusR10	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400,
gsusR11	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR12	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13
gsusR13	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA
gsusR14	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>G. uraniireducens</i> Rf4, <i>G. sp. M21</i> , <i>G. bemidjensis</i> Bem.
gsusR15	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400,
gsusR16	δ-protobacteria	<i>Geobacter</i> , <i>Pelobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>P. propionicus</i> DSM 2379, <i>G. uraniireducens</i> Rf4, <i>G. bemidjensis</i> Bem, <i>G. daltonii</i> FRC-32, <i>G. lovleyi</i> SZ
gsusR17	δ-protobacteria γ-protobacteria	<i>Geobacter</i> , <i>Desulfurivibrio</i> <i>Pelobacter</i> , <i>Desulfobacterium</i> , <i>Acidihalobacter</i> , <i>Thioalkalivibrio</i> , <i>Allochromatium</i>	<i>G. sulfurreducens</i> PCA, <i>G. anodireducens</i> SD-1, <i>G. sp. M18</i> , <i>G. daltonii</i> FRC-32, <i>G. sp</i> DSM9738, <i>D. alkaliphilus</i> AHT2, <i>P. propionicus</i> DSM 2379, uncultured
gsusR18	δ-protobacteria	<i>Geobacter</i>	<i>Desulfobacterium sp</i> , <i>A. prosperus</i> V6, <i>T. nitratireducens</i> DSM 14787, <i>A. vinosum</i> DSM 180
gsusR19	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400,
gsusR20	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1,
gsusR21	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400,
gsusR22	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR23	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400,

gsusR24	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>G. bemidjensis</i> Bem,
gsusR25	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR26	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400,
gsusR27	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13,
gsusR28	δ-protobacteria	<i>Geobacter</i> , <i>Pelobacter</i> , <i>Desulfuromonas</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>P. propionicus</i> DSM 2379, <i>G. uraniireducens</i> Rf4, <i>G. bemidjensis</i> Bem, <i>G. daltonii</i> FRC-32, <i>D. soudanensis</i> WTL
gsusR29	δ-protobacteria γ-protobacteria	<i>Geobacter</i> , <i>Desulfurivibrio</i> <i>Pelobacter</i> , <i>Acidihalobacter</i> , <i>Thioalkalivibrio</i>	<i>G. sulfurreducens</i> PCA, <i>G. anodireducens</i> SD-1, <i>G. sp.</i> M18, <i>G. daltonii</i> FRC-32, <i>G. sp.</i> DSM9738, <i>D. alkaliphilus</i> AHT2, <i>P. propionicus</i> DSM 2379, <i>A. prosperus</i> V6, <i>T. nitratireducens</i> DSM 14787
gsusR30	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA
gsusR31	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400,
gsusR32	δ-protobacteria	<i>Geobacter</i> , <i>Desulfurivibrio</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>D. alkaliphilus</i> AHT2
gsusR33	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR34	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. pickeringii</i> G13,
gsusR35	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR36	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR37	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15
gsusR38	δ-protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400

gsusR39	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400
gsusR40	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13, <i>G. uraniireducens</i> Rf4
gsusR41	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR42	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400
gsusR43	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. metalirreducens</i> GS-15, <i>G. pickeringii</i> G13,
gsusR44	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1
gsusR45	δ -protobacteria	<i>Geobacter</i>	<i>G. sulfurreducens</i> PCA, <i>G. sulfurreducens</i> AM1, <i>G. sulfurreducens</i> KN400, <i>G. anodireducens</i> SD-1, <i>G. bemidjensis</i> Bem,

Solo en gsusR17 y gsusR29 se observó similitud de secuencia en bacterias fuera de la clase de δ -protobacteria. Es interesante destacar que, en ambos casos, la secuencia no mostró conservación en otras cepas de *G. sulfurreducens* pero sí con organismos pertenecientes a la clase γ -protobacteria y con otros δ -protobacteria. La secuencia de gsusR17 fue la que mostró mayor nivel de conservación (Figura 19) y al analizar si esta conservación se extendía a los genes colindantes, solo se observó conservación del contexto genómico en *P. propionicus* y *T. nitratireducens*. Aunque se ha propuesto que las regiones donde se localizan sRNAs tienen diferentes niveles de conservación con respecto a sus genes colindantes (Tsai et al. 2015) aún no quedan claros los mecanismos de evolución que dan origen a los sRNA bacterianos y si estos se dan independientes de su contexto genómico.

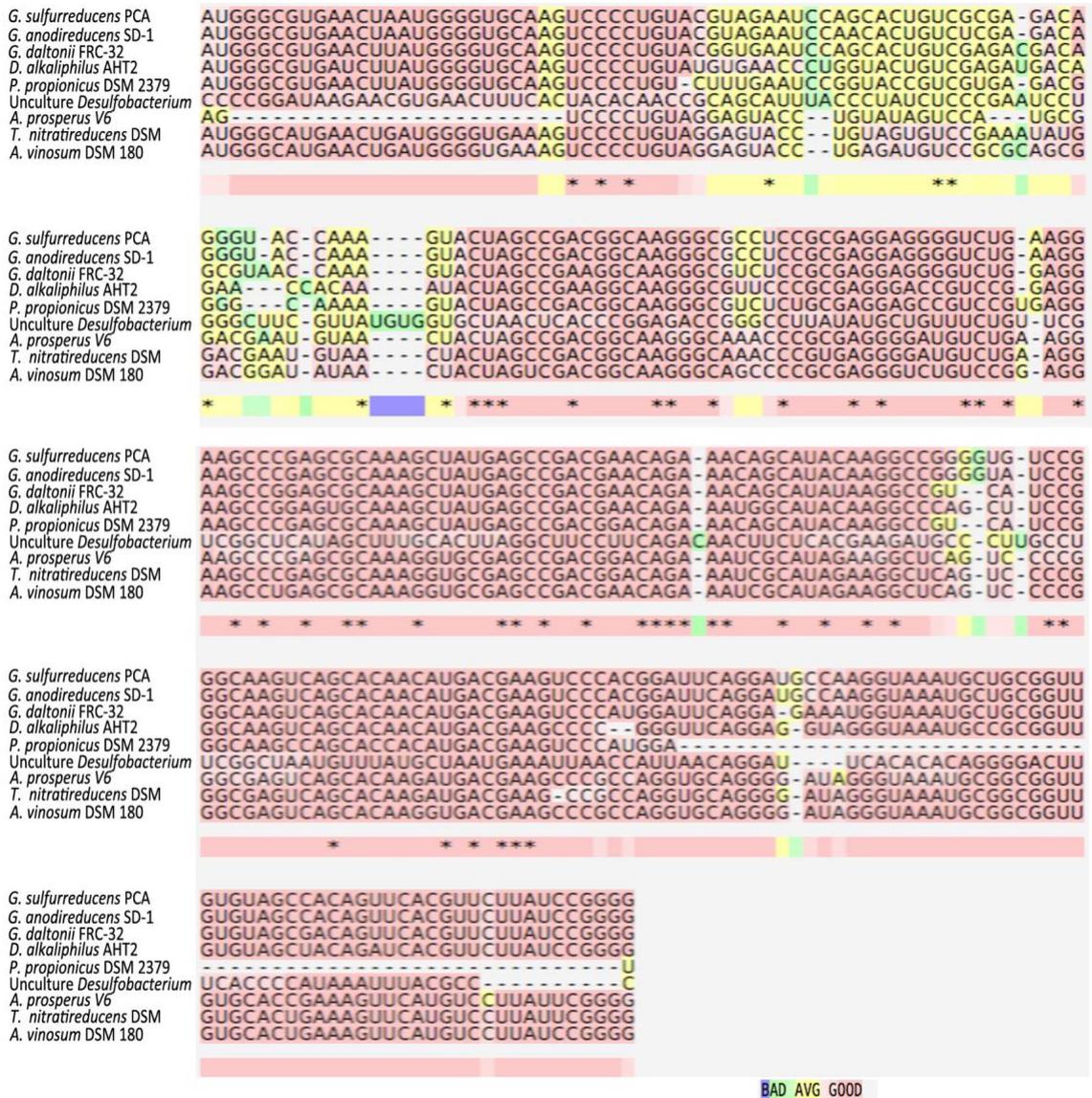


Figura 19. Alineamiento de secuencias múltiples realizado con la herramienta M-Coffee (<http://tcoffee.org.cat/>).

5.2.5. Identificación de posibles genes blanco.

A pesar del constante aumento en la identificación de sRNAs en muchos grupos bacterianos, la caracterización de los mecanismos de acción permanece desconocida para la mayoría de los sRNAs identificados a la fecha. Una primera aproximación para conocer la función de estos sRNAs es la predicción de sus posibles genes blanco mediante análisis *in silico*. En años recientes se han desarrollado herramientas bioinformáticas poderosas dirigidas a este fin (Backofen and Hess 2010; Pain et al. 2015; Siqueira et al. 2016). Para este análisis se utilizó la herramienta TargetRNA2 (Kery et al. 2014) y se utilizaron las secuencias de sRNAs que fueron validados por diferentes experimentos. En este análisis se observó que la mayoría de

los sRNAs mostraron múltiples genes blanco; sin embargo, solo se consideraron aquellos genes blancos que presentaron el mayor índice de probabilidad (Tabla 7).

Tabla 7. Lista de predicciones de posibles blancos de los sRNAs identificados.

sRNAs	Posible blanco	Función del gen blanco
gsusR02	GSU2586	Hypothetical protein
gsusR06	GSU0188	Hypothetical protein
gsusR08	GSU0196	acyl-CoA thioesterase
gsusR17	GSU0828	RND family efflux pump outer membrane protein
gsusR27	GSU1100	Hypothetical protein
gsusR32	GSU2806	Nitrogenase molybdenum-cofactor synthesis protein NifE
gsusR34	GSU2261	Lipid-A-disaccharide synthase
gsusR36	GSU006	prophage regulatory protein
gsusR38	GSU2586	Hypothetical protein
gsusR40	GSU2451	Peptide methionine sulfoxide reductase
gsusR41	GSU1175	Queuine tRNA-ribosyltransferase
gsusR45	GSU0355	Hypothetical protein

Aunque estos resultados son una primera aproximación, resulta interesante observar que los genes blanco corresponden a diferentes grupos funcionales. Dadas las características metabólicas de *G. sulfurreducens* resultan interesantes aquellos genes blanco correspondientes a reductasas y proteínas de membrana. Sin embargo, sería necesario profundizar en este tipo de análisis con otras herramientas bioinformáticas en conjunto con una validación experimental.

CAPÍTULO VI

Conclusiones generales y Perspectivas

6.1. Conclusiones generales.

- El empleo de herramientas globales de transcriptómica brindó la posibilidad de enriquecer el conocimiento de la estructura transcripcional de *G. sulfurreducens*.
- Mediante el empleo de diferentes metodologías fue posible la identificación y validación de los primeros sRNAs reportados para *G. sulfurreducens*, lo cual brinda un importante aporte al conocimiento del sRNAoma² de las especies del género *Geobacter*.
- Se observó que al menos 6 de los sRNAs estudiados mostraron expresión diferencial en ambas condiciones de crecimiento, de los cuales 5 fueron sobre expresados durante la reducción de Fe(III).
- No se encontraron homólogos de los sRNAs identificados en este estudio y los análisis de conservación de secuencia indicaron que se encuentran poco conservados a nivel de género e inclusive en algunos casos solo están presentes en la cepa PCA. Solo 7 de los 45 sRNA identificados mostraron conservación de secuencia mas allá del género *Geobacter*.
- La predicción de posibles blancos sugiere que los sRNAs pudieran estar involucrados en la regulación de una amplia variedad de procesos celulares.

² sRNA transcriptoma

6.2. Perspectivas.

Debido a que este es el primer estudio de identificación de sRNAs en el género *Geobacter*, la caracterización e identificación de este tipo de RNAs reguladores se encuentra en sus inicios, por lo que son varias las líneas que se pueden seguir para ahondar en su caracterización funcional:

1. Establecer si los sRNAs son funcionales mediante la generación de mutantes que permitan reconocer algún fenotipo.
2. Debido a que se ha reportado que la proteína chaperona de RNA Hfq se encuentra implicada en los mecanismos de regulación mediados por sRNAs, sería interesante por un lado generar una mutante de Hfq para entender su relevancia en *G. sulfurreducens* y por otro lado realizar ensayos de unión a RNA para determinar su interacción con los sRNAs identificados.
3. Continuando con estudios globales, podrían realizarse experimentos de tipo RIP-seq o CLIP-seq para poder determinar la unión de estos transcritos a proteínas específicas.
4. Ampliar la búsqueda de sRNAs en otras condiciones de crecimientos y en otras especies del género *Geobacter*.

REFERENCIAS

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST : a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25(17):3389–402.

Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010; 11(10):R106.

Babski J, Haas KA, Näther-Schindler D, Pfeiffer F, Förstner KU, Hammelmann M, et al. Genome-wide identification of transcriptional start sites in the haloarchaeon *Haloferax volcanii* based on differential RNA-Seq (dRNA-Seq). *BMC Genomics*; 2016;17(1):629.

Backofen R, Hess WR. Computational prediction of sRNAs and their targets in bacteria. *RNA Biol.* 2010;7(1):33–42.

Bardill JP, Zhao X, Hammer BK. The *Vibrio cholerae* quorum sensing response is mediated by Hfq-dependent sRNA/mRNA base pairing interactions. *Mol Microbiol.* 2011; 80(5):1381–94.

Berghoff BA, Glaeser J, Sharma CM, Vogel J, Klug G. Photooxidative stress-induced and abundant small RNAs in *Rhodobacter sphaeroides*. *Mol Microbiol.* 2009;74:1497–512.

Besemer J, Borodovsky M. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* 2005; 1:33

Bobrovskyy M, Vanderpool CK. Regulation of bacterial metabolism by small RNAs using diverse mechanisms. *Annu Rev Genet.* 2013; 47:209–32.

Bond DR, Lovley DR. Electricity Production by *Geobacter sulfurreducens* Attached to Electrodes. *Society.* 2003;69(3):1548–55.

Bronsard J, Pa G, Sassi M, Mauro T, Yoann A. identification in *Staphylococcus aureus*

highlights an unusual sRNA gene cluster with one encoding a secreted peptide. *Sci Rep.* 2017; 1–17.

Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, et al. Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.* 2012; 41:D226–32.

Butler J, Young N, Lovley D. Evolution of electron transfer out of the cell: comparative genomics of six *Geobacter* genomes. *BMC Genomics.* 2010; 11(40).

Butler JE, Young ND, Lovley DR, Lovley D, Holmes D, Nevin K, et al. Evolution of electron transfer out of the cell: comparative genomics of six *Geobacter* genomes. *BMC Genomics.* 2010; 11(1):40.

Caccavo F, Lonergan DJ, Lovley DR, Davis M, Stolz JF, Mcinerneyl MJ. Oxidizing Dissimilatory Metal-Reducing Microorganism. *Microbiology.* 1994; 60(10):3752–9.

Cartron ML, Maddocks S, Gillingham P, Craven CJ, Andrews SC. Feo--transport of ferrous iron into bacteria. *Biometals.* 2006; 19(2):143–57.

Coates JD, Phillips EJP, Lonergan DJ, Jenter H, Lovley DR. Isolation of *Geobacter* species from diverse sedimentary environments. *Appl Environ Microbiol.* 1996; 62(5):1531–6.

Coppi M V, Leang C, Sandler SJ, Lovley DR. Development of a Genetic System for *Geobacter sulfurreducens*. *Appl Environ Microbiol.* 2001; 67(7):3180–7.

Ding Y-HR, Hixson KK, Aklujkar M a, Lipton MS, Smith RD, Lovley DR, et al. Proteome of *Geobacter sulfurreducens* grown with Fe(III) oxide or Fe(III) citrate as the electron acceptor. *Biochim Biophys Acta.* 2008; 1784(12):1935–41.

Ding Y-HR, Hixson KK, Giometti CS, Stanley A, Esteve-Núñez A, Khare T, et al. The proteome of dissimilatory metal-reducing microorganism *Geobacter sulfurreducens* under various growth conditions. *Biochim Biophys Acta.* 2006; 1764(7):1198–206.

Embree M, Qiu Y, Shieu W, Nagarajan H, O’Neil R, Lovley D, et al. The iron stimulon and fur regulon of *Geobacter sulfurreducens* and their role in energy metabolism. *Appl Environ Microbiol.* 2014; 80(9):2918–27.

Esteve-núñez A, Núñez C, Lovley DR, Esteve A. Preferential Reduction of Fe (III) over Fumarate by *Geobacter sulfurreducens*. *J Bacteriol.* 2004; 186(9):2897–9.

Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database : towards a more sustainable future. *Nucleic Acids Res.* 2016; 44:279–85.

Gaimster H, Chalklen L, Alston M, Munnoch JT, Richardson DJ, Gates AJ, et al. Genome-Wide Discovery of Putative sRNAs in *Paracoccus denitrificans* Expressed under Nitrous Oxide Emitting Conditions. *Front Microbiol.* 2016; 7:1806.

Galushko AS, Schink B. Oxidation of acetate through reactions of the citric acid cycle by *Geobacter sulfurreducens* in pure culture and in syntrophic coculture. *Arch Microbiol.* 2000; 174(5):314–21.

Geelhoed JS, Henstra AM, Stams AJM. Carboxidotrophic growth of *Geobacter sulfurreducens*. *Appl Microbiol Biotechnol.* 2016; 100:7–99.

Georg J, Hess WR. cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev.* 2011; 75(2): 286–300.

Gómez-Lozano M, Marvig RL, Molin S, Long KS. Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*. *Environ Microbiol.* 2012; 14(8):2006–16.

Gottesman S, Storz G. Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol.* 2010 Oct.

Haas BJ, Chin M, Nusbaum C, Birren BW, Livny J. How deep is deep enough for RNA-

Seq profiling of bacterial transcriptomes? BMC Genomics. 2012; 13(1): 734.

Helt G a, Nicol JW, Erwin E, Blossom E, Blanchard SG, Chervitz S, et al. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. BMC Bioinformatics. 2009; 10(20):266.

Hernández-Eligio A, Andrade Á, Soto L, Morett E, Juárez K. The unphosphorylated form of the PilR two-component system regulates pilA gene expression in *Geobacter sulfurreducens*. Environ Sci Pollut Res. 2016; 1–9.

Holmes DE, Mester T, O’Neil R a, Perpetua L a, Larrahondo MJ, Glaven R, et al. Genes for two multicopper proteins required for Fe(III) oxide reduction in *Geobacter sulfurreducens* have different expression patterns both in the subsurface and on energy-harvesting electrodes. Microbiology. 2008; 154:1422–35.

Juárez K, Kim B-C, Nevin K, Olvera L, Reguera G, Lovley DR, et al. PilR, a transcriptional regulator for pilin and other genes required for Fe(III) reduction in *Geobacter sulfurreducens*. J Mol Microbiol Biotechnol. 2009; 16(3-4):146–58.

Kacharia FR, Millar JA, Raghavan R. Emergence of New sRNAs in Enteric Bacteria is Associated with Low Expression and Rapid Evolution. J Mol Evol. 2017; 84(4):204–13.

Karlin S, Brocchieri L, Mrázek J, Kaiser D. Distinguishing features of delta-proteobacterial genomes. Proc Natl Acad Sci. 2006; 103(30):11352–7.

Kellenberger C a, Wilson SC, Hickey SF, Gonzalez TL, Su Y, Hallberg ZF, et al. GEMM-I riboswitches from *Geobacter* sense the bacterial second messenger cyclic AMP-GMP. Proc Natl Acad Sci. 2015; 6: 1–6.

Kery MB, Feldman M, Livny J, Tjaden B. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. Nucleic Acids Res. 2014; 42: W124–9.

Kopf M, Klähn S, Scholz I, Hess WR, Voß B. Variations in the non-coding transcriptome

as a driver of inter-strain divergence and physiological adaptation in bacteria. *Sci Rep.* 2015; 5: 9560.

Kröger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp K, et al. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci.* 2012; 109(20):E1277–86.

Kwenda S, Gorshkov V, Ramesh AM, Naidoo S, Rubagotti E, Birch PRJ, et al. Discovery and profiling of small RNAs responsive to stress conditions in the plant pathogen *Pectobacterium atrosepticum*. *BMC Genomics.* 2016; 17(1): 47.

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012; 9(4): 357–9.

Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penadés JR, Valle J, Solano C. Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Natl Acad Sci.* 2011; 108(50): 20172-7.

Leang C, Qian X, Mester T, Lovley DR. Alignment of the c-Type Cytochrome OmcS Along Pili of *Geobacter sulfurreducens*. *Appl Environ Microbiol.* 2010; 76(12): 4080–4.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25(14): 1754–60.

Li L, Huang D, Cheung MK, Nong W, Huang Q, Kwan HS. BSRD: a repository for bacterial small regulatory RNA. *Nucleic Acids Res.* 2013; 41: D233–8.

Liu JM, Camilli A. A broadening world of bacterial small RNAs. *Curr Opin Microbiol.* 2010; 13(1): 18–23.

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. *Methods.* 2001; 25(4): 402–8.

Lloréns-Rico V, Cano J, Kamminga T, Gil R, Latorre A, Chen W-H, et al. Bacterial antisense RNAs are mainly the product of transcriptional noise. *Sci Adv.* 2016; 2(3): 1–10.

Lovley DR, Holmes DE, Nevin KP. Dissimilatory Fe(III) and Mn(IV) reduction. *Adv Microb Physiol.* 2004; 49(2): 219–86.

Lovley DR, Philips EJP. Novel Mode of Microbial Energy Metabolism : Organic Carbon Oxidation Coupled to Dissimilatory Reduction of Iron or Manganese. 1988; 54(6): 1472–80.

Lovley DR, Ueki T, Zhang T, Malvankar NS, Shrestha PM, Flanagan K, et al. *Geobacter*: the microbe electric's physiology, ecology, and practical applications. *Advances in microbial physiology.* 2011. 1-100 p.

Lovley DR. Bug juice: harvesting electricity with microorganisms. *Nat Rev Microbiol.* 2006; 4(7): 497–508.

Lovley DR. Cleaning up with genomics: Applying molecular biology to bioremediation. *Nat Rev Microbiol.* 2003; 1: 35–44.

Lovley DR. Electromicrobiology. *Annu Rev Microbiol.* 2012; 66(1): 391–409.

Mahadevan R, Bond DR, Butler JE, Coppi V, Palsson BO, Schilling CH, et al. Characterization of Metabolism in the Fe(III) Reducing Organism *Geobacter sulfurreducens* by Constraint-Based Modeling. *Appl Environ Microbiol.* 2006; 72(2): 1558–68.

Malvankar NS, Lovley DR. Microbial nanowires: A new paradigm for biological electron transfer and bioelectronics. *ChemSusChem.* 2012; 5(6): 1039–46.

Malvankar NS, Vargas M, Nevin KP, Franks AE, Leang C, Kim B-C, et al. Tunable

metallic-like conductivity in microbial nanowire networks. *Nat Nanotechnol.* 2011; 6(9): 573–9.

McClure R, Tjaden B, Genco C. Identification of sRNAs expressed by the human pathogen *Neisseria gonorrhoeae* under disparate growth conditions. *Front Microbiol.* 2014; 5: 1–12.

Mellin JR, Cossart P. Unexpected versatility in bacterial riboswitches. *Trends Genet.* 2015; 31(3): 150–6.

Melton ED, Swanner ED, Behrens S, Schmidt C, Kappler A. The interplay of microbially mediated and abiotic reactions in the biogeochemical Fe cycle. *Nat Rev Microbiol.* 2014; 12(12): 797–809.

Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, et al. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One.* 2009; 4(10): e7526.

Méthé B, Nelson KE, Eisen J, Paulsen IT, Nelson W, Heidelberg JF, et al. Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments. *Science.* 2003; 302(5652): 1967–9.

Methe BA, Webster J, Nevin K, Butler J, Lovley DR. DNA Microarray Analysis of Nitrogen Fixation and Fe(III) Reduction in *Geobacter sulfurreducens*. *Appl Environ Microbiol.* 2005;71(5):2530–8.

Michaux C, Verneuil N, Hartke A, Giard JC. Physiological roles of small RNA molecules. *Microbiology.* 2014 ;160: 1007–19.

Nies DH. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol Rev.* 2003; 27(2-3):313–39.

Núñez C, Adams L, Childers S, Lovley DR. The RpoS Sigma Factor in the Dissimilatory

Fe(III) Reducing Bacterium *Geobacter sulfurreducens*. J Bacteriol. 2004; 186(16): 5543–6.

O’Neil R, Holmes DE, Coppi M V, Adams L, Larrahondo MJ, Ward JE, et al. Gene transcript analysis of assimilatory iron limitation in *Geobacteraceae* during groundwater bioremediation. Environ Microbiol. 2008; 10(5): 1218–30.

Oelgeschläger E, Rother M. Carbon monoxide-dependent energy metabolism in anaerobic bacteria and archaea. Arch Microbiol. 2008; 190(3): 257–69.

Oglesby-Sherrouse AG, Murphy ER. Iron-responsive bacterial small RNAs: variations on a theme. Metallomics. 2013; 5(4): 276–86.

Pain A, Ott A, Amine H, Rochat T, Bouloc P, Gautheret D. An assessment of bacterial small RNA target prediction programs. RNA Biol. 2015; 12(5): 509–13.

Qiu Y, Cho B-K, Park YS, Lovley D, Palsson B, Zengler K. Structural and operational complexity of the *Geobacter sulfurreducens* genome. Genome Res. 2010; 20(9): 1304–1311.

Raghavan R, Sloan D, Ochman H. Antisense transcription is pervasive but rarely conserved in enteric bacteria. MBio. 2012; 3(4): 1–7.

Richter AS, Backofen R. General features of bacterial small RNA-mRNA interactions? RNA Biol. 2012; 9(7): 954–65.

Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26(1): 139–40.

Romeo T, Vakulskas CA, Babitzke P. Post-transcriptional regulation on a global scale : form and function of Csr / Rsm systems. Environ Microbiol. 2013; 15: 313–24.

Rotaru A, Shrestha PM, Liu F, Ueki T, Nevin K, Summers ZM, et al. Interspecies

Electron Transfer via Hydrogen and Formate Rather than Direct Electrical Connections in Cocultures of *Pelobacter carbinolicus* and *Geobacter sulfurreducens*. *Appl Environ Microbiol.* 2012; 78(21): 7645–51.

Salgado H, Peralta-Gil M, Gama-Castro S, Santos-Zavaleta A, Muñiz-Rascado L, García-Sotelo JS, et al. RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res.* 2013; 41: D203–13.

Schroeder CLC, Narra HP, Rojas M, Sahni A, Patel J, Khanipov K, et al. Bacterial small RNAs in the Genus *Rickettsia*. *BMC Genomics.* 2015; 16(1): 1075.

Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature.* 2010; 464(7286): 250–5.

Sharma CM, Vogel J. Differential RNA-seq: the approach behind and the biological insight gained. *Curr Opin Microbiol.* 2014; 19: 97–105.

Silva SM, Voordouw J, Leita C. Function of formate dehydrogenases in *Desulfovibrio vulgaris* Hildenborough energy metabolism. *Microbiology.* 2013:1760–9.

Siqueira FM, de Moraes GL, Higashi S, Beier LS, Breyer GM, de Sá Godinho CP, et al. *Mycoplasma* non-coding RNA: identification of small RNAs and targets. *BMC Genomics.* 2016; 17(8): 743.

Sorek R, Cossart P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet.* Jan; 11(1): 9–16.

Summers AO. Damage control: regulating defenses against toxic metals and metalloids. *Curr Opin Microbiol.* 2009; 12(2): 138–44.

Sun D, Wang A, Cheng S, Yates M, Logan BE, Logan BE. *Geobacter anodireducens* sp. nov., an exoelectrogenic microbe in bioelectrochemical systems. *Int J Syst Evol*

Microbiol. 2014; 3485–91.

Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-seq: a matter of depth. *Genome Res.* 2011; 21(12): 2213–23.

Thomason MK, Bischler T, Eisenbart SK, Förstner KU, Zhang A, Herbig A, et al. Global transcriptional start site mapping using dRNA-seq reveals novel antisense RNAs in *Escherichia coli*. *J Bacteriol.* 2014; 197(1): 18–28.

Thomason MK, Storz G. Bacterial antisense RNAs: how many are there, and what are they doing? *Annu Rev Genet.* 2010; 44: 167–88.

Tremblay PL, Summers ZM, Glaven RH, Nevin KP, Zengler K, Barrett CL, et al. A c-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in *Geobacter sulfurreducens* revealed by adaptive evolution. *Environ Microbiol.* 2011; 13(1): 13–23.

Tsai CH, Liao R, Chou B, Contreras LM. Transcriptional analysis of *Deinococcus radiodurans* reveals novel small RNAs that are differentially expressed under ionizing radiation. *Appl Environ Microbiol.* 2015; 81(5): 1754–64.

Tsai C, Liao R, Chou B, Palumbo M, Contreras M. Genome-Wide Analyses in Bacteria Show Small-RNA Enrichment for Long and Conserved Intergenic Regions. *J Bacteriol.* 2015; 197(1): 40–50.

Ueki T, DiDonato LN, Lovley DR. Toward establishing minimum requirements for extracellular electron transfer in *Geobacter sulfurreducens*. *FEMS Microbiol Lett.* 2017; (May): 1–7.

Van Assche E, Van Puyvelde S, Vanderleyden J, Steenackers HP. RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front Microbiol.* 2015; 6: 1–16.

Vogel J, Luisi BF. Hfq and its constellation of RNA. *Nat Rev Microbiol.* 2011; 9(8): 578–89.

Voigt K, Sharma CM, Mitschke J, Joke Lambrecht S, Voß B, Hess WR, et al. Comparative transcriptomics of two environmentally relevant cyanobacteria reveals unexpected transcriptome diversity. *ISME J.* 2014; 17:2056–68.

Wade JT, Grainger DC. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat Rev Microbiol.* 2014; 12(9): 647–53.

Wagner EGH, Romby P. Small RNAs in Bacteria and Archaea: Who They Are, What They Do, and How They Do It. *Advances in Genetics.* 2015. 133-208 p.

Waters LS, Storz G. Regulatory RNAs in bacteria. *Cell.* 2009; 136(4): 615–28.

Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, et al. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* 2010; 11(3): R31.

Weiss V, Medina-Rivera A, Huerta AM, Santos-Zavaleta A, Salgado H, Morett E, et al. Evidence classification of high-throughput protocols and confidence integration in RegulonDB. Database. 2013; bas059.

Yan B, Núñez C, Ueki T, Esteve-Núñez A, Puljic M, Adkins RM, et al. Computational prediction of RpoS and RpoD regulatory sites in *Geobacter sulfurreducens* using sequence and gene expression information. *Gene.* 2006; 384: 73–95.

Yang TH, Coppi M V, Lovley DR, Sun J. Metabolic response of *Geobacter sulfurreducens* towards electron donor/acceptor variation. *Microb Cell Fact.* 2010; 9(1): 90.

Zorgani MA, Quentin R, Lartigue M-F. Regulatory RNAs in the Less Studied Streptococcal Species: From Nomenclature to Identification. *Front Microbiol.* 2016; 7: 1161.

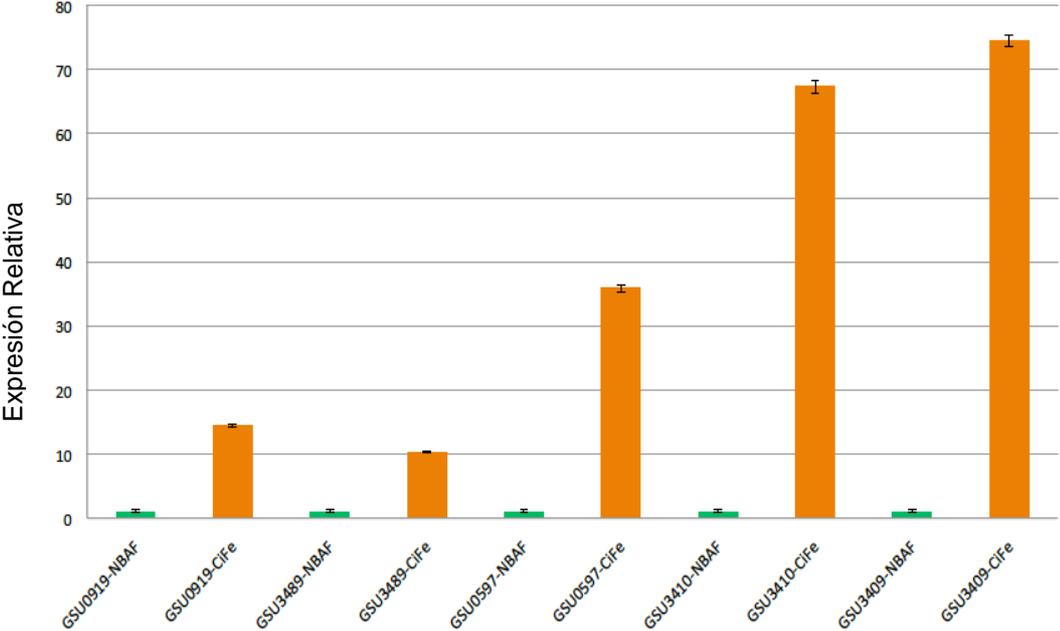
Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003; 31(13): 3406–15.

ANEXOS

Lista de oligonucleótidos utilizados en este trabajo. Los nombres entre paréntesis corresponden al sRNA que amplifican.

Nombre	Secuencia(5' → 3')	Experimento en el que fue utilizado
GSU0919Fw	GGATCGGCACACTCACATTC	Validación por q-PCR de experimentos RNaseq
GSU0919Rev	TACCCCTGAAAACCCCTCTG	Validación por q-PCR de experimentos RNaseq
GSU3489Fw	TTCTCATGATGCTCGGCTTT	Validación por q-PCR de experimentos RNaseq
GSU3489Rev	CTAGTGCGGTGCCCTTCCCT	Validación por q-PCR de experimentos RNaseq
GSU0597Fw	GCACTGATCATCGTGTCCCA	Validación por q-PCR de experimentos RNaseq
GSU0597Rev	GGCCTGATAAACAGTCCCTTG	Validación por q-PCR de experimentos RNaseq
GSU3410Fw	CAATGGTCTTGGGAGCGTA	Validación por q-PCR de experimentos RNaseq
GSU3410Rev	AGCCGAGGAAGATCCATACC	Validación por q-PCR de experimentos RNaseq
GSU3409Fw	CATCATGCTGATGGGATCG	Validación por q-PCR de experimentos RNaseq
GSU3409Rv	CAGTGAGTCGTGGCTCCTC	Validación por q-PCR de experimentos RNaseq
IG_7_F (gsusR02)	CGGTGAACGGACCAAGAT	Northern Blot y q-PCR
IG_7_R (gsusR02)	CTTGACGTACCAAGGAGTTACA	Northern Blot y q-PCR
IG_11_Fw (gsusR17)	GACAGGGTACCAAAGTACTAGC	Northern Blot , q-PCR y co-transcripción
IG_11_Rv (gsusR17)	CCCGGATAAGAACGTGAAC	Northern Blot y q-PCR
11RvCoTr (gsusR17)	TGACCGTGTCGAAGAACTTG	Co-transcripción
IG_14_Fw (gsusR40)	ACGATCCGGTTCGCCTTTTT	Northern Blot , q-PCR y co-transcripción
IG_14_Rv (gsusR40)	CTCGGGTCGATGGGTGTATT	Northern Blot y q-PCR
IG_14Cotr (gsusR40)	GTTCAGGATGTTGATCGGGTAG	Co-transcripción
IG_21_Fw (gsusR41)	GTATTACAATTGCGCGAGAAT	Northern Blot , q-PCR y co-transcripción
IG_21_Rv (gsusR41)	AAACTGGCTTCACGCTCTTT	Northern Blot y q-PCR
21RvCoTr (gsusR41)	GGAGAGGTGGAGAAACAGAAC	Co-transcripción
IG_47F (gsusR34)	GGGCTACGCGGAGAAA	Northern Blot y q-PCR
IG_47R (gsusR34)	GGACTCCCCTCCGAAA	Northern Blot y q-PCR
IG_60Fw (gsusR08)	TGGCCTTTGCTGGCTTAT	Northern Blot y q-PCR
IG_60RV (gsusR08)	CGACCAAGGGATCAATCCAATA	Northern Blot y q-PCR
IG_70_F (gsusR38)	ATTGTTCCCTGTCGGTGAAG	Northern Blot y q-PCR
IG_70_R (gsusR38)	AAGTCAACTCCTGCTAGG	Northern Blot y q-PCR
IG_160Fw (gsusR06)	GGATTTCAATTGCTTCGGGTTC	Northern Blot , q-PCR y co-transcripción
IG_160Rv (gsusR06)	TTCAGACGCTTGCCAGATATAC	Northern Blot y q-PCR
160CoTrRv (gsusR06)	TCGTGTTGCAGAAGATGAGAG	Co-transcripción
IG_202_F (gsusR45)	GACAATCTGATGCTTGACCTG	Northern Blot y q-PCR
IG_202_R (gsusR45)	CGCCTCCCGGAAGATATG	Northern Blot y q-PCR
IG_303F (gsusR27)	GTGGTATTCATGCAGGTATTCTAATG	Northern Blot y q-PCR
IG_303R (gsusR27)	GAAGGGAAAGCTCATCTATCC	Northern Blot y q-PCR
IG_404F (gsusR32)	GAATGGCGGGCGTCGTTG	Northern Blot y q-PCR
IG_404R (gsusR32)	CCTGGTTGCAATTTCCCATGA	Northern Blot y q-PCR
IG_721Fw (gsusR36)	CCGTTTCATTGCTTCGGT	Northern Blot , q-PCR y co-transcripción
IG_721Rv (gsusR36)	TCCTCGCCAGAGACAT	Northern Blot y q-PCR
721FwCoTr (gsusR36)	CGCACCGTAAACGAGATAG	Co-transcripción
IG21Rv (gsusR41)	GAAAGACGCTGAAGCCAGTTTCC	5'RACE
IG303fw (gsusR27)	GTGGTATTCATGCACGTATTCTAATG	5'RACE
IG303rv (gsusR27)	CGCGGTTACCGTGGTTGTC	5'RACE

Validación por el método de q-PCR de 5 de los genes con los valores de expresión diferencial mas altos obtenidos por en los datos de RNA-seq



RESEARCH LETTER – Physiology & Biochemistry

Global transcriptional start site mapping in *Geobacter sulfurreducens* during growth with two different electron acceptors

Getzabeth González¹, Aurora Labastida¹, Verónica Jiménez-Jacinto², Leticia Vega-Alvarado³, Maricela Olvera¹, Enrique Morett^{1,†} and Katy Juárez^{1,*}

¹Departamento de Ingeniería Celular y Biotatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Campus Morelos, Av. Universidad 2001, Cuernavaca Morelos, C.P. 62210, México,

²Unidad Universitaria de Secuenciación Masiva y Bioinformática, Instituto de Biotecnología, Universidad Nacional Autónoma de México Campus Morelos, Av. Universidad 2001, Cuernavaca Morelos, C.P. 62210, México and ³Centro de Ciencias Aplicadas y Desarrollo Tecnológico, Universidad Nacional Autónoma de México, Circuito exterior s/n, Ciudad Universitaria, Coyoacán, D.F., C.P. 04510, México

*Corresponding author: Departamento de Ingeniería Celular y Biotatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Campus Morelos, Av. Universidad 2001, Cuernavaca Morelos, C.P. 62210, México. Tel: +52 (777) 329-1605; E-mail: katy@ibt.unam.mx

†Present address: Instituto Nacional de Medicina Genómica. Periférico Sur No. 4809, Col. Arenal Tepepan, Delegación Tlalpan, D.F. C.P. 14610, México.

One sentence summary: The present work is a detailed study of the transcriptional landscape of *Geobacter sulfurreducens* at single nucleotide resolution in two relevant growth conditions.

Editor: Robert Gunsalus

ABSTRACT

Geobacter sulfurreducens is an anaerobic soil bacterium that is involved in biogeochemical cycles of elements such as Fe and Mn. Although significant progress has been made in the understanding of the electron transfer processes in *G. sulfurreducens*, little is known about the regulatory mechanisms involved in their control. To expand the study of gene regulation in *G. sulfurreducens*, we carried out a genome-wide identification of transcription start sites (TSS) by 5'RACE and by deep RNA sequencing of primary mRNAs in two growth conditions. TSSs were identified along *G. sulfurreducens* genome and over 50% of them were located in the upstream region of the associated gene, and in some cases we detected genes with more than one TSS. Our global mapping of TSSs contributes with valuable information, which is needed for the study of transcript structure and transcription regulation signals and can ultimately contribute to the understanding of transcription initiation phenomena in *G. sulfurreducens*.

Keywords: *Geobacter sulfurreducens*; fumarate and Fe(III) reduction; RNA; transcriptional start sites

INTRODUCTION

Geobacter species are delta-Proteobacteria abundant in anaerobic soils and aquatic sediments. These microorganisms have

the ability to exchange electrons with the extracellular environment (Lovley 2002). *Geobacter* species are able to carry out dissimilatory reduction of Fe(III), Mn(IV) and other metals

Received: 1 March 2016; Accepted: 13 July 2016

© FEMS 2016. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

coupled to organic matter degradation (Lovley, Holmes and Nevin 2004). This relevant feature of these bacteria has been exploited for bioremediation of soils contaminated with organic or metal pollutants. *Geobacter* species have successfully been employed for in situ bioremediation of U(VI) (Holmes et al. 2002) and aromatic hydrocarbons (reviewed in Lovley et al. 2011), as well as in the reduction of chlorinated contaminants (Strycharz et al. 2008). Besides their potential role in cleaning up the environment, *Geobacter sulfurreducens* has been employed in the production of electricity through the extracellular transfer of electrons produced from the degradation of organic matter to electrodes (Bond and Lovley 2003). The extracellular electron transfer mechanism has extensively been investigated in *G. sulfurreducens* (Lovley et al. 2011), and previous studies showed that cytochromes located both at the inner and the outer membrane, as well as in the periplasm, are involved in this process (Afkar et al. 2005; Butler, Young and Lovley 2010). In addition to cytochromes, *G. sulfurreducens* produces conductive pili to transfer electrons to different acceptors including Fe(III) oxides (Reguera et al. 2005; Malvankar and Lovley 2014). Although models of metabolism and electron transfer have been defined in *G. sulfurreducens*, the regulatory networks involved in their genetic control are not yet well understood. Therefore, identification of regulatory elements that control the expression of the relevant genes involved in electron transfer is a key step for the development of new strategies of bioremediation and electricity production in which *G. sulfurreducens* participates (Lovley 2002). Bioinformatics approaches (Krushkal et al. 2007) and genome-wide studies (Qiu et al. 2010; Ueki and Lovley 2010) have been followed to identify regulatory elements involved in extracellular electron transfer mechanism; however, information corresponding to different growth conditions as well as the exploration of different technical approaches is still needed.

In recent years, global mapping of transcription start sites (TSSs) in bacteria employing high-throughput sequencing has proven to be a powerful tool for studying microbial transcriptomes (Mendoza-Vargas et al. 2009; Sorek and Cossart 2010; Filiatrault et al. 2011; Salgado et al. 2013; Sharma and Vogel 2014). The identification of TSSs has not only contributed to promoter and transcript structure characterization, but it has also revealed a high complexity in prokaryotic transcriptional organization. This includes TSSs at internal and also at some antisense positions, as well as transcription in non-coding regions. In addition, new protocols aimed at discriminating processed (mRNA 5^{prime}: -P) from primary transcripts (mRNA 5'-PPP) have been described. These strategies are intended to achieve a more accurate genome-wide identification of TSS by limiting the presence of degraded mRNA species to avoid transcriptional noise signals (Sharma et al. 2010; Salgado et al. 2013).

In this work, we carried out a genome-wide mapping of TSSs in *G. sulfurreducens* in two growth conditions: fumarate or Fe(III) as electron acceptors and acetate as electron donor. The use of soluble Fe(III) as a final electron acceptor has served as a model to simulate the Fe(III)-oxide conditions in anoxic soils and subsurface environment, where *Geobacteraceae* are the predominant Fe(III)-reducing organisms (Ding et al. 2006), while fumarate has been used as an alternative electron acceptor, which improve the growth of *G. sulfurreducens* in the laboratory (Estevez-Núñez et al. 2005). By studying these conditions, we provide data that are relevant for the study of genes involved in electron transfer in this environmentally and biotechnologically relevant bacteria.

MATERIALS AND METHODS

Bacterial strains and culturing conditions

Geobacter sulfurreducens wild-type strain PCA (ATCC51573) (Cacavo et al. 1994) was routinely cultured anaerobically at 30°C with acetate as the electron donor and either fumarate (40 mM) or Fe(III)-citrate (50 mM) as electron acceptors, as previously described (Coppi et al. 2001). RNA Later (Ambion) was added to cultures and cells were harvested at exponential phase. Bacterial pellets were flash frozen and stored at -80°C.

5' RACE

5' Rapid Amplification of cDNA Ends (5'RACE) experiments were performed according to the method described by Mendoza-Vargas et al. (2009). Briefly, total RNA from *G. sulfurreducens* wild-type strain was isolated from the two growth conditions mentioned above. RNA was treated with tobacco acid pyrophosphatase (TAP, Epicenter) to generate 5'-monophosphate ends before ligation to a 5' RNA adapters (5'-GUUCAGAGUUCUACAGUCCGACGAUC-3') with T4 RNA Ligase 1 (ssRNA Ligase, New England Biolabs). For the construction of cDNA libraries, a primer with a random sequence at the 3' end (5'-CAAGCAGAAGACGGCATAACGANNNNNN-3') was used for the generation of the cDNA in a reverse transcription reaction employing SuperScript III (Invitrogen, Carlsbad, CA, USA). Finally, a PCR reaction was performed using specific primers for the genes of interest. PCR reactions were subjected to 8% polyacrylamide gel electrophoresis (PAGE), bands were excised and purified from the gel. Purified PCR products were sequenced by the Sanger method in an Applied Biosystems 3100 Genetic Analyzer/ABI PRISM instrument. The sequences were aligned with *G. sulfurreducens* genome, and the putative TSSs were identified as the first nucleotide immediately adjacent to the 5' RNA adapter.

RNA preparation, library construction and sequencing of 5' ends

Total RNA from cultures in the two growth conditions described above was isolated with RNeasy minicolumns (QIAGEN Inc., CA, USA). Residual DNA was digested by treatment with DNaseI (Fermentas, Life Technologies, MA, USA) following manufacturer's instructions. RNA was subsequently cleaned and concentrated with RNeasy minicolumns (QIAGEN Inc., CA, USA) and rRNA was removed using RiboMinus kit (Invitrogen). Quality and concentration of RNA were assessed by Bioanalyzer 2100, Chip RNAnano (Agilent Technologies, Santa Clara, CA, USA). The enrichment for 5' PPP mRNA was carried out as described in previous reports (Gama-Castro et al. 2011; Salgado et al. 2013). Briefly, a biotinylated RNA adapter was ligated to 5'-P RNA with T4 RNA ligase (New England BioLabs, MA, USA), and the biotinylated molecules were removed. Then to allow for the ligation of the RNA-seq library 5' adapter, the enriched 5'-PPP mRNA was converted to its 5'-P form by treatment with TAP. 5'-P mRNA was reverse transcribed with SuperScript II Reverse Transcriptase (INVITROGEN, Carlsbad, CA, USA) and then PCR amplified for 10 cycles. During the retrotranscription and PCR steps, the Illumina platform-specific adapters were added to the sample. Illumina libraries (transcript sizes between 100 and 300 nt) were purified after electrophoresis using 6% TBE PAGE gels. Library quality and concentration were evaluated using the Agilent 2100 Bioanalyzer DNA 1000 assay. Preparation of libraries was performed in the UUSMD (UNAM, México) using the TruSeq Sample Prep

kit (Illumina, San Diego, CA, USA) following the manufacturer's instructions and was sequenced using an Illumina Genome Analyzer Iix (36 bp single-read run).

Bioinformatics analyses

All reads obtained were mapped to the *G. sulfurreducens* reference genome (NC.002939) using the Bowtie2 aligner program (Langmead and Salzberg 2012) with the following parameters: zero mismatches in a seed of length 29, taking the best alignment of multiple hits. From the aligned reads, low-quality reads (Mean Phred quality lower than 20) and reads corresponding to rRNA were eliminated using Perl custom scripts. For each experiment and from the total number of genomic 5' positions, the mean value of the reads was calculated. Only positions with read numbers equal to or above the mean were considered for subsequent analyses.

Since it is not uncommon that reads map at adjacent positions, when this happened these positions were grouped into clusters. Then, TSSs (both single positions and clusters) were assigned to different categories depending on their location with respect to genes: (a) UpSense, from -150 bp to 50 bp up to the start of the gene; (b) UpSenseExtended, upstream to gene over 150 bp; (c) InSense, into the coding region of the gene; (d) UpAntiSense, in the upstream region of the gene but in the opposite orientation; (e) UpAntiSenseExtended, upstream to gene over 150 bp and in the opposite direction; and (f) InAntisense, into the coding region of the gene but in the opposite orientation. Data were visualized using Integrated Genome Browser (Nicol et al. 2009).

RESULTS AND DISCUSSION

The aim of this study was to identify TSSs in the most commonly used growth conditions in the laboratory for a more accurate annotation of genes and regulatory elements in *Geobacter sulfurreducens*.

In a first approach, we performed a direct mapping of TSSs by 5'RACE experiments of genes related to extracellular electron transfer, Fe(III) reduction and with regulatory functions. No more than 30 TSSs have been identified by traditional methods, as primer extension or 5' RACE in *G. sulfurreducens*. Here, we expanded this number with the identification of TSSs associated with 79 genes in two growth conditions, some of which correspond to cytochromes and global regulators (Table 1).

As a second approach for the identification of TSSs, we carried out 5' RNA end mapping by a high-throughput method on cultures grown with soluble Fe(III) or fumarate as the final electron acceptors (Fig. 1). As mentioned in the bioinformatics analyses section, in order to remove positions with low number of reads that might correspond to transcriptional noise, positions with a number of reads below the mean value were discarded. Although the number of positions decreased significantly, the remaining positions conserved the highest percentage of detected reads in the experiments (Table S1, Supporting Information). With these criteria, we identified TSSs associated with 2196 of the 3711 genes of this microorganism.

As described in the Materials and Methods section, 5' RNA ends in adjacent positions detected by global mapping were grouped into clusters, which could be considered as a single TSS. 2479 and 1247 clusters were detected during Fe(III) and fumarate reduction, respectively (Supplementary Data, Supporting Information). Although the aim of this work was not to report differential expression of TSSs, we found that some TSSs

were only detected in one of the two growth conditions and 389 were present in both growth conditions (Fig. 2A). Notwithstanding the number of matched positions in both culture conditions is not too high, it should be noted that intersected TSSs have the highest number of reads (Fig. 2B).

In a previous work, Qiu et al (2010) reported the identification of 1374 TSSs during fumarate reduction in *G. sulfurreducens*, 26 of those were located at exactly the same positions as ours. However, when we compared the immediate vicinity of our TSSs with those reported by Qiu et al (2010), the coincidences increased considerably to 603 (Fig. S1, Supporting Information), indicating that TSSs mapping is prone to minor shifts in chromosomal positions, likely depending on the methodology, laboratory or a specific strain used. Also, we sought for co-occurrence of TSSs with promoters derived from RNAP-binding regions reported in Qiu et al (2010), and we found 668 TSSs with a promoter region either to RpoN and RpoD. In this regard and according to RegulonDB, which is one of the most complete data collection of TSSs for *Escherichia coli*, a TSS could be considered with 'strong evidence' when its position is supported by at least two independent types of evidences or experiments (Weiss et al. 2013). These TSSs together with those that were detected by 5'RACE should be assigned as TSSs with 'strong evidence' since they were validated by independent experiments.

As an example of TSSs identified by two different strategies, Fig. 3 shows a TSS identified for the *pilB* gene by genome-wide mapping and 5' RACE. In this example, besides the TSS identified at the UpStream region, a TSS corresponding to the InAntisense category was identified (Fig. 3) in both growth conditions at the same position and with a similar number of reads (9 and 7, respectively).

Another example of TSS identified by two experimental approaches and therefore assigned as with strong evidence is shown in Fig. 4. This TSS is assigned to GSU3421 gene and corresponds to one of the two Crp/Fnr homologs present in *Geobacter* genome (Krushkal et al. 2007). Particularly in *E. coli*, FNR is the master transcriptional regulator of the transition between aerobic and anaerobic growth (Tolla and Savageau 2010) but in *G. sulfurreducens* the regulatory role of this gene is being studied.

All TSSs detected in this were classified depending on their genetic context (Fig. 5). Based on this classification, we found that 51% and 64% of the TSSs identified in the two growth conditions belong to the UpSense category, which is the region where the canonical TSSs are located. Interestingly, in the crude data before making the cut for the mean value of reads, the highest number of TSSs were located in the InSense region (more than 50% in both growth conditions), suggesting that positions with reads below the mean corresponded to spurious transcription. We noticed that when TSSs detected in both growth conditions were categorized, the UpSense elements were increased to 66%. Data showed that when Fe(III) reduction was carried out, the number of TSS belonging to the InSense category increased; nevertheless, more experiments are needed to define whether or not these 5'ends are real TSSs or might correspond to mRNA degradation.

In recent years, it has been reported that antisense transcription is highly abundant and conspicuous in bacteria, suggesting that RNA complementary to mRNA has important functions in gene regulation and genome evolution (Sharma et al. 2010; Lasa et al. 2011; Kim et al. 2012; Raghavan, Sloan and Ochman 2012; Thomason et al. 2014; Wade and Grainger 2014). However, the biological significance of the antisense transcripts remains unclear. In a previous work in *E. coli*, we reported that antisense transcription is not reproducible through different experiments,

Table 1. Examples^a of TSS identified by 5' RACE methodology in two different growth conditions. (*) indicates the exact position in *G. sulfurreducens* genome of the TSS identified.

Locus	Gene name	Product	Position*	Electron acceptor
GSU2811	<i>hsc</i>	cytochrome c	3089557	Soluble Fe(III)
GSU0618	<i>omcE</i>	cytochrome c	654676	Soluble Fe(III)
GSU2076	<i>omcZ</i>	cytochrome c	2283005	Soluble Fe(III)
GSU1491	<i>pilB</i>	Pilus biogenesis protein	1634818	Soluble Fe(III)
GSU2043	<i>pilD</i>	Type 4 prepilin-like protein leader peptide	2238772	Soluble Fe(III)
GSU0108		ATP synthase F0, B' subunit, putative	123597	Soluble Fe(III)
GSU0145	<i>recA</i>	RecA	161047	Soluble Fe(III)
GSU0174		Acetyl-CoA hydrolase transferase	191212	Soluble Fe(III)
GSU0305	<i>hypB</i>	Hydrogenase accessory protein	337425	Soluble Fe(III)
GSU0360	<i>ompA</i>	Domain protein	391321	Soluble Fe(III)
GSU1165	<i>ptsP</i>	Phosphoenolpyruvate-protein phosphotransferase	1271392	Soluble Fe(III)
GSU1885	<i>hprK</i>	HPr(Ser) kinase/phosphatase	2064353	Soluble Fe(III)
GSU2751	<i>dcuB</i>	C4-dicarboxylate transporter, anaerobic	3027503	Soluble Fe(III)
GSU1887	<i>rpoN</i>	RNA polymerase sigma-54 factor	2066647	Soluble Fe(III)
GSU1903		3-Isopropylmalate dehydratase	2082084	Soluble Fe(III)
GSU1379	<i>fur</i>	Ferric uptake regulation protein	1508348	Soluble Fe(III)
GSU1525		RNA polymerase sigma-70 factor family	1671174	Soluble Fe(III)
GSU3304	<i>lamB</i>	Porin family protein, putative	3627454	Soluble Fe(III)
GSU0670		cytochrome c family protein	704740	Soluble Fe(III)
GSU2504		cytochrome c family protein	2759927	Soluble Fe(III)
GSU2606	<i>aroA</i>	3-Phosphoshikimate 1-carboxyvinyltransferase	2875030	Soluble Fe(III)
GSU3401		Branched-chain amino acid ABC transporter	3744562	Soluble Fe(III)
GSU2809		Transcriptional regulator Fur family	3088832	Soluble Fe(III)
GSU2645		cytochrome c family protein	2916894	Soluble Fe(III)
GSU1469		Keto/oxoacid ferredoxin oxidoreductase	1611302	Soluble Fe(III)
GSU2808		cytochrome c family protein	3088036	Soluble Fe(III)
GSU2513		cytochrome c family protein	2772585	Soluble Fe(III)
GSU1809	<i>ftsH-2</i>	Cell division protein FtsH	1975677	Soluble Fe(III)
GSU3445	<i>nuoA-2</i>	NADH dehydrogenase I, A subunit	3791574	Soluble Fe(III)
GSU1761		cytochrome c family protein	1922942	Soluble Fe(III)
GSU1345		Rrf2 family protein	1472552	Fumarate
GSU2645		cytochrome c family protein	2916894	Fumarate
GSU2100	<i>katG</i>	Catalase/peroxidase	2308752	Fumarate
GSU1469		Keto/oxoacid ferredoxin oxidoreductase	1611302	Fumarate
GSU2808		cytochrome c family protein	3088036	Fumarate
GSU0298		Sigma-54 dependent DNA-binding response regulator	327551	Fumarate
GSU2513		cytochrome c family protein	2772585	Fumarate
GSU1908		Phosphatidylserine decarboxylase, putative	2087254	Fumarate
GSU0065		Cytidine/deoxycytidylate deaminase family protein	81353	Fumarate
GSU1711		Integrative genetic element Gsu5, resolvase	1877510	Fumarate
GSU1809	<i>ftsH-2</i>	Cell division protein FtsH	1975677	Fumarate
GSU3445	<i>nuoA-2</i>	NADH dehydrogenase I, A subunit	3791574	Fumarate
GSU1761		cytochrome c family protein	1922942	Fumarate
GSU3421	<i>fnr-1</i>	Transcriptional regulator	3767006	Fumarate

^aThe complete list of TSSs is attached in Supplementary File.

the positions in this category are the least conserved among different biological replicates (Salgado et al. 2013). In agreement with this observation, in this study only about 4% of TSSs corresponded to the Antisense categories, and when we grouped the TSSs detected in both experiments, the TSSs that belong to InAntisense category decreased further (Fig. 5B). Additionally, we found that four antisense TSSs showed a promoter region previously reported (Qiu et al. 2010); despite this low number, to date there is no related information for antisense regulation in *G. sulfurreducens* and these examples could be interesting for later studies.

We also observed that some genes have more than one TSS assigned and in some cases they are present along the genes and operons. The presence of multiple sites had al-

ready been previously reported in *G. sulfurreducens* and in others organisms (Leang and Lovley 2005; Mendoza-Vargas et al. 2009). Although it is not known if these transcripts have any biological role or are the product of non-functional promoters, we found an interesting case: the cytochrome type c *ppcD*. Yan et al. (2006) reported three putative promoters regions with a RpoS signal and we found TSSs associated with these promoter regions, three detected during fumarate reduction and one TSSs in Fe(III) reduction suggesting that *ppcD* expression is controlled by different TSSs depending growth condition.

In summary, this work presents a detailed study of the transcriptional landscape of *G. sulfurreducens* at single nucleotide resolution in two relevant growth conditions. With the

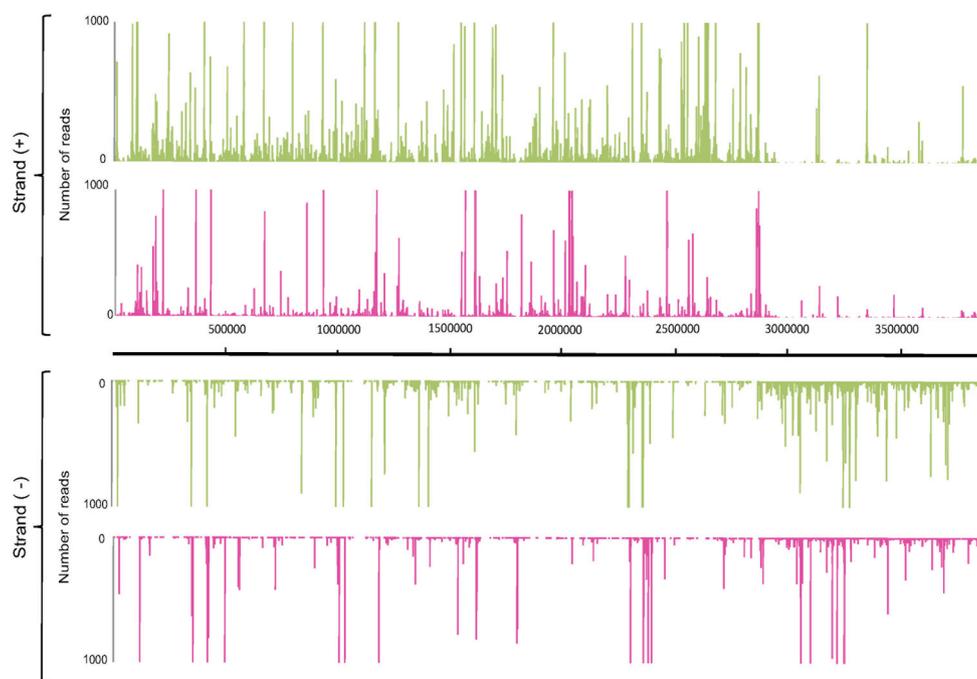


Figure 1. Visualization of TSSs identified in two growth conditions. The X-axis corresponds to the genomic coordinates of *G. sulfurreducens* and the Y-axis corresponds to the number of reads per position. Pink graphs correspond to fumarate as final electron acceptor and green graphs to Fe(III) as final acceptor.

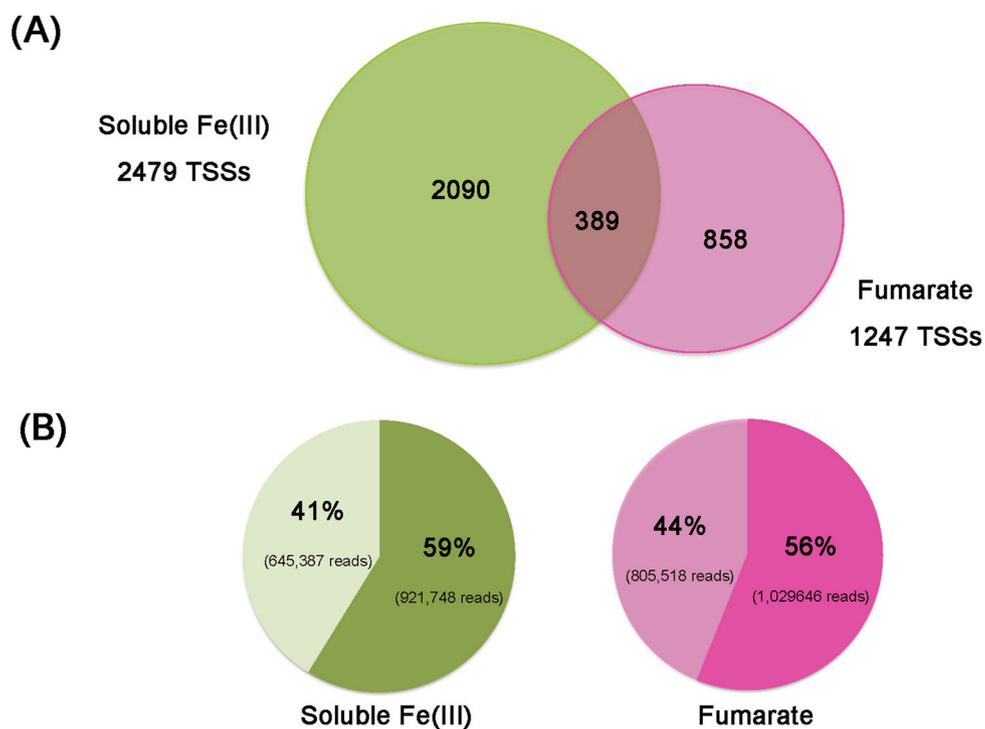


Figure 2. (A) Venn diagram showing the intersection of TSSs grouped in clusters. (B) Percentage of reads corresponding to TSSs present in only one or both growth conditions (light or solid tone, respectively).

information generated in this study, it would be possible to identify genetic features and regulatory elements not yet described for this organism. The identification of TSSs profile in *G. sulfurreducens* provides new biological information that is being used for the identification of new regulatory elements such as non-coding sRNA (Gonzalez et al. 2015) as well in the char-

acterization of transcriptional regulators involved in electron transfer (Hernández-Eligio et al. 2016). The genetic characteristics described in this work must be taken into account in further efforts for the understanding and the optimization of metabolic activities involved in bioremediation and electricity production.

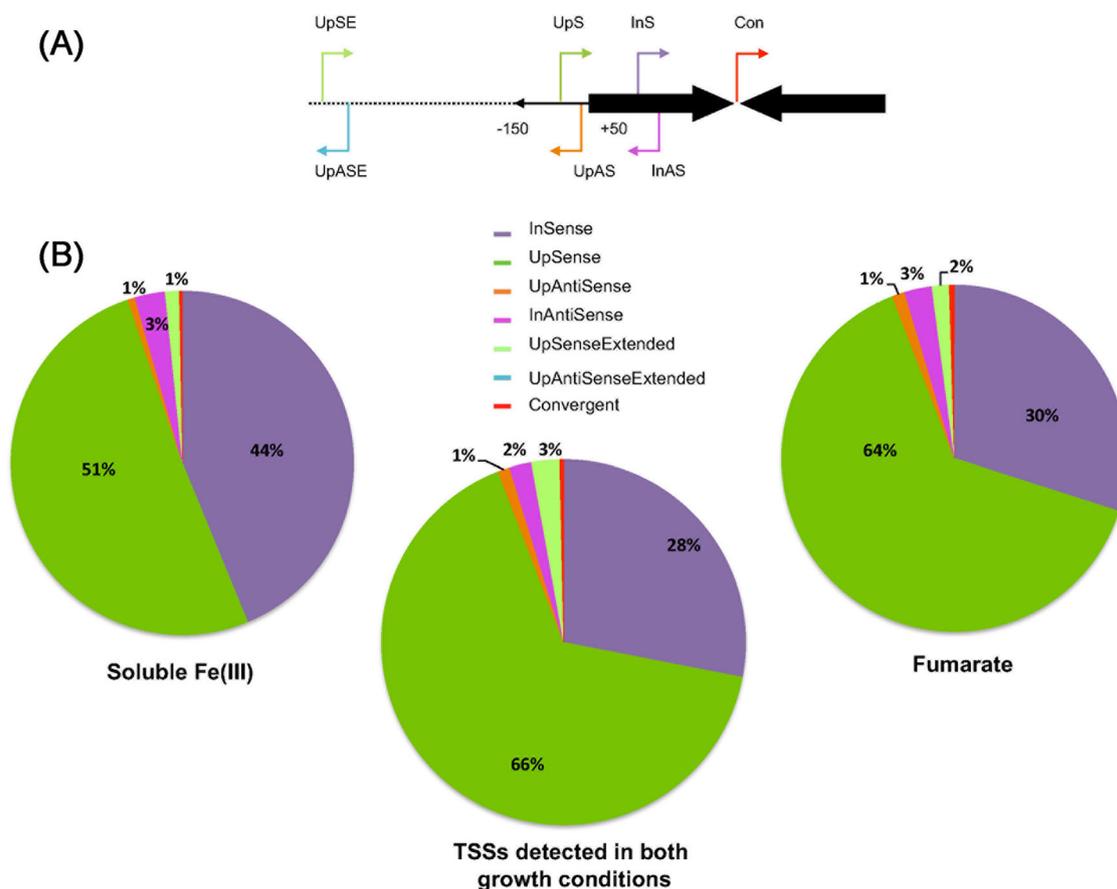


Figure 5. Categorization of the TSSs identified. (A) Schematic representation of TSSs categories used in this study: UpSense is located from -150 pb to $50+$ from the translation start site. InSense is within coding region. UpSenseExtended is located in a upstream region in a distance greater than -150 pb of the gene. UpAntisense, InAntisense and UpAntiSenseExtended are in the same regions but in the opposite strand of the annotated gene. Convergent is located between convergent genes. (B) TSSs corresponding to the two growth conditions and the TSSs detected in both conditions.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSLE online.

ACKNOWLEDGEMENTS

We thank Instituto de Biotecnología-UNAM for access to their computer cluster, Prof. Derek Lovley (UMASS) for financial support to this work, Leticia Olvera for technical support and Ricardo Grande, Unidad Universitaria de Secuenciación Masiva de DNA, UNAM, for sequencing services.

FUNDING

This research was supported in part by the Office of Science (BER), US Department of Energy (DOE) and Cooperative Agreements No. DEFC02-02ER63446PI Derek Lovley. DGAPA, Universidad Nacional Autónoma de México (UNAM) IN-209312, Consejo Nacional de Ciencia y Tecnología (CONACYT), Grants No. 179684 and No. 179997. GG was the recipient of a PhD scholarship from CONACYT.

Conflict of interest. None declared.

REFERENCES

- Afkar E, Reguera G, Schiffer M et al. A novel Geobacteraceae-specific outer membrane protein J (OmpJ) is essential for electron transport to Fe(III) and Mn(IV) oxides in *Geobacter sulfurreducens*. *BMC Microbiol* 2005;5:41–52.
- Bond DR, Lovley DR. Electricity production by *Geobacter sulfurreducens* attached to electrodes. *Appl Environ Microb* 2003;69:1548–55.
- Butler J, Young N, Lovley DR. Evolution of electron transfer out of the cell: comparative genomics of six *Geobacter* genomes. *BMC Genomics* 2010;11:40–52.
- Caccavo F, Lonergan DJ, Lovley DR et al. Oxidizing dissimilatory metal-reducing microorganism. *Appl Environ Microb* 1994;60:3752–9.
- Coppi MV, Leang C, Sandler SJ et al. Development of a genetic system for *Geobacter sulfurreducens*. *Appl Environ Microb* 2001;67:3180–7.
- Ding R, Hixson KK, Giometti CS et al. The proteome of dissimilatory metal-reducing microorganism *Geobacter sulfurreducens* under various growth conditions. *Biochim Biophys Acta* 2006;1764:1198–206.
- Estevez-Núñez A, Núñez C, Lovley DR. Preferential reduction of Fe(III) over fumarate by *Geobacter sulfurreducens*. *J Bacteriol* 2005;186:2897–9.

- Filiatrault MJ, Stodghill PV, Myers CR et al. Genome-wide identification of transcriptional start sites in the plant pathogen *Pseudomonas syringae* pv. tomato str. DC3000. *PLoS One* 2011;**6**:e29335.
- Gama-Castro S, Salgado H, Peralta-Gil M et al. RegulonDB version 7.0: transcriptional regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (Gensor Units). *Nucleic Acids Res* 2011;**39**:D98–105.
- Gonzalez G, Labastida A, Jimenez V et al. Genes involved in electron transfer mechanism in *G. sulfurreducens*. In: Poggi-Varaldo HM (ed.). *Environmental Biotechnology and Engineering-2014*, Vol 3. Mexico: Cinvestav, 2015, 1269–76.
- Hernández-Eligio JA, Andrade A, Soto L et al. The unphosphorylated form of the PilR two-component system regulates pilA gene expression in *Geobacter sulfurreducens*. *Environ Sci Pollut Res* 2016, DOI: 10.1007/s11356-016-6192-5.
- Holmes DE, Finneran KT, Neil RAO et al. Enrichment of members of the family *Geobacteraceae* associated with stimulation of dissimilatory metal reduction in uranium-contaminated aquifer sediments. *Appl Environ Microb* 2002;**68**:2300–6.
- Kim D, Hong JSJ, Qiu Y et al. Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella pneumoniae* by genome-wide transcription start site profiling. *PLoS Genet* 2012;**8**:e1002867.
- Krushkal J, Yan B, DiDonato LN et al. Genome-wide expression profiling in *Geobacter sulfurreducens*: identification of Fur and RpoS transcription regulatory sites in a relGsu mutant. *Funct Integr Genomic* 2007;**7**:229–55.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9.
- Lasa I, Toledo-Arana A, Dobin A et al. Genome-wide antisense transcription drives mRNA processing in bacteria. *P Natl Acad Sci USA* 2011;**108**:20172–7.
- Leang C, Lovley DR. Regulation of two highly similar genes, omcB and omcC, in a 10 kb chromosomal duplication in *Geobacter sulfurreducens*. *Microbiology* 2005;**151**:1761–7.
- Lovley DR. Analysis of the genetic potential and gene expression of microbial communities involved in the *in situ* bioremediation of uranium and harvesting electrical energy from organic matter. *OMICS* 2002;**6**:331–9.
- Lovley DR, Holmes DE, Nevin KP. Dissimilatory Fe(III) and Mn(IV) Reduction. *Adv Microb Physiol* 2004;**49**:219–86.
- Lovley DR, Ueki T, Zhang T et al. *Geobacter*: the microbe electric's physiology, ecology, and practical applications. *Adv Microb Physiol* 2011;**59**:1–100.
- Malvankar NS, Lovley DR. Microbial nanowires for bioenergy applications. *Curr Opin Biotechnol* 2014;**27**:88–95.
- Mendoza-Vargas A, Olvera L, Olvera M et al. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One* 2009;**4**:e7526.
- Nicol JW, Helt GA, Blanchard SG et al. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* 2009;**25**:2730–1.
- Qiu Y, Cho BK, Park YS et al. Structural and operational complexity of the *Geobacter sulfurreducens* genome. *Genome Res* 2010;**20**:1304–11.
- Raghavan R, Sloan DB, Ochman H. Antisense transcription is pervasive but rarely conserved in enteric bacteria. *mBio* 2012;**3**:1–7.
- Reguera G, McCarthy KD, Mehta T et al. Extracellular electron transfer via microbial nanowires. *Nature* 2005;**43**:1098–101.
- Salgado H, Peralta-Gil M, Gama-Castro S et al. RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res* 2013;**41**:D203–13.
- Sharma CM, Hoffmann S, Darfeuille F et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 2010;**464**:250–5.
- Sharma CM, Vogel J. Differential RNA-seq: the approach behind and the biological insight gained. *Curr Opin Microbiol* 2014;**19**:97–105.
- Sorek R, Cossart P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet* 2010;**11**:9–16.
- Strycharz SM, Woodard TL, Johnson JP et al. Graphite electrode as a sole electron donor for reductive dechlorination of tetrachlorethene by *Geobacter lovleyi*. *Appl Environ Microb* 2008;**74**:5943–7.
- Thomason MK, Bischler T, Eisenbart SK et al. Global transcriptional start site mapping using dRNA-seq reveals novel antisense RNAs in *Escherichia coli*. *J Bacteriol* 2014;**197**:18–28.
- Tolla DA, Savageau MA. Regulation of aerobic-to-anaerobic transitions by the FNR cycle in *Escherichia coli*. *J Mol Biol* 2010;**397**:893–905.
- Ueki T, Lovley DR. Genome-wide gene regulation of biosynthesis and energy generation by a novel transcriptional repressor in *Geobacter* species. *Nucleic Acids Res* 2010;**38**:810–21.
- Wade JT, Grainger DC. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat Rev Microbiol* 2014;**12**:647–53.
- Weiss V, Medina-Rivera A, Huerta AM et al. Evidence classification of high-throughput protocols and confidence integration in RegulonDB. *Database* 2013;**2013**:bas059.
- Yan B, Núñez C, Ueki T et al. Computational prediction of RpoS and RpoD regulatory sites in *Geobacter sulfurreducens* using sequence and gene expression information. *Gene* 2006;**384**:73–95.