UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA GERAL
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

TESE DE DOUTORADO

Diversidade filogenética e funcional do microbioma, com ênfase nos genes envolvidos no metabolismo do arsênio, de sedimento impactado por mineração

ORIENTADA: PATRÍCIA COSTA LIMA DA SILVA

ORIENTADORA: ANDRÉA MARIA AMARAL NASCIMENTO

CO-ORIENTADOR: EDMAR CHARTONE DE SOUZA

BELO HORIZONTE

DEZEMBRO – 2014

PATRÍCIA COSTA LIMA DA SILVA

TESE DE DOUTORADO

Diversidade filogenética e funcional do microbioma, com ênfase nos genes envolvidos no metabolismo do arsênio, de sedimento impactado por mineração

Tese de doutorado apresentada ao curso de pós-graduação do departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, como requisito a obtenção do título de Doutora em Genética.

ORIENTADORA: ANDRÉA MARIA AMARAL NASCIMENTO
CO-ORIENTADOR: EDMAR CHARTONE DE SOUZA

BELO HORIZONTE

DEZEMBRO – 2014

Agradecimento Especial

Dedico este trabalho ao meu marido Marco e ao meu filho Artur que foram a minha força e perseverança, me fazendo sorrir todos os dias. Aos meus pais pelo exemplo de vida e por sempre me apoiarem. Amo vocês!

AGRADECIMENTOS

Agradeço a Deus, por me dar força, saúde e perseverança para continuar indo em frente.

Ao meu filho Artur que alegra meus dias com seus beijos e sorrisos. Tudo se iluminou quando você nasceu! Obrigada por ter me escolhido para ser sua mãe! Ao meu marido Marco, meu companheiro de todas as horas, por sempre me apoiar e me incentivar. Eu te amo muito!

Aos meus pais Marco Antônio e Olga, por vocês serem o meu exemplo de luta e resignação, por toda paciência, dedicação e amor. Amo vocês demais! Se eu estou aqui hoje, devo a vocês!

Aos meus irmãos Carlos e Adriana, à minha cunhada Chirlei e meu cunhado Eduardo, por todo o carinho e amizade. Às minhas sobrinhas Maria Clara e Isabella por fazerem os meus sábados mais felizes.

Aos meus sogros, Janete e Flamarion, e a minha cunhada Carol, por sempre apoiar e incentivar meu trabalho. Mais do que isso, por se importarem comigo como se fossem meus pais, por cuidarem do meu filho nos momentos em que eu não podia estar ao lado dele. Eu nunca vou esquecer a gratidão que tenho por amarem tanto o Artur. Muito obrigada!

Aos meus tios que tanto torceram por mim, Paulo, Ângela, Nayra, Imaculada, obrigada!

À minha orientadora e amiga Andréa por toda paciência, dedicação e sabedoria. Muito obrigada por ter acreditado em mim e por ser meu exemplo de profissional. Sou uma pessoa de sorte por ter encontrado em você mais do que uma orientadora, e sim uma amiga!

Ao professor Edmar Chartone de Souza pela dedicação e conhecimentos passados.

Aos meus amigos do LGM - Paulo, Raiana, Alexandre, Luciana, Cris, Marcelo, Maria Luiza, Barbara, Mercedes e Ana Paula - pelo companheirismo. Um agradecimento especial á Mariana e a Magna, pela amizade e por todos os momentos alegres que compartilhamos no laboratório. Adoro vocês!

À professora Mônica Bucciarelli Rodriguez pelo carinho, amizade, puxões de orelha e conhecimentos passados. Gosto demais de você!!!!

À professora Adlane Villas-Boas, por me abrir as portas do laboratório, pela amizade e conhecimentos passados.

Às técnicas Paixão e Maria Rosa pela paciência, amizade e pelas contribuições essenciais para o desenvolvimento dos experimentos e na preparação de materiais.

Aos parceiros de outros laboratórios, Flávio e Ana (Laboratório de Parasitologia - Fiocruz) e Laura e Guilherme (CEBIO - Fiocruz), pela grande contribuição na execução deste trabalho.

À FAPEMIG, CNPq e a CAPES pelo apoio financeiro.

LISTA DE FIGURAS

INTRODUÇÃO

CAPÍTULO 1

Figure 3 - Evolutionary relationships of AsV-resistant bacteria (MS-AsV) 16S rRNA sequences. A total of 40 nucleotide sequences and 721 sites were analyzed. The phylogeny was reconstructed by maximum likelihood and TrN+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green. Different background colors highlight three well-supported clades: *Firmicutes*, *Actinobacteria*, and *Proteobacteria*. *Thermodesulfobacteria* was used as outgroup.

Figure 4 - Evolutionary relationships of arsC sequences. A total of 48 nucleotide sequences and 352 sites were analyzed. The phylogeny was reconstructed by maximum likelihoodand TrN+I+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green, and operational taxonomic unities (OTUs) from As gene clone libraries in blue. Different background colors highlight *Actinobacteria* and three *Proteobacteria* classes – *Gamma-*, *Beta*, and *alpha-proteobacteria*.

Figure 5 - Evolutionary relationships of arrA sequences. A total of 47 nucleotide sequences and 242 sites were analyzed. The phylogeny was reconstructed by maximum likelihoodand GTR+I+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green, and operational taxonomic unities (OTUs) from As gene clone libraries in blue. Different background colors highlight three bacterial phyla - *Proteobacteria*, *Firmicutes*, and *Chrysiogenetes*.

Figure 6 - Evolutionary relationships of aioA sequences. A total of 72 nucleotide sequences and 543 sites were analyzed. The phylogeny was reconstructed by maximum likelihood and GTR+I+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green, and operational taxonomic unities (OTUs) from As gene clone libraries in blue. Different background colors highlight two *Proteobacteria* classes – *beta*- and *alpha*-proteobacteria.

Figure S1 - Map showing the sampling site. Crosshatch, red and yellow areas represent mining, urban, and sampling areas, respectively.

CAPÍTULO 2

Figure 1 - Taxonomic composition of bacterial (A) and archaeal (B) taxa from MSS microbiome based on the Greengenes database. Other bacteria: *Gemmatimonadetes*, *Cyanobacteria*, *OP3*, *OP11*, *Spirochaetes*, *TM7*, *Chlorobi*, *WS3*, *Elusimicrobia*, *GN04*, *TM6*, *GN02*, *Tenericutes*, *Armatimonadetes*, *BRC1*, *NC10*, *WPS*-2, *Fibrobacteres*, *Fusobacteria*, *H-178*, *FCPU426*, *Kazan*-3B-28, *WS5*, *NKB19*, *Thermi*, *AC1*, *TPD*-58, *WS6*, *Synergistetes*, *OP8*, *WS2*, *ZB3*, *SC4*, *OP1*, *SBR1093*, *SR1*, *Lentisphaerae*, *GAL15*, *PAUC34f*, *LCP*-89 and *MVS*-104.

Figure 2 - Fragment recruitment plots of the MSS contigs by Candidatus *Nitrospira defluvi* and *Nitrosopumilus maritimus* SCM1 genomes.

Figure 3 - SEED subsystems distribution of the MSS metagenome based on MG-RAST annotation.

Figure 4 - Significant SEED subsystem differences as a result of a Fisher exact test between the MSS and RAW metagenomes conducted with the STAMP program. Enrichment of SEED subsystem in the RAW metagenome has a positive difference between proportions (blue circles), while enrichment of SEED subsystem in the MSS metagenome has a negative difference between proportions (orange circles).

Figure 5 - Significant metal resistance genes differences as a result of a Fisher exact test between the MSS and RAW metagenomes conducted with the STAMP program. Enrichment of metal resistance genes in the RAW metagenome has a positive difference between proportions (blue circles), while enrichment of metal resistance genes in the MSS metagenome has a negative difference between proportions (orange circles).

Figure 6 – Average well color development (AWCD) calculated from the consumption of carbon sources of anaerobic and aerobic microbial communities.

Figure S1 - Rarefaction curve of number of OTUs observed with an evolutionary distance of 0.03, 0.05 and 0.10.

Figure S2 – Fragment recruitment plots of the MSS contigs by bacterial and archaeal genomes. A - *Anaeromyxobacter dehalogenans* 2CP-1; B - *Chitinophaga pinensis* DSM2588; C - *Geobacter metallireducens*; D - *Leptothrix chlolodnii*; E - *Sideroxydans lithotrophicus*; F - *Thiobacillus denitrificans*; G - *Thiomonas* 3As; H - Candidatus *Nitrososphaera gargensis*; I - *Cenarchaeum symbiosum*.

Figure S3 - The bacterial standard curve (A and B) and the archaeal qPCR standard curve (C and D).

Figure S4 - The C_t values from the 16S rRNA gene amplifications. A and B represent bacterial and archaeal amplifications, respectively.

Figure S5 – Nitrogen cycle representation obtained in the Keeg Mapper analysis of MG RAST web server based on SEED database. The red square represents the presence of enzyme sequence in the MSS metagenome.

Figure S6 – The Keeg Map obtained using the MG-RAST web server based on SEED database.

LISTA DE TABELAS

CAPÍTULO 1

CAPÍTULO 2

LISTA DE ABREVIATURAS

Ag	Prata
As	Arsênio
As^0	Arsênio elementar
AsIII	Arsenito
As-III	Arsenieto
AsV	Arseniato
AsH_3	Arsina
Au	Ouro
$(CH_3)_3As^+$	Tetrametilarsônio
C	Carbono
Cd	Cádmio
CIM	Concentração Inibitória Minima
cm	Centímetro
Co	Cobalto
CO_2	Dióxido de carbono
CONAMA	Conselho Nacional do Meio Ambiente
Cr	Cromo
Cu	Cobre
DMA	Ácido dimetilarsínico
DNA	Ácido desoxirribonucléico
Eh	Potencial oxidação-redução
Fe	Ferro
FeAsS	Arsenopirita
g	Grama
GlpF	Aquaporina
Hg	Mercúrio
kDa	Kilodalton
L	Litro
mg	Miligrama
MMA	Ácido monometilarsônico
N	Nitrogênio
Ni	Níquel
orfs	Open reading frame
pb	Pares de bases
Pb	Chumbo

PCR	Reação em Cadeia da Polimerase
pH	Potencial hidrogeniônico
qPCR	Quantitative real-time PCR
rRNA	RNA ribossômico
RNA	Ácido ribonucleico
Sb	Antiamônio
t	Tonelada
Zn	Zinco
µg	Micrograma

Sumário

Resumo

O Quadrilátero Ferrífero (Minas Gerais, Brasil) é uma das maiores regiões de mineração do mundo, sendo historicamente explorada por mais de 300 anos. Desde então, muitos metais e metaloides tóxicos foram liberados no ambiente, levando a contaminação de corpos d'água. Considerando que a comunidade procariótica influencia a biodisponibilidade destes elementos tóxicos e sua importância no equilíbrio ecológico de ambientes, investigou-se no microbioma a diversidade taxonômica e funcional no sedimento do Córrego da Mina, historicamente contaminado por metais. Abordagens clássica de cultivo – com ênfase em bactérias resistentes ao arsênio (As) e genes responsáveis pela sua transformação –, e metagenômica, usando plataforma de sequenciamento de nova geração, foram empregadas. Cento e vinte e três isolados bacterianos oxidadores de AsIII e redutores de AsV foram identificados, representados por 20 gêneros. Este estudo descreve pela primeira vez os gêneros *Thermomonas* e *Pannonibacter* como transformadores de As. A caracterização fenotípica dos isolados revelou que 72% foram redutores de AsV e 20% oxidadores de AsIII. Além disso, a caracterização genotípica revelou a presença dos genes *arsC*, *aioA* e *arrA* nos isolados: 85% dos isolados abrigaram o gene *arsC*, 20% o gene *aioA* e 7% o gene *arrA*. A análise filogenética dos genes *arsC* e *aioA* dos isolados e das bibliotecas de clones sugerem que os isolados obtidos representam bactérias ambientalmente importantes na especiação do As. Além disso, o perfil taxonômico revelado pela análise do metagenoma evidenciou uma comunidade complexa, com dominância dos filos Proteobacteria e Parvarcheota. Genomas de bactérias e arqueias foram reconstruídos baseados em anotações do banco de dados SEED, destacando-se os genomas de Candidatus *Nitrospira defluvii* e *Nitrosopumilus maritimus*. A presença dessas espécies no sedimento sugere importante participação no ciclo de C e N. Reconstrução funcional revelou um conjunto diversificado de genes para a assimilação de amônia e amonificação. Estes processos estão envolvidos na manutenção do ciclo de N no sedimento. Além disso, a anotação funcional do metagenoma revelou uma grande diversidade de genes de resistência a metais. Verificou-se, ainda, elevada diversidade metabólica no sedimento do Córrego da Mina, sugerindo que a contaminação histórica por metais, com ênfase em As, não mais afeta a comunidade procariótica. Portanto, os resultados aqui relatados adicionam novos conhecimentos na composição taxonômica e funcional microbiana de sedimentos contaminados por metais.

Abstract

The Iron Quadrangle (Minas Gerais, Brazil) is one of the world's largest mining regions, being historically explored for over 300 years. Since then, many toxic metals and metalloids were released into the environment, leading to contamination of water bodies. Since prokaryotic community influences the bioavailability of these toxic elements and their importance in the ecological balance of various environments, we investigated the taxonomic and functional diversity in the microbiome from Mina Stream, historically metal-contaminated. Classical approache of cultivation - with emphasis on As-resistant bacteria and genes responsible for As transformation -, and metagenomic sequencing using next-generation platform were applied. One hundred and twenty-three bacterial isolates were identified as AsIII oxidizers and AsV reducers, represented by 20 genera. This study describes for the first time *Thermomonas* and *Pannonibacter* as As-transforming genera. Phenotypic characterization of isolates revealed that 72% were AsV reducers and 20% were AsIII oxidizers. In addition, the genotype characterization revealed the presence of *arsC*, *arrA* and *aioA* genes in the bacterial isolates: 85% of the isolates harbored *arsC* gene, 20% *aioA* gene and 7% *arrA* gene. Phylogenetic analysis of *arsC* and *aioA* genes obtained from the isolates and clones suggest that these isolates represent environmentally important bacteria acting on the As transformation. Moreover, taxonomic profile obtained by metagenome analysis revealed a complex community, with dominance of Proteobacteria and Parvarcheota. Bacterial and archaeal genomes were reconstructed based on SEED subsystems database, especially Candidatus *Nitrospira defluvii* and *Nitrosopumilus maritimus*. Their presence implicated them in C and N cycling in the MSS. Functional reconstruction revealed a large diverse set of genes for ammonium assimilation and ammonification. These processes have been implicated in the maintenance of N cycle of the sediment. Functional annotation unveiled a high diversity of metal resistance genes. Furthermore, a high metabolic diversity was detected in the sediment of Mina Stream, suggesting that the historical metal contamination is no longer affecting it. Finally, the results reported herein may contribute to expand the current knowledge of the microbial taxonomic and functional composition of metal-contaminated sediments.

Estrutura da Tese

Esta tese apresenta a seguinte organização: justificativa do trabalho, seguida de uma introdução geral com uma revisão bibliográfica abordando os principais temas propostos, os objetivos gerais e específicos. Em seguida, os artigos aceito e submetido serão apresentados na forma de capítulos: capítulo 1 - **Bacteria and genes involved in arsenic speciation in sediment impacted by long-term gold mining;** capítulo 2 - **Metagenome of a microbial community inhabiting a metal-rich tropical stream sediment**. Finalmente, uma discussão e conclusão geral será apresentada, seguida de referências bibliográficas e anexos.

1 - Justificativa

O Brasil se destacou desde o início da sua colonização pela abundância de seus recursos naturais. Inicialmente, sofreu com a extração da madeira do pau-brasil e, logo em seguida, com a exploração da riqueza mineral presente no território brasileiro. A atividade mineradora de ouro e de diamante foi de alta relevância nos aspectos sócio-econômicoe políticodo Brasil-Colônia, durante quase todo o século XVIII, levando a um grande avanço territorial com as expedições para o interior do país. Internacionalmente conhecida como uma atividade que desenvolve a economia do país, a mineração tem forte impacto na economia de grandes nações, como os Estados Unidos, Austrália e Canadá, dentre outros. No Brasil, a exploração mineral é até os dias de hoje de extrema importância sendo um dos setores básicos da economia e contribui para o desenvolvimento da sociedade, já que os minérios explorados estão em todos os produtos utilizados pelo homem (Mello et al., 2006).

O crescente desenvolvimento tecnológico e industrial no Brasil aumenta a demanda pelos metais, cuja extração só é possível através da escavação de rochas e de solos, gerando, em consequência, a liberação de rejeitos industriais metálicos de Fe, Pb, As, Cr, Zn, Cd, Cu, Hg, dentre outros. Devido a sua facilidade de solubilização e posterior mobilização, podem causar danos significativos ao ambiente e à saúde humana (Moraes & Jordão, 2002; Ávila & Monte-Mór, 2007; Rabadjieva et al., 2008). A exploração dos recursos minerais de forma desordenada sempre esteve associada à destruição da natureza, e são frequentes as imagens dos impactos provocados principalmente por minas a céu aberto (Pompêo et al., 2004).

No Brasil, a região do Quadrilátero Ferrífero (MG) apresenta uma contaminação histórica das águas e sedimentos de rios e córregos por As devido aos rejeitos gerados ao longo de 300 anos de atividade mineradora nessa área (Borba et al., 2000; Deschamps et al., 2002; Borba et al., 2004; Bundschuh et al., 2012). A Bacia do Rio das Velhas, principalmente a região de Nova Lima, localizada no Quadrilátero Ferrífero, apresenta altas concentrações de As em seus rios e córregos, tornando necessário estudos sobre a comunidade de microrganismos nesse ambiente.

A análise molecular da diversidade taxonômica e funcional da comunidade de procariotos em ambientes contaminados por metais amplia o conhecimento dos seus ciclos biogeoquímicos, como, por exemplo, a mobilidade ou imobilidade de determinados metais, como o As. Além disso, pode revelar interações que mantêm a comunidade procariótica nesses ambientes impactados, podendo levar à identificação de novas espécies e de genes de interesse biotecnológicos ainda não descritos. Embora existam estudos sobre a composição de comunidades de bactérias resistentes a As em ambientes impactados por mineração de ouro, pouco se sabe sobre essas bactérias em sedimento historicamente impactado por As (Anderson & Cook, 2004; Chang et al., 2008; Drewniak et al., 2008; Cai et al., 2009; Oliveira et al., 2009; Cavalca et al., 2010).

Sabe-se que os sedimentos exercem um papel importante no transporte e armazenamento de contaminantes refletindo na ecologia e biodiversidade do sistema aquático. Na maioria desses ambientes, as concentrações de metais dissolvidos na água são baixas, pela precipitação de sólidos ou adsorção de partículas suspensas depositadas no sedimento. Além disso, as bactérias exercem um papel essencial no ciclo biogeoquímico do As, influenciando sua especiação e biodisponibilidade. Portanto, é relevante conhecer a diversidade e as funções metabólicas dessa comunidade, podendo revelar novos genes e vias metabólicas de interesse biotecnológico, incluindo possíveis indicadores.

2 – Introdução

2.1 - Arsênio

O arsênio (As) é o 20º elemento mais abundante da crosta terrestre. Na natureza encontra-se em minérios associado principalmente à compostos sulfurosos de Fe, e outros como Au, Ag, Cu, Sb, Ni e Co, sendo a arsenopirita (FeAsS) o minério mais abundante encontrado na crosta. O intemperismo das rochas contendo As e a extração de metais pelas mineradoras, que só é possível através da quebra de rochas e escavação de solos, pode levar à liberação desse metalóide no solo e corpos de água adjacentes (Lièvremont et al., 2009; Tsai et al., 2009). Apesar da principal fonte de contaminação do As ser natural, as atividades antrópicas, principalmente a mineração e o refino de metais, contribuem e aceleram a liberação desse metalóide no ambiente. Outras atividades humanas também podem resultar na liberação de As no ambiente, como o seu uso na produção farmacêutica, no processamento de madeira, na fabricação de vidro e em fungicidas utilizados na agricultura (Lièvremont et al., 2009). O uso do As na indústria e na agricultura vem sendo descontinuado, mas resíduos da contaminação causada por essas atividades ainda podem ser encontrados no ambiente (Páez-Espino et al., 2009).

O As é um tóxico bastante conhecido pelo homem, ocupando o primeiro lugar na lista de contaminantes de água potável (Slyemi & Bonnefoy, 2012). Poluição da água é um dos principais desafios para a saúde pública, principalmente devido a seu potencial carcinogênico em baixas doses (Bhattacharya et al., 2007; McClintock et al., 2012; Kruger et al., 2013). Sua exposição prolongada pode ter efeito teratogênico e clastogênico (Chen et al., 2005).

O As é encontrado no ambiente em diferentes estados oxidativos: arseniato (AsV), arsenito (AsIII), arsênio elementar (As^0) e arsenieto (As-III), sendo considerado um elemento bastante móvel no ambiente devido a essa grande variação em seus estados oxidativos, permitindo sua presença em diferentes condições ambientais (Tsai et al., 2009). O ciclo biogeoquímico do As está ilustrado na Figura 1.

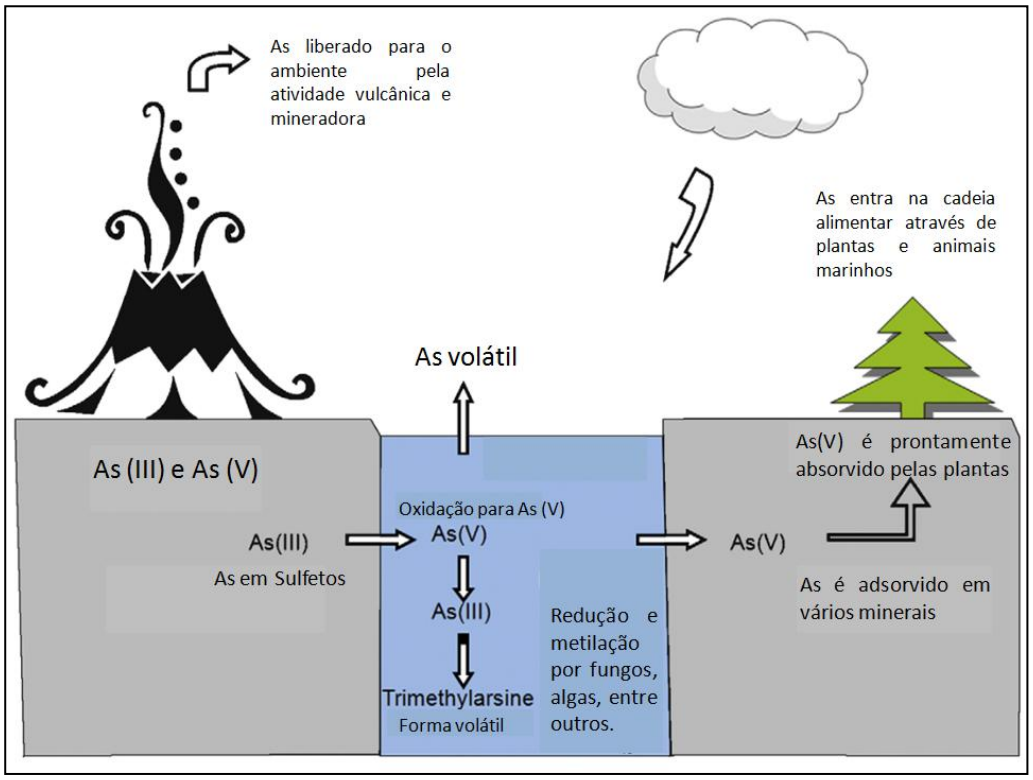

Figura 1- Ciclo biogeoquímico do arsênio. Fonte: Tsai et al., 2009, modificado.

O ciclo biogeoquímico deste elemento envolve diversos processos físico-químicos e biológicos nos quais os procariotos desenpenham papel fundamental (Lièvremont et al, 2009). As formas inorgânicas AsV e AsIII são as principais formas que ocorrem no ambiente aquático e a proporção relativa desses dois estados de oxidação depende de vários fatores, como processos biológicos e condições físico-químicas (Escalante et al., 2009; Liao et al., 2011). As^0 e arsina (AsH_3) são encontrados em ambientes extremamente redutores. O AsV é a forma inorgânica de As em ambientes aeróbicos com alto Eh (potencial oxidação-redução) e pH baixo, enquanto que, inversamente, AsIII se torna dominante em ambientes anaeróbicos, com baixo Eh e pH alto. Além de ser menos tóxico que AsIII, AsV tem a tendência de ser fortemente adsorvido ligando-se com minerais inorgânicos, especialmente minério de ferro e óxidos de alumínio, sendo imobilizado no sedimento, enquanto que, AsIII é considerado mais solúvel mantendo-se biodisponível na água (Oremland & Stoltz, 2005). Portanto, dependendo das características abióticas do ambiente, como o potencial redox e o pH, estados oxidativos distintos do As podem predominar.

Além das condições físico-químicas, diversos estudos fazem a ligação da especiação do As, ou seja, a transformação de uma forma inorgânica em outra, com as atividades microbianas no ambiente (Paez-Espino et al., 2009; Lièvremont et al., 2009). A proporção dos vários estados oxidativos do As em um determinado ambiente depende de reações envolvidas no metabolismo de microrganismos, especialmente de bactérias, como a redução do AsV, a oxidação do AsIII e várias reações de metilação (Figura 1). Essas reações protegem os microrganismos dos efeitos tóxicos desse metalóide e podem contribuir para o aumento ou redução da contaminação, dependendo da estrutura da comunidade bacteriana presente no ambiente (Lièvremont et al., 2009). Além disso, diversos microrganismos têm mostrado capacidade de metilar As originando monometil, dimetil e trimetil (Orelamd & Stolz, 2003). Essas arsinas metiladas são voláteis e rapidamente liberadas para a atmosfera (Figura 1). Portanto, o estudo da composição da comunidade bacteriana nesses ambientes é de extrema relevância para entender a especiação do As no ambiente. Além disso, pode-se entender sua participação no ciclo biogeoquímico do As, possibilitando o uso de bactérias indígenas ou geneticamente modificadas na biorremediação do ambiente contaminado por arsênio.

2.2 - Arsênio um poluente global

As altas concentrações de As nos corpos d'água e solo em diversas regiões do mundo são decorrentes, principalmente, de atividades antrópicas, como mineração e indústria metalúrgica, e intemperismo de rochas ricas em minérios contendo As (Figura 2) (Smedley & Kinniburgh, 2002). Vários países têm enfrentado problemas ambientais causados pela contaminação por As, sendo detectados níveis acima do limite permitido pela Organização Mundial de Saúde (10 μg/L para a água potável) (WHO, 2001).

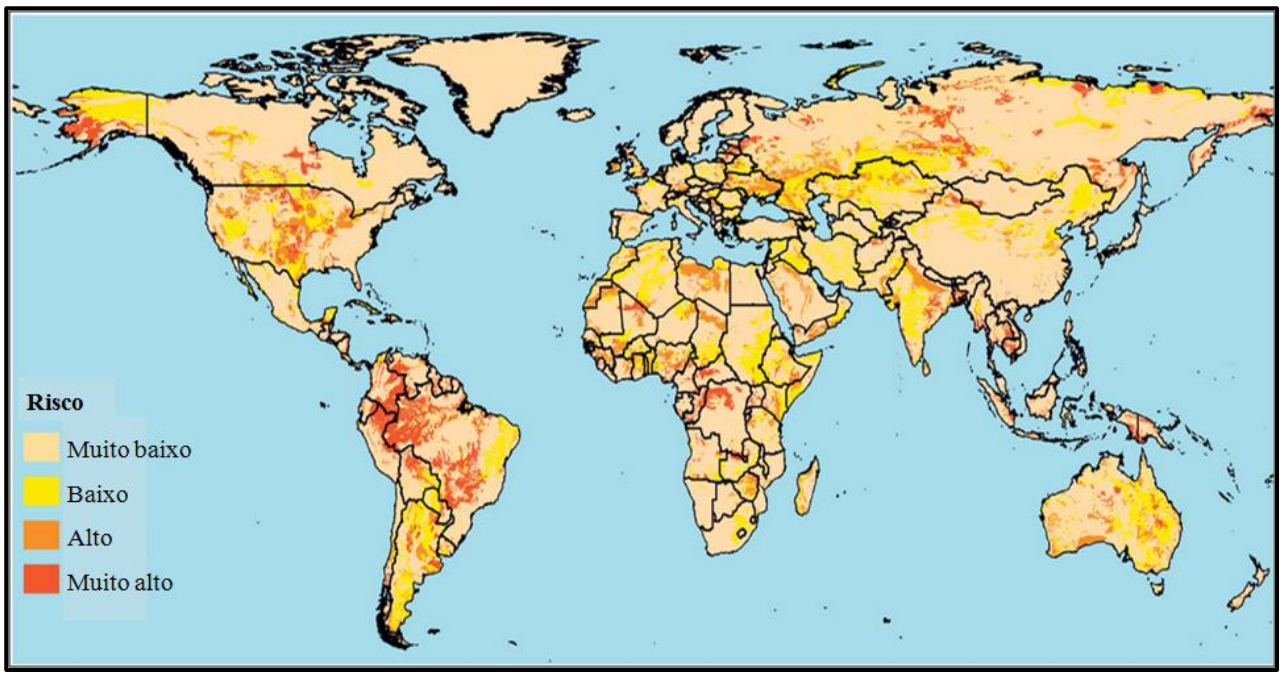

Figura 2 - Mapa das regiões contaminadas com arsênio e risco da exposição crônica nessas áreas. Fonte: Schwarzenbach et al., 2010, modificado.

Diversos países tiveram problemas de saúde pública devido à liberação de As causada pelo intemperismo de rochas contendo esse metalóide. O caso mais grave de contaminação por intemperismo, e que foi responsável pelo maior problema de saúde pública do mundo, devido a altas concentrações de As em águas utilizadas para consumo, ocorreu devido à contaminação de água subterrânea em Bangladesh, e West Bengal (Índia), onde pelo menos 100 milhões de pessoas ficaram expostas ao risco de câncer e outras doenças relacionadas à ingestão de altas concentrações de As (Orelamd & Stolz, 2005). Além disso, o intemperismo foi responsável pela contaminação de águas no Nepal, Paquistão, Vietnã, Camboja e Austrália, causando dificuldades ao abastecimento dos recursos hídricos nessas regiões (Bhattacharya et al., 2007).

A contaminação de recursos hídricos por As causada por atividades antrópicas também já foi relatada em diversos locais ao redor do mundo, como Guam – uma ilha do Oceano Pacifico –, Estados Unidos e Canadá. Nas Américas do Sul e Central ocorre liberação do As por ambas as formas, intemperismo e atividades antrópicas, abrangendo países como México, Argentina,

Bolívia, Nicarágua, Brasil, dentre outros (Figura 3) (Matschullat, 2000; Nordstrom, 2002; Smedley et al., 2002; 2005; Barragner-Bigot, 2004; Bhattacharya et al., 2007).

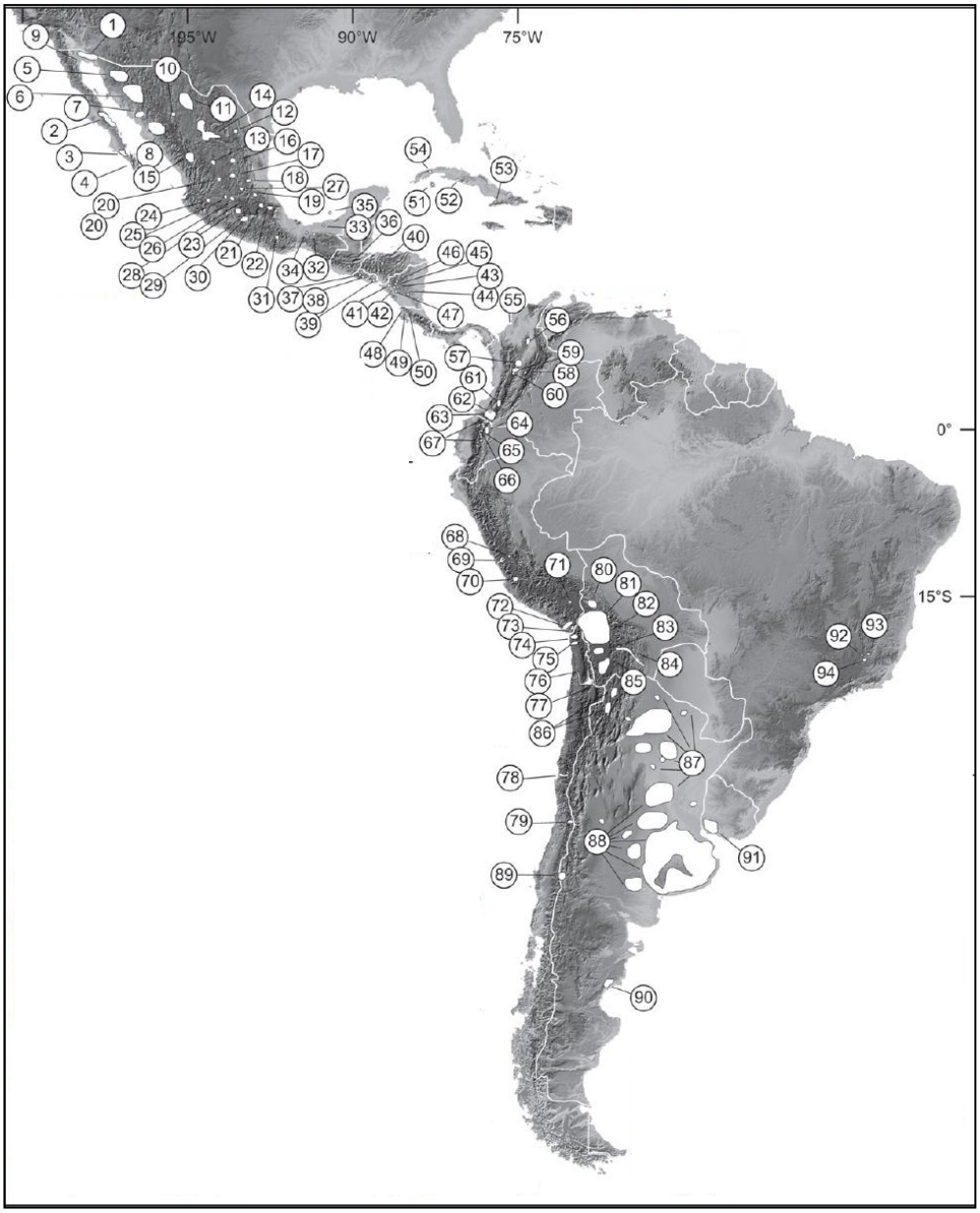

Figura 3 – Mapa mostrando as regiões das Américas Central e do Sul onde foram detectadas altas concentrações de arsênio. As regiões destacadas no Brasil são: Nova Lima (92), Santa Bárbara (93) e Ouro Preto/Mariana (94). Fonte: Bundschuh et al., 2012, modificado.

2.3 - Quadrilátero Ferrífero

O Quadrilátero Ferrífero, situado no Estado de Minas Gerais, é considerado uma das regiões mais ricas em minérios do mundo, principalmente, minério de ferro e ouro, abrigando diversas minas de ouro ativas ou atualmente abandonadas (Daus et al., 2005; Bundschuh et al., 2012). Essa região vem sendo explorada desde o final do século XVII, tendo como o auge da exploração mineral o século XVIII e, portanto, sofre com o despejo histórico de rejeitos e com a contaminação por metais pesados associados a eles em córregos, rios, lagoas, águas subterrâneas e sedimentos na região há mais de 300 anos, levando a impactos ambientais significativos (Matschullat et al., 2000; Costa et al., 2003).

A presença de altos níveis de As nessa região deve-se, principalmente, a fontes antrópicas relacionadas a atividades de mineração e refino de minério de depósitos auríferos sulfetados, como a arsenopirita (FeAsS) (Borba et al., 2003). O Quadrilátero Ferrífero contribuiu com a produção de 1.300 t de ouro nos últimos três séculos e, considerando a razão As/Au nos minérios, que pode variar de 300 a 3.000, estima-se que entre 390.000 t a 3.900.000 t de As podem ter sido liberadas para o ambiente (Borba et al., 2000; Deschamps et al., 2002). O Quadrilátero Ferrífero é reconhecido como uma área de alta contaminação por As, possuindo risco potencial de exposição crônica (Garelick & Jones, 2008; Schwarzenbach et al., 2010; Bundschuh et al., 2012) (Figuras 2 e 3).

Dentre as principais e mais antigas regiões de mineração de ouro do Quadrilátero Ferrífero estão as regiões de Nova Lima, Santa Bárbara e Ouro Preto/Mariana (Figura 3). As concentrações de As em Nova Lima são extremamente altas devido ao despejo de rejeitos de mineração que eram depositados ao longo do Córrego Cardoso, um afluente do Rio das Velhas (Bundschuh et al., 2012). Estudos realizados pelo Instituto Mineiro de Gestão das Águas (IGAM) na Bacia do Rio das Velhas revelaram concentrações elevadas de metais tóxicos acima dos limites determinados pela lei brasileira (IGAM, 2012). As maiores concentrações de As detectadas nessa bacia são encontradas em seu alto curso, sendo muito acima do permitido pelo Conselho Nacional

do Meio Ambiente (CONAMA) – órgão que determina as leis ambientais (0,5 mg/L de As – Resolução N^{o} 430, 2011). Além das fontes antrópicas, fontes naturais de As no Quadrilátero Ferrífero podem levar à contaminação e estão relacionadas às rochas que contêm depósitos auríferos sulfetados, formados por arsenopirita e pirita. Essas rochas sofrem intemperismo, seguido de lixiviação e drenagem do As, resultando na contaminação de águas e sedimentos (Borba et al, 2004).

A contaminação das águas presentes no Quadrilátero Ferrífero é extremamente prejudicial à saúde já que o abastecimento público de água é feito através das bacias hidrográficas situadas nessa região (Borba et al, 2004). Estudo realizado para determinar a concentração de As na urina de mais de 300 pessoas, principalmente crianças de 8 a 14 anos, residentes em três cidades pertencentes ao Quadrilátero Ferrífero (Nova Lima, Santa Bárbara e Ouro Preto), revelou altas concentrações desse metalóide na urina amostrada, variando de 2 a 106 μg/L. Segundo Matschullat et al. (2007), concentrações de As na urina variando de 15 a 40 μg/L são consideradas críticas para a saúde. Destaca-se, ainda, que as amostras coletadas na região de Nova Lima apresentaram as maiores concentrações de As, recomendando a necessidade o monitoramento da contaminação dos recursos hídricos dessa região.

A contaminação por metais pesados de águas utilizadas para o consumo por diversas cidades torna essencial a pesquisa do efeito dessa contaminação sobre comunidades de seres vivos. Considerando que a comunidade microbiana é essencial em diversos processos ecológicos, como a cadeia alimentar aquática, qualquer alteração ambiental deve ser primeiramente observada nesta comunidade. Os microrganismos têm papel essencial na dissolução e transformação de diversos metais em sua forma biologicamente ativa. Portanto, torna-se relevante o conhecimento da estrutura e diversidade desses microrganismos em ambientes contaminados por metais pesados e metaloides.

2.4 - Resistência bacteriana ao arsênio

Os microrganismos desenvolveram mecanismos dinâmicos de resposta ao As na natureza. Dependendo da espécie de microrganismo presente no ambiente estas respostas podem ser: quelação, compartimentalização, exclusão ou transfomação bioquímica (Paez-Espino et al., 2009; Tsai et al., 2009; Slyemi & Bonnefoy, 2012). A maioria dos organismos vivos desenvolveu mecanismos de resistência ao As, mas apenas alguns são capazes de utilizá-lo para obter energia. Além disso, alguns microrganismos são capazes de metilar as formas inorgânicas ou desmetilar os compostos orgânicos contendo As (Páez-Espino et al., 2009).

Apesar de diversos microrganismos utilizarem as formas inorgânicas de As como aceptor ou doador de elétrons, as células bacterianas não possuem transportadores específicos para a captação de As devido à sua extrema toxicidade, levando suas formas inorgânicas a entrarem na célula utilizando outros transportadores (Tsai et al., 2009; Slyemi & Bonnefoy, 2012). Devido à similaridade química do AsV com o fosfato, o primeiro entra na célula através dos transportadores de fosfato, como os transportadores Pst e Pit que permitem a entrada de AsV em *E. coli*. Já a forma inorgânica AsIII entra na célula através de aquaporinas (GlpF) em *E. coli* (Figura 4) (Páez-Espino et al., 2009; Kruger et al., 2013).

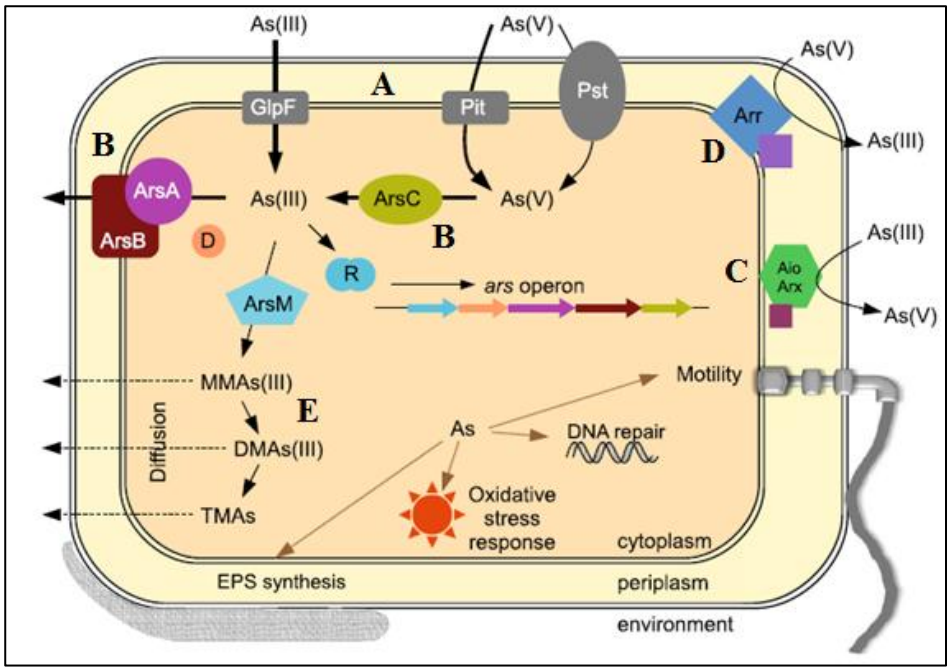

Figura 4 - Mecanismos de entrada e de resistência ao arsênio na célula bacteriana. (A) AsV entra na célula através de transportadores de fosfato e AsIII entra na célula através de aquaporinas. (B) Dentro da célula, AsV é reduzido a AsIII pela enzima ArsC. O AsIII produzido é eliminado da célula pela bomba de efluxo ArsB. Além disso, (C) AsIII atua como doador de elétrons no processo de oxidação a AsV. (D) AsV também pode atuar como aceptor final de elétrons durante a respiração celular. (E) Arsênio inorgânico também pode ser transformado em espécies orgânicas através do processo de metilação. Fonte: Krueger, 2013.

A metilação é um mecanismo de resistência considerado um processo de detoxificação das células por meio da metilação das formas inorgânicas AsV e AsIII. Em eucariotos superiores, a enzima glutationa reduz AsV a AsIII, sendo o último metilado e transformado nas formas orgânicas, ácido dimetilarsínico (DMA - $(CH_3)_2AsO(OH)$), ácido monometilarsônico (MMA - $CH_3AsO(OH)_2$) e tetrametilarsônio ($(CH_3)_3As^+$). A reação de metilação é diferente nas células bacterianas, onde AsIII é primeiramente oxidado a AsV e metilado em seguida (Stolz et al., 2002; Tsai et al, 2009). As formas orgânicas metiladas de As são voláteis e liberadas rapidamente para o ambiente. Apesar da metilação ser considerada apenas um mecanismo de detoxificação, estudos já

demonstraram que nem sempre as formas metiladas são menos tóxicas, como, por exemplo, o óxido metilarsina, o qual é mais tóxico do que AsIII para a levedura *Cryptococcus humiculus* (Bentley & Chasteen, 2002; Slyemi & Bonnefoy, 2012). Além do processo de metilação, alguns microrganismos são conhecidos por realizarem o processo contrário, ou seja, desmetilar compostos orgânicos metilados de arsênio transformando em AsV que será usado como fonte de energia (Slyemi & Bonnefoy, 2012).

O mecanismo de resistência bacteriana baseado na redução de AsV e na oxidação de AsIII pode ser classificado em duas categorias: transformações do As que consistem em reações de detoxificação da célula bacteriana e transformações que geram energia para o crescimento celular (Figura 4) (Stolz et al., 2002; Silver & Phung, 2005).

Atualmente, duas formas de redução de AsV já foram descritas, sendo a primeira um mecanismo de detoxificação e a segunda um mecanismo de geração de energia. O operon *ars* (Figura 4), responsável pelo sistema de detoxificação, é o mais frequente entre as bactérias na natureza. Ele é constituído por três genes (*arsR*, *arsB* e *arsC*) podendo conter, em algumas bactérias, até cinco genes (*arsR*, *arsD*, *arsA*, *arsB* e *arsC*) (Lièvremont et al., 2009). Os genes *arsA* e *arsB* codificam uma bomba de efluxo transmembrana que exporta AsIII da célula, reduzindo a sua concentração intracelular. O gene *arsC* codifica a enzima arseniato redutase, que é responsável pela biotransformação de AsV para AsIII. Os genes *arsR* e *arsD* codificam um repressor da transcrição e uma chaperona que tem como função transferir AsIII para a bomba de efluxo, respectivamente (Silver & Phung, 2005; Kaur et al., 2009; Páez-Espino et al., 2009; Liao et al, 2011). As enzimas arseniato redutase são pequenas e pertencem a duas famílias distintas que evoluíram convergentemente: thioredoxin e glutaredoxin, constituídas por enzimas codificadas por homólogos do gene *arsC* encontrado nos plasmideos pI258 de *Staphylococcus aureus* e R773 de *Escherichia coli*, respectivamente (Krueger et al., 2013). Essas duas famílias diferem em suas estruturas e nos mecanismos de redução, mas ambas são codificadas pelo gene *arsC* (Slyemi &

Bonnefoy, 2012). O operon *ars* foi descoberto primeiramente em plasmídeos, mas pode ser encontrado no cromossomo de diversos organismos (Stolz, 2010).

Além do mecanismo de detoxificação, outro processo de redução de AsV, conhecido como redução dissimilatória, já foi descrito para diversos filos bacterianos, incluindo bactérias anaeróbias obrigatórias e facultativas e algumas arqueias (Lièvremont et al, 2009). *Geospirillum arsenophilus*, isolada de sedimentos anóxicos, foi a primeira bactéria descrita na qual o processo de respiração dissimilatória acontece (Ahmann et al., 1994). A enzima arseniato redutase dissimilatória ArrAB é composta de duas subunidades: maior, ArrA (100 kDa), e menor, ArrB (30 kDa) (Slyemi & Bonnefoy, 2012). A arseniato redutase é uma enzima periplasmática que funciona durante a respiração anaeróbia, utilizando AsV como aceptor final de elétrons (Malarsan et al, 2004). As duas subunidades são codifiadas pelos genes *arrA* e *arrB* que estão agrupados em um operon. Variação no número de genes dentro do operon *arr* já foi detectada em algumas bactérias, especialmente na análise do genoma de *Desulfitobacterium hafniense*, onde foram encontradas cinco *orfs* codificando: ArrC, uma proteína de membrana que serve como âncora para a ArrAB; uma chaperona ArrD, que parece estar envolvida na inserção de cofator em ArrA; ArrR, ArrS, e ArrT consideradas três proteínas reguladoras (Reyes et al., 2008).

A oxidação microbiana do AsIII foi descrita pela primeira vez em 1918 e é mediada por duas arsenito oxidase distintas: AioBA e ArxAB. Os genes responsáveis pela oxidação de arsenito já foram descritos em diversas espécies bacterianas heterotróficas e quimioautotróficas (Hamamura et al, 2009). A oxidação de AsIII é catalizada pela enzima AioBA codificada pelos genes *aioB* e *aioA*, anteriormente chamados de *aoxA/B*, *aroB/A* e *asoB/A* (Lett et al., 2012). Essas bactérias utilizam, através de respiração aeróbia e anaeróbia, AsIII como doador de elétrons para a redução de oxigênio e nitrato (Lièvremont et al, 2009). A energia produzida é usada para fixar CO_2, fornecendo carbono necessário para o crescimento. Em populações onde bactérias heterotróficas e quimioautotróficas estão presentes, as primeiras conseguem sobreviver utilizando

substâncias orgânicas produzidas pelas bactérias quimioautotróficas (Hamamura et al, 2009; Lièvremont et al, 2009).

Recentemente, uma nova arsenito oxidase, ArxAB, foi descrita em *Alkalilimnicola ehrlichii* e *Ectothiorhodospira* sp. (Zargar et al., 2010; Zargar et al., 2012). O operon *arx*, contendo os genes *arxA* e *arxB*, só é expresso na presença de condições anaeróbicas e quando AsIII está presente no ambiente (Zargar et al., 2010). Estudos filogenéticos detectaram uma relação próxima entre as enzimas ArxAB, ArrAB e AioBA, sendo as duas primeiras mais próximas uma da outra (Slyemi & Bonnefoy, 2012; Kruger et al., 2013). Essa proximidade filogenética fez com que Oremland e colaboradores (2009) sugerissem que a enzima ArxAB seria a forma ancestral de ArrAB e AioBA.

2.5 - Diversidade de procariotos

A descoberta, o cultivo e a pesquisa de novos microrganismos vêm sendo estudada pelos microbiologistas desde os primórdios da bacteriologia. Robert Koch (1880) foi o responsável pelo surgimento da bacteriologia, a partir do desenvolvimento de técnicas laboratoriais. O cultivo de microrganismos foi de extrema importância para a consolidação da Microbiologia e as informações contidas atualmente em livros-texto vieram do estudo de organismos em cultura pura (Handelsman, 2004). O fascínio em estudar as interações desses seres microscópicos e unicelulares com o seu habitat se deve ao fato de constituírem a maior parte, e ainda grandemente desconhecida, da biota terrestre. A diversidade procariótica é o resultado de cerca de 3,8 bilhões de anos de evolução, sendo esses seres extremamente versáteis metabolicamente, fundamentais nos ciclos biogeoquímicos, processos de biorremediação, síntese de produtos naturais, biocatálise e conversão de energia, possibilitando a ocupação de todos os habitats da Terra (Rondon et al., 2000; Torsvik et al., 2002).

Entretanto, com o surgimento das técnicas moleculares e de métodos independentes de cultivo revelou-se que o cultivo não detectava a real diversidade dos microrganismos. Além disso, a ênfase nos estudos com cultura pura levava à separação dos microrganismos de suas comunidades e direcionava o foco para seu comportamento em ambientes biologicamente simples, como a placa de Petri e os tubos de ensaio. Assim, por um longo período, a análise integrada das comunidades microbianas complexas foi prejudicada (Shapiro & Dworkin, 1997). Apesar do cultivo não ser a técnica adequada em estudos de diversidade, a obtenção de culturas de microrganismos ainda é muito importante e permanece como um grande desafio para o estudo das comunidades microbianas. Deve-se salientar que o cultivo é necessário para que se possa conhecer a atividade metabólica e fisiológica dos microrganismos e sua participação em importantes processos ambientais, além de ser necessária para a identificação de novas espécies (Garland, 1996; Foley et al, 2008).

Até a década de 1980 a classificação taxonômica baseava-se apenas em comparações fenotípicas, incluindo características morfológicas, fisiológicas, metabólicas e químicas das células. Para superar a dificuldade e as limitações encontradas nas técnicas de cultivo, técnicas moleculares, como a reação em cadeia pela polimerase (PCR) (Mullis, 1987) e o sequenciamento de DNA (Sanger, 1977), tiveram forte impacto, sendo extremamente valiosas em diversas áreas da ciência, em particular na classificação microbiana. A introdução do sequenciamento do gene de RNA ribossômico (rRNA) 16S e seu uso como marcador molecular trouxe uma dimensão filogenética à taxonomia bacteriana e tem fornecido informações consideráveis sobre a taxonomia, a ecologia e a evolução de espécies de procariotos encontradas nas amostras ambientais, sem a necessidade de isolamento e cultivo (Woose, 1987; Rappé & Giovannoni, 2003; Tringe & Hugenholtz, 2008; Nascimento, 2011).

Os RNAs ribossômicos (rRNA) são moléculas antigas e conservadas, sendo elementos importantes do ribossomo e responsáveis pela síntese protéica de todos os organismos, e de amplo interesse evolutivo tanto em procariotos como em eucariotos. Nos procariotos essa organela é

formada por duas subunidades ribossômicas: a subunidade 30S, composta por proteínas e o rRNA 16S, e a subunidade 50S, composta de proteínas e dos rRNAs 5S e 23S (Woese, 1987). Para garantir a produção dos três tipos de RNA em quantidades iguais, os genes que dão origem aos rRNAs de ambas as subunidades, 5S (120 pb), 16S (1542 pb) e 23S (2904 pb), estão agrupados em um único operon (Figura 5). Apesar da maioria das espécies de procariotos possuir no máximo duas cópias desse operon, algumas espécies podem conter múltiplas cópias em seu genoma, variando de três a oito e, em poucas espécies, chegando a 15 cópias (Tourova, 2003).

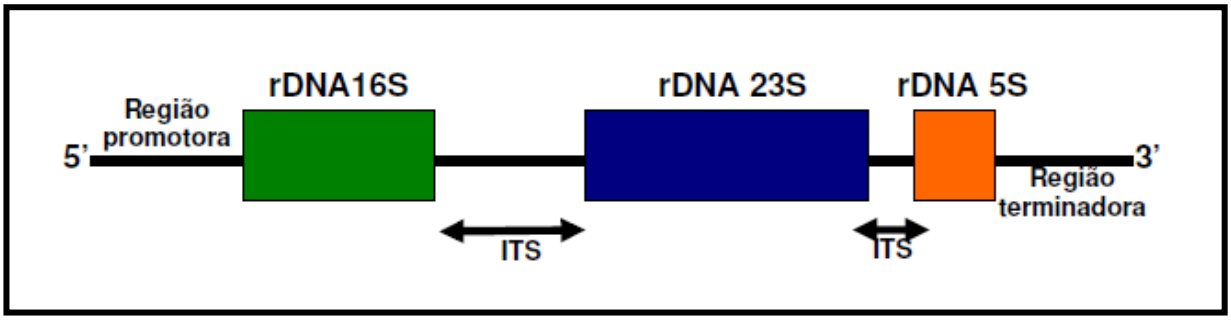

Figura 5 – Esquema do operon ribossômico dos procariotos.

Dentre os três genes ribossômicos, os genes dos RNAs 16S e 23S possuem em sua estrutura primária uma combinação de regiões conservadas, variáveis e altamente variáveis, permitindo a diferenciação de procariotos, com resolução suficiente, tanto em nível de domínio e filo quanto em nível de gênero e até mesmo espécies (Yarza et al., 2014). O gene de rRNA 16S possui 9 regiões, V1-V9, que diferem em tamanho e em relação a sua utilidade para a taxonomia microbiana (Figura 6). A escolha dos iniciadores para a análise de microbiomas é a parte mais crítica do processo já que o uso de iniciadores inadequados pode levar a sub-representação de grupos bacterianos, gerando conclusões biológicas questionáveis (Klindworth et al., 2013). Recentemente, pesquisadores avaliaram estas regiões para definir qual delas melhor definia a diversidade bacteriana e chegaram à conclusão que as regiões V3 e V4, analisadas juntas, cobrem um amplo espectro de filos bacterianos, sendo as regiões ideais para estudo de microbioma

(Klindworth et al., 2013). Estas características tornam essas moléculas ótimos cronômetros evolutivos para estudos filogenéticos dos procariotos. Ao longo das últimas décadas, o gene de rRNA 16S vem tendo destaque nos estudos de diversidade, evolução e ecologia dos procariotos, sendo o marcador taxonômico mais sequenciado, tornando-se o atual pilar da classificação de bactérias e arqueias (Madigan et al., 2004; Pontes et al., 2007; Yarza et al., 2014).

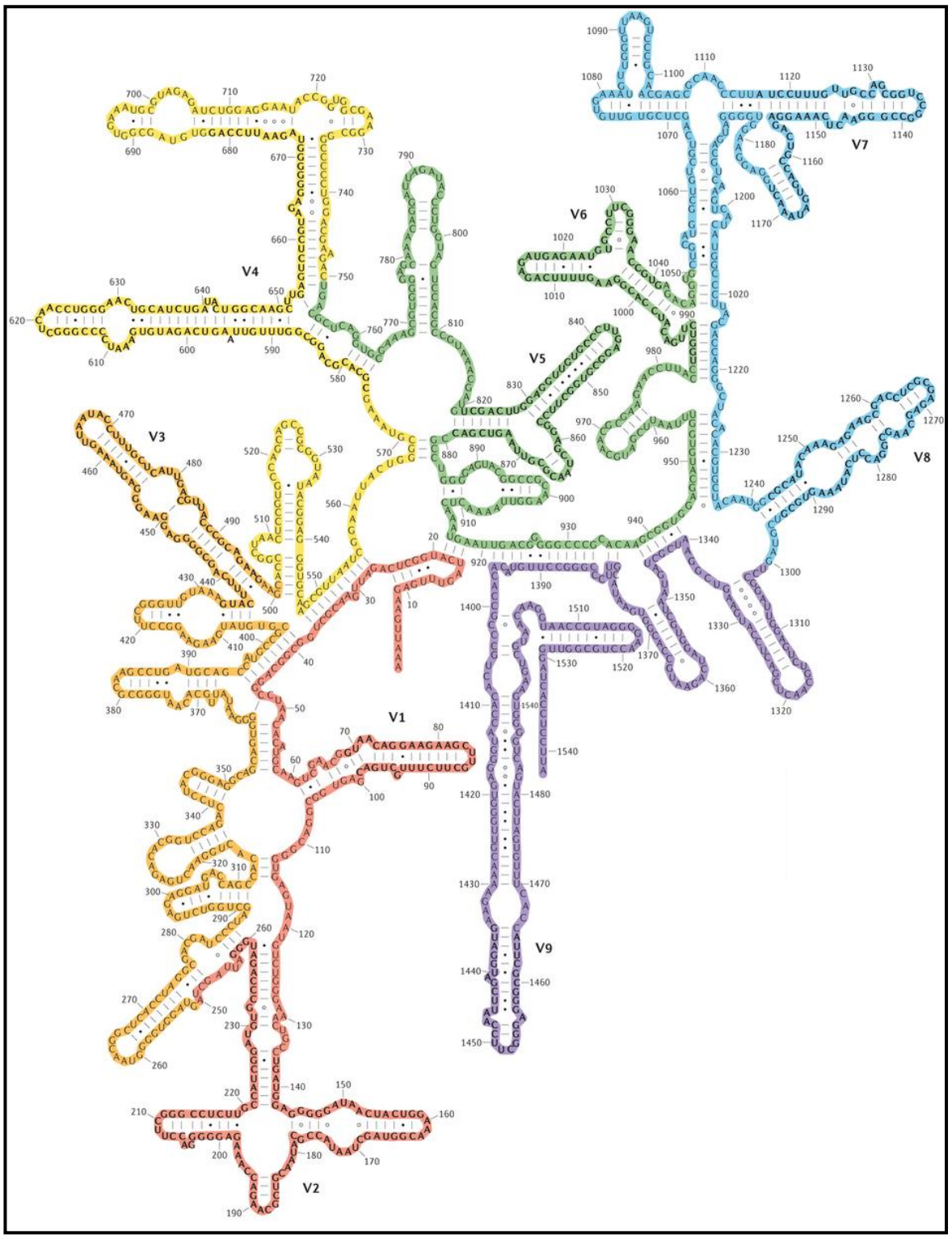

Figura 6 – Estrutura secundária do gene de rRNA 16S de *Escherichia coli*. As regiões variáveis estão destacadas por cores: rosa – V1 e V2; laranja – V3; amarelo – V4; verde – V5 e V6; azul – V7 e V8 e roxo – V9. Fonte: Yarza et al., 2014.

De acordo com a análise dos genes de rRNA, a vida se divide em três domínios: Bacteria, Archaea e Eukarya, sendo que os dois primeiros, os procariotos, são constituídos exclusivamente

por microrganismos. Com o surgimento da PCR e do sequenciamento do gene rRNA 16S foi possível para Woese e colaboradores, em 1987, distinguir 12 grandes filos do domínio Bacteria usando metodologia baseada no cultivo bacteriano. Atualmente, os filos descritos por Woese são denominados Proteobacteria, Actinobacteria, Firmicutes, Chlamydiae, Planctomycetes, Cyanobacteria, Bacteroidetes, Chlorobi, Spirochaetes, Chloroflexi, Deinococcus-Thermus e Thermotogae. Desde 1987, outros 14 filos foram descobertos e caracterizados através de técnicas dependentes de cultivo e, dentre eles, estão filos bem conhecidos como Fusobacteria, Acidobacteria, Nitrospira, Aquificae (Rappé & Giovannoni, 2003). Em 1996, a partir de sequências de rRNA 16S obtidas de amostras ambientais, técnicas independentes de cultivo permitiram a descrição de 26 novos filos, sem representantes cultivados (Hugenholtz *et al*, 1998; Handelsman, 2004). Atualmente, existem nos bancos de dados cerca de 84 filos bacterianos que foram descobertos utilizando técnicas dependentes e independentes de cultivo (GreenGenes, última atualização em outubro de 2014).

O domínio Archaea é constituído por cinco filos, Crenarchaeota, Euryarchaeota, Korarchaeota, Thaumarchaeota e Nanoarchaeota, os quais foram obtidos por meio do sequenciamento do rDNA 16S, utilizando técnicas dependentes e independentes de cultivo (GreenGenes, última atualização em outubro de 2014). Recentemente, cinco novos filos foram propostos: Aigarchaeota (Nunoura et al., 2010), Diapherotrites (pMC2A384), Parvarchaeota, Aenigmarchaeota (DSEG) e Nanohaloarchaeota (Rinke et al., 2013). O conhecimento da diversidade desse domínio aumentou consideravelmente com a abordagem independente de cultivo, devido à dificuldade para cultivar seus membros (Madigan et al., 2004).

Analisar a comunidade procariótica, que apresenta a maior diversidade biológica da Terra, pode ampliar o entendimento da interação desses microrganismos em ambientes naturais contaminados com metais. Além disso, é de grande interesse o conhecimento da diversidade microbiana, incluindo informações da biogeografia e biotecnologia ambiental. Neste contexto, os

microrganismos nativos podem ser utilizados na biorremediação e como bioindicadores de ambiente impactado pela atividade mineradora.

2.6 – A ecologia microbiana na era Genômica

Os microrganismos dificilmente vivem isolados em seu ambiente natural, estando sempre interagindo entre si, com outros seres vivos e com fatores abióticos (pH, temperatura, dentre outros), formando comunidades complexas no ambiente que habitam (Shapiro e Dworkin, 1997; Rudi et al., 2007). Além disso, os microrganismos são os seres vivos mais abundantes da Terra (10^{30} células), sendo considerados essenciais para a manutenção da biosfera (Whitman et al., 1998). Até o momento, a extensão do conhecimento de como eles agem no ambiente é limitado. Portanto, é essencial gerar informações para entender o papel desses microrganismos e os processos nos quais eles estão envolvidos em seu habitat.

A Metagenômica é a ciência que utiliza a Genômica, ciência que surgiu há cerca de 30 anos permitindo o mapeamento e análise de genomas. A Metagenômica estuda o conjunto genômico de todos os organismos que habitam um determinado ambiente amostrado e suas diversas interações (Medini et al., 2008; Xu, 2011). O uso do gene de rRNA 16S como uma ferramenta para analisar filogeneticamente a comunidade microbiana de uma amostra ambiental foi realizado incialmente por Pace e colaboradores (Pace et al., 1985). Apesar disso, apenas em 1998, o termo Metagenômica foi usado pela primeira vez para nomear a técnica independente de cultivo utilizada para investigar comunidades microbianas em amostras ambientais no nível de DNA (Handelsman et al. 1998).

O objetivo principal de um projeto Metagenômico é fornecer um modelo descritivo e preditivo da taxonomia e do metabolismo de um ecossistema para tentar descrever o conjunto completo de interações dos microrganismos com o ambiente (Gilbert & Dupont, 2011). Analisar o metagenoma das comunidades presentes em ambientes contaminados com metais, como arsênio,

pode ampliar o entendimento das interações entre os microrganismos desse ambiente e seu papel na mobilização desses metais. Além disso, conhecer a diversidade microbiana é de grande interesse para a biotecnologia ambiental, já que os microrganismos nativos podem ser utilizados como bioindicadores ou em processos de biorremediação. Atualmente, a Metagenômica pode ser dividida em três áreas de pesquisa distintas de acordo com a metodologia utilizada: análise *shotgun* de todos os genomas presentes em uma determinada amostra ambiental; prospecção de funções microbianas específicas no ambiente; e pesquisa utilizando marcadores filogenéticos ou funcionais para adqurir informações da comunidade microbiana na amostra ambiental (Jones et al., 2009; Gupta e Vakhlu, 2011; Suenaga, 2012).

A Genômica e a Metagenômica microbiana vêm se destacando no mundo científico com o surgimento das novas tecnologias de sequenciamento. O sequenciamento de DNA surgiu no final da década de 1970 (Sanger 1977) e evoluiu tornando essa técnica mais eficiente, rápida e barata. O sequenciamento de terminação de cadeia, conhecida como "Sanger", é atualmente considerada a mais tradicional dentre as técnicas de sequenciamento disponíveis. Essa técnica produz sequências de 800 a 1000 pb utilizando dideoxinucleotídeos que atuam como terminadores de cadeia, impedindo a continuação da reação (Peixoto, 2011; Liu et al, 2012). Diversos metagenomas foram efetuados utilizando o sequenciamento de "Sanger" para investigar ambientes naturais, como drenagem ácida de mina, Mar de Sargasso, dentre outros (Venter et al., 2004; Edwards et al., 2007; Gilbert e Dupont, 2011). Embora esses estudos tenham sido definidos como Metagenômicos, eles foram limitados na amplitude da diversidade amostrada devido ao alto custo do sequenciamento tradicional (Mardis, 2008).

Atualmente, uma nova geração de sequenciadores automatizados, considerados *high throughput* sequencing (sequenciamento em larga escala), e o desenvolvimento de ferramentas de bioinformática tornaram o sequenciamento de genomas e metagenomas mais acessível,rápido e menos oneroso (Xu, 2011). O uso dos sequenciadores de nova geração tem definindo o futuro dos projetos em Metagenômica, além de gerar informações muito mais abrangentes sobre as

comunidades microbianas e suas funções, quando comparado com os projetos de Metagenômica utilizando a tecnologia de "Sanger". Além disso, os sequenciadores de nova geração permitem desvendar o vasto potencial dos microrganismos, cultiváveis e não cultiváveis, incluindo os mais raros, possibilitando a descoberta de mais genes biotecnologicamente interessantes e novas vias metabólicas (Chistoserdova, 2010).

As tecnologias de sequenciamento de nova geração são capazes de sequenciar enomes quantidades de DNA em pouco tempo, sendo atualmente utilizadas as seguintes plataformas: 454 GS FLX da Roche, Sistema HiSeq e MiSeq da Illumina, Applied Biosystems SOLiD™ System e Ion Torrent Personal Genome Machine (PGM) (Mardis, 2008; Jünemann et al., 2012). As diferenças fundamentais entre os sequenciadores de nova geração são o comprimento das sequências e o perfil de erro exclusivo para cada plataforma. O rendimento dos sequenciadores de nova geração, com relação ao número de sequências e o total de bases geradas por corrida, é significativamente maior do que o sequenciamento de "Sanger" e varia dependendo do sequenciador de nova geração utilizado (Mardis, 2008).

A plataforma SOLiD (Sequencing by Oligo Ligation Detection) usa bibliotecas de fragmentos ligados a um adaptador e PCR em emulsão para a amplificação dos fragmentos antes do sequenciamento. Diferente de outras plataformas, a reação de sequenciamento do SOLiD é catalisada por uma DNA ligase, e não uma polimerase. Outra característica dessa plataforma é a presença de quatro fluoróforos apresentando 16 combinações de nucleotídeos gerando, assim, sequências em formato *colorspace*. Esse formato permite uma alta precisão das sequências geradas por essa plataforma (Mardis, 2008).

As plataformas Illumina adotam o sequenciamento por síntese, ou seja, utilizam DNA polimerase e nucleotídeos marcados com diferentes fluoróforos. Os fragmentos de DNA ligados a adaptadores fixos são desnaturados em DNA fita simples, inseridos e fixados à *flowcell*, lâmina de vidro onde ocorre o sequenciamento. Em seguida, é feita a amplificação dos fragmentos de DNA ligados à *flowcel* através de uma estrutura em forma de "ponte" para formar *clusters* que contém

fragmentos de DNA clonados (Figura 7). Os nucleotídeos com corantes fluorescentes diferentes são incorporados e o sinal gerado é captado e interpretado pelo sequenciador (Liu et al, 2012).

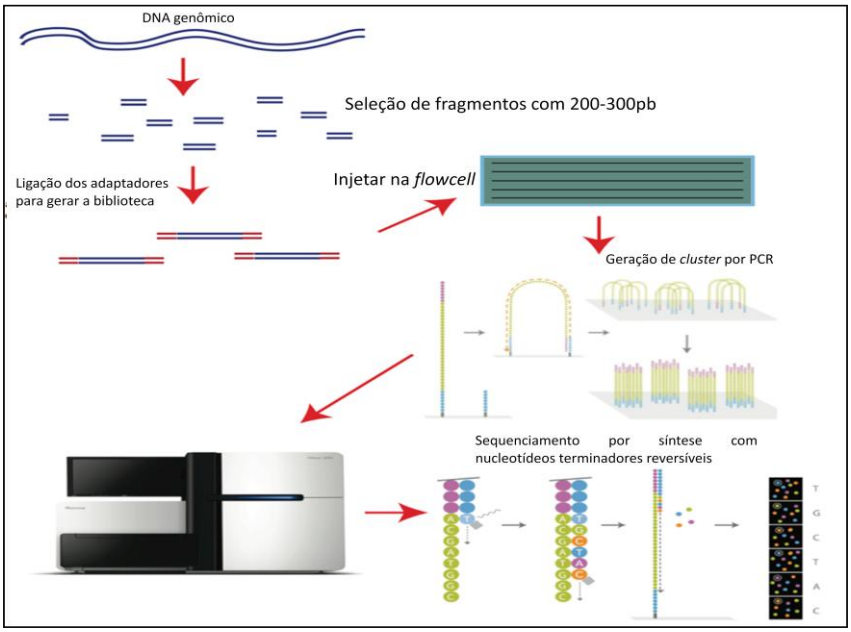

Figura 7 - Passos da plataforma de sequenciamento MiSeq Illumina. Fonte: Brown, 2012, modificado.

Neste contexto, o presente estudo revelou a diversidade do microbioma e dos genes envolvidos no metabolismo do arsênio em sedimento do Córrego da Mina, historicamente contaminado por metais pela atividade mineradora.

3 - Objetivos

3.1 - Objetivo Geral

Investigar a diversidade procariótica de sedimento de Córrego da Mina contaminado, principalmente por arsênio, utilizando abordagens clássica de cultivo, com ênfase em bactérias resistentes ao arsênio e os genes responsáveis pela sua transformação, e metagenômica usando sequenciamento de nova geração.

3.2 - Objetivos Específicos

Capítulo 1 - Bacteria and genes involved in arsenic speciation in sediment impacted by long-term gold mining

- Caracterizar as propriedades físico-químicas e determinar as concentrações de metais presentes nas amostras de água e sedimento do Córrego da Mina;
- Obter culturas do sedimento enriquecidas com AsIII e AsV e isolar bactérias resistentes ao arsênio inorgânico;
- Determinar a concentração inibitória mínima (CIM) de AsIII e AsV;
- Caracterizar os isolados quanto à oxidação de AsIII e/ou redução de AsV;
- Caracterizar o genótipo dos isolados através da presença dos genes *arsC*, *arrA* e *aioA*;
- Classificar taxonomicamente e analisar filogeneticamente os isolados usando o gene de rRNA 16S;
- Obter bibliotecas metagenômicas dos genes *arsC, arrA* e *aioA* do sedimento do Córrego da Mina;
- Analisar filogeneticamente as sequências destes genes obtidas por abordagens cultivável e metagênomica.

3.2.2 - Capitulo 2 - Metagenome of a microbial community inhabiting a metal-rich tropical stream sediment

- Analisar a diversidade metabólica da comunidade microbiana através do perfil fisiológico usando o Biolog Ecoplate;
- Obter DNA total da amostra de sedimento do Córrego da Mina e sequenciar o microbioma presente no sedimento utilizando os sequenciadores de nova geração Applied Biosystems SOLiD^{TM} System e MiSeq da Illumina;
- Otimizar e quantificar as comunidades de bactéria e arqueias presentes no sedimento do Córrego da Mina através de qPCR do gene de rRNA 16S;
- Analisar taxonomicamente e funcionalmente o microbioma da comunidade de microrganismos do sedimento;

4 – Capítulos

4.1 - Capítulo 1

Bacteria and genes involved in arsenic speciation in sediment impacted by long-term gold mining

OPEN ACCESS Freely available online

PLOS ONE

Bacteria and Genes Involved in Arsenic Speciation in Sediment Impacted by Long-Term Gold Mining

Patrícia S. Costa[1], Larissa L. S. Scholte[2], Mariana P. Reis[1], Anderson V. Chaves[1], Pollyanna L. Oliveira[1], Luiza B. Itabayana[1], Maria Luiza S. Suhadolnik[1], Francisco A. R. Barbosa[1], Edmar Chartone-Souza[1], Andréa M. A. Nascimento[1]*

1 Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais; Belo Horizonte, Brazil, **2** Grupo de Genômica e Biologia Computacional, Centro de Pesquisas René Rachou (CPqRR), Fundação Oswaldo Cruz (FIOCRUZ), Belo Horizonte, Brazil

Abstract

The bacterial community and genes involved in geobiocycling of arsenic (As) from sediment impacted by long-term gold mining were characterized through culture-based analysis of As-transforming bacteria and metagenomic studies of the *arsC*, *arrA*, and *aioA* genes. Sediment was collected from the historically gold mining impacted Mina stream, located in one of the world's largest mining regions known as the "Iron Quadrangle". A total of 123 As-resistant bacteria were recovered from the enrichment cultures, which were phenotypically and genotypically characterized for As-transformation. A diverse As-resistant bacteria community was found through phylogenetic analyses of the 16S rRNA gene. Bacterial isolates were affiliated with *Proteobacteria*, *Firmicutes*, and *Actinobacteria* and were represented by 20 genera. Most were AsV-reducing (72%), whereas AsIII-oxidizing accounted for 20%. Bacteria harboring the *arsC* gene predominated (85%), followed by *aioA* (20%) and *arrA* (7%). Additionally, we identified two novel As-transforming genera, *Thermomonas* and *Pannonibacter*. Metagenomic analysis of *arsC*, *aioA*, and *arrA* sequences confirmed the presence of these genes, with *arrA* sequences being more closely related to uncultured organisms. Evolutionary analyses revealed high genetic similarity between some *arsC* and *aioA* sequences obtained from isolates and clone libraries, suggesting that those isolates may represent environmentally important bacteria acting in As speciation. In addition, our findings show that the diversity of *arrA* genes is wider than earlier described, once none *arrA*-OTUs were affiliated with known reference strains. Therefore, the molecular diversity of *arrA* genes is far from being fully explored deserving further attention.

Citation: Costa PS, Scholte LLS, Reis MP, Chaves AV, Oliveira PL, et al. (2014) Bacteria and Genes Involved in Arsenic Speciation in Sediment Impacted by Long-Term Gold Mining. PLoS ONE 9(4): e95655. doi:10.1371/journal.pone.0095655

Editor: Celine Brochier-Armanet, Université Claude Bernard - Lyon 1, France

Received June 3, 2013; **Accepted** March 31, 2014; **Published** April 22, 2014

Funding: Sources of funding: FAPEMIG APQ 00801/12, CNPq n°472411/2012-8, CNPq/INCT no 15206-7. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: amaral@ufmg.br

Introduction

Arsenic occurs naturally in the earth's crust and is widely distributed in the environment [1,2]. Natural mineralization and microorganisms enhance arsenic mobilization in the environment, but human interventions, such as gold mining, have aggravated the environmental arsenic contamination arousing health concerns. Water pollution by arsenic is one of the major challenges for public health, primarily due to its carcinogenic potential at low doses [3,4,5,6]. According to Nordstrom [7] over 50 million people in the world are at risk from drinking arsenic-contaminated water. Moreover, given that arsenic has a variety of valence states (+V, +III, 0, −III) with different physicochemical properties, the removal of arsenic from contaminated water bodies is yet a challenge.

In nature, microorganisms have developed different response mechanisms to metabolize As, mainly via reduction and oxidation reactions, leading to its speciation [8]. Previous studies have regarded As speciation as a result of microbial activity in the environment, including some derived from gold-mining activities [1,2]. However, few bacterial genera involved in As transformation have been found at any of the sites studied [9–13]. Thus, a more comprehensive knowledge on the structure of the bacterial community involved in As-transformation in gold-mining sites remains warranted.

The arsenate (AsV) reducing pathways known are the detoxification (*arsC* gene) and the dissimilatory respiration (*arrA/B* genes). The organization of *ars* operons varies greatly between taxa, and the core genes include *arsR*, *arsB* and *arsC*, whereas *arsD* and *arsA* genes can eventually be found [1]. The *ars*C gene encodes the enzyme AsV reductase, which is located in the cytoplasm and is responsible for the biotransformation of AsV to AsIII. This enzyme together with a transmembrane efflux pump, encoded by *arsA* and *arsB* genes, is the most common As transformation mechanism in the environment [2,14–16]. Moreover, *arrA/B* genes encode a periplasmic AsV reductase that works during anaerobic respiration using AsV as the final electron acceptor for energy generation [17]. The AsV dissimilatory respiration reduction has already been described for many bacterial phyla, including obligatory and facultative anaerobic bacteria and some archaea [1].

The microbial oxidation of AsIII was first reported in 1918 and can be mediated by two distinct enzymes: AioBA, hardly studied, and ArxAB, recently described by Zargar et al. [18]. Both enzymes

have been found in several heterotrophic and chemolithoautotrophic bacterial species [19–22]. Aerobic AsIII oxidation is catalyzed by arsenite oxidase, which uses O_2 as terminal electron acceptor, and is encoded by *aioB/A* genes, formerly referred to as *aoxA/B*, *aroB/A* and *asoB/A* genes [20,23]. ArxAB detected in AsIII oxidizing bacteria in anoxic conditions, in which nitrate or chlorate reduction is coupled to AsIII oxidation in the chemolithotrophs [24,25]. Interestingly, members of the genus *Ectothiorhodospira* are able to use AsIII as electron donor for anoxygenic phototrophic growth [26]. According to Zargar et al. [18] the *arxA* gene is more closely related to *arrA* than to *aioA* genes.

In this research, we bioprospected As-resistant bacteria from As-enrichment culture of sediments collected from a stream located at the Brazilian gold mining area known as the Iron Quadrangle (IQ, Minas Gerais state), one of the world's largest mining regions. Much concern exists about As-contamination of gold-mining sites in this area because it is estimated that at least 390,000 tons of As have been released into this area since the beginning of gold-mining activity in the 17th century [27]. We also investigated the diversity of As-transforming genes using metagenomic strategies. This included the genes for arsenite oxidase (*aioA*) and arsenate reductases (*arsC* and *arrA*).

Materials and Methods

Ethics Statement

For sampling in Mina stream, no specific permit was required for the described field study. The study location is not privately-owned or protected in any way and we confirm that the field study did not involve endangered or protected species.

Study Area and Sampling

Mina stream (19°58′46.80″S–43°49′17.07″W) is a natural body of water located at the Velhas River Basin (IQ, Minas Gerais state, Brazil) and characterized as backwater (Figure S1). This stream was chosen because is located near a historically impacted gold-mining area. Moreover, previous investigations [28] reported As concentrations superior to those permitted by Brazilian law (Conselho Nacional do Meio Ambiente – CONAMA) and by Canadian Environmental Quality Guidelines (Canadian Council of Ministers of the Environment– CCME).

Bulk water and superficial sediment samples (up to 1.0 cm depth) were collected on 13 July 2011, during the dry season. The typical sediment core can be divided into three zones: oxic, suboxic and anoxic [29]. According to literature the thick oxic zone can extend from several mm up to 10 cm [30,31]. In this work the sampling site was shallow (20 cm) and therefore highly influenced by the nutrients and oxygen concentrations of the water body. The analyzed sediment was taken from the upper part, representing the oxic zone. Samples were collected aseptically at three points at 1m distance from each other, subsequently pooled in a single sample, and stored at 4°C for bacterial analysis or at −20°C for chemical and molecular analyses.

To assess the bulk water conditions physicochemical characteristics such as temperature, pH, and dissolved oxygen (DO) concentration were measured *in situ* with a multiprobe (Horiba, model U-22) [30]. Concentrations of total nitrogen (TN), total phosphorus (TP), ammonium (NH_4^+-N), nitrite (NO_2-N), nitrate (NO_3-N), and soluble reactive phosphorus (PO_4-P) were measured as previously described [32,33]. Metal and metalloid concentrations of water and sediment samples were determined by using an inductively coupled plasma-optical emission spectrometer (ICP-OES, Optima 7300 DV, PerkinElmer).

Arsenic Enrichment and Isolation

Sediment (10 g) samples were added to Erlenmeyer flasks containing 100 mL of CDM medium (0.012 mM Fe_2SO_4, 7 mM Na_2SO_4, 0.0574 mM K_2HPO_4, 9.5 mM $NaHCO_3$, 18.7 mM NH_4Cl, 8.12 mM $MgSO_4$, 0.457 mM $CaCl_2$ and 44.6 mM sodium lactate as organic carbon source, pH 7.2) with either 2 mM sodium arsenite or 10 mM sodium arsenate and incubated at 28°C for seven days. Then, serial 10-fold dilutions of the enrichment cultures were plated onto CDM agar media (1.5% agar) amended with 2 mM sodium arsenite or 10 mM sodium arsenate to selectively enrich and isolate AsIII- and AsV-resistant bacteria. Plates were incubated at 28°C for five days. The resulting colonies were repeatedly streaked on the same medium to accomplish their purification. The bacterial isolates from AsIII- and AsV-resistant bacteria (named MS-AsIII and MS-AsV, respectively) were stored at −20°C in 25% glycerol.

DNA Extraction from the Cultures and Sediment

Genomic DNA was extracted and purified from each MS-AsIII and MS-AsV isolate using a protocol previously described [34]. Additionally, metagenomic DNA was extracted from 10 g (wet weight) of sediment using the PowerSoil DNA Extraction Kit (MO BIO Laboratories, USA) according to the manufacturer's instructions. Total DNA from the MS-AsIII and MS-AsV isolates and sediment were quantified by absorbance at 260 nm using a NanoDrop Spectrophotometer (NanoDrop Technologies). DNA purity was assessed using the A260/A280 and A260/A230 ratios. DNA was stored at −20°C until further processing.

PCR Amplification and Construction of Clone Libraries

Briefly, touchdown PCR was carried out by amplifying bacterial MS-AsIII and MS-AsV isolates 16S rRNA gene fragments using the conditions previously described by Freitas *et al.* [35]. The reactions were performed using the bacterial-targeted primer set 8F (5′-AGAGTTTGATYMTGGCTCAG-3′) and 907R (5′-CCGTCAATTCMTTTRAGT-3′) [36]. Taq DNA polymerase and dNTPs were purchased from Fermentas (Canada) and used in all the PCR reactions.

Metagenomic and genomic DNA were used as template for PCR employing the *arsC*, *arrA* and *aioA* genes for construction of clone libraries and genotypic characterization of the bacterial MS-AsIII and MS-AsV isolates. PCR reactions targeting the *arsC*, *arrA* and *aioA* genes were carried out using primers and conditions as previously described by Sun *et al.* [37], Malasarn *et al.* [17] and Hamamura *et al.* [20], respectively. The *arsC* gene examined was of the glutaredoxin-dependent arsenate reductase enzyme, ArsC, from *Escherichia coli* R773 plasmid. The primer chosen has been successfully applied in several investigations of a variety of environmental samples [37,38].

The amplicons of *arsC*, *arrA*, and *aioA* genes were gel-purified using the Silica Bead DNA Gel Extraction Kit (Fermentas, Canada). PCR products were cloned into the vector pJET1.2/blunt (Fermentas, Canada), and propagated with *Escherichia coli* XL1-Blue electrocompetent cells according to the manufacturer's instructions.

Sequencing and Phylogenetic Analysis

Partial 16S rRNA, *arsC*, *arrA*, and *aioA* gene sequences were obtained using BigDye Terminator Cycle Sequencing kit (Life Technologies, USA) according to the manufacturer's instructions. The nucleotide sequences were quality checked and submitted to GenBank with accession numbers from KC577613 to KC577798. The 16S rRNA gene sequences were analyzed through blastn

(http://www.ncbi.nlm.nih.gov) and Classifier search tool (http://rdp.cme.msu.edu) to determine their phylogenetic affiliation. The *arsC*, *arrA*, and *aioA* gene sequences were compared with those available at the GenBank databases using blastn and blastx tools (http://www.ncbi.nlm.nih.gov) to retrieve potential homologs. Operational taxonomic unities (OTUs) from As gene clone libraries were defined with DOTUR software [39] using a cut-off threshold of ≥97% identity. Coverage of the clone libraries was calculated using the equation $C = 1-(n/N)\times100$, where n is the number of unique OTUs and N is the number of sequences analyzed in the library [40].

In total, five fasta files were obtained containing *arsC*, *arrA*, *aioA*, MS-AsIII 16S rRNA, and Ms-AsV 16S rRNA gene sequences. Due to the short length of *arsC* and *arrA*amino acid sequences obtained in this study, added to the high similarity of some OTUs and isolates, we decided to reconstruct the phylogenetic relationships of As metabolism genes using nucleotide sequences to increase the phylogenetic signal and avoid overparameterization.

Sets of nucleotide sequences were independently aligned using MAFFT 7 with iterative refinement by the G-INS-i strategy [41]. Multiple sequence alignments were manually refined using Jalview [42]. To optimize the datasets for evolutionary analyses we removed redundancy and sequences too distantly related using the Decrease Redundancy tool available as a resource at ExPaSy (www.expasy.org). The Decrease Redundancy parameters were set as 99 for "% max similarity" and 30 for "% min similarity". Identical sequences were clustered as single OTUs and filtered alignments were further used in phylogenetic analyses. Identifiers of filtered sequences were later included into the phylogenetic tree. To reconstruct phylogenetic trees we used the maximum likelihood method (ML) as implemented in PhyML [43]. For the phylogenetic reconstruction we tested seven different evolutionary models (HKY85, JC69, K80, F81, F84, TN93, and GTR) using the jModelTest 2 software [44]. The evolutionary model best fitting the data was determined by comparing the likelihood of tested models according to the Akaike Information Criterion (AIC). Statistical support value for each node was computed by approximate likelihood ratio test (aLTR). Trees were visualized and edited using the FigTree software (tree.bio.ed.ac.uk/software/figtree).

Susceptibility and Arsenic Transformation Tests

Minimum inhibitory concentrations (MIC) were established, in triplicate, by the agar dilution method in CDM with 1×10^5 CFU ml^{-1} as standard inoculums. CDM plates were supplemented with increasing concentrations (from 2 mM to 1024 mM) of AsIII or AsV and incubated at 28°C for seven days. MIC was defined as the lowest AsIII or AsV concentration that completely inhibited bacterial growth.

The ability to oxidize AsIII and reduce AsV was investigated using a qualitative screening according to [45]. To achieve that, bacterial MS-AsIII and MS-AsV isolates were grown in CDM broth with 100 mg l^{-1} 2 mM sodium arsenite or 100 mg l^{-1} sodium arsenate until an optical density of 0.4 at 595 nm was reached. After that, 20 µl of 0.01 mol l^{-1} of potassium permanganate solution were added in 1 ml of bacterial culture. The data were interpreted according to the change in medium color, i.e., a pink color indicated a positive oxidation of AsIII and a yellow color indicated a positive reduction of AsV.

Results

Environmental Parameters

The physicochemical characteristics of the water and sediment samples from the Mina stream are presented in Tables 1 and 2. Data displayed on Table 1 revealed that metal concentrations in the Mina stream exceeded the maximum allowable concentrations established by Brazilian and Canadian environmental regulations [46,47] for sediment and water. Al, Mn, Fe, Cu, As and Zn were the metals present in the highest concentrations in the sediment sample analyzed.

The physicochemical analysis revealed that the Mina stream can be characterized as a mesothermal oxidized environment with highly oxygenated and circum-neutral waters (Table 2). Nitrogen and phosphorus ratio was greater than nine (Table 2). According to Salas & Martino [48], this ratio indicates that the phosphorus was the most limiting nutrient and that the stream can be classified as eutrophic.

Phylogenetic Affiliation

In total, 123 bacterial isolates were recovered from the enrichment cultures (68 and 55 from the MS-AsIII and MS-AsV, respectively). Partial 16S rRNA gene sequences used for phylogenetic analysis were approximately 600 bp long and spanned the V2 to V4 variable regions. MS-AsIII and MS-AsV isolates were categorized into three phyla: *Proteobacteria* (56% and 59%, respectively, includes *alpha*, *beta*, and *gamma-Proteobacteria*), *Firmicutes* (36% in both enrichment cultures), and *Actinobacteria* (8% and 5%). Twenty genera represented these phyla in the Mina stream sample analyzed. Differences in the bacterial composition between the MS-AsIII and MS-AsV enrichment cultures were detected (Table 3 and Figure 1). The resulting Venn diagram shows that a higher bacterial diversity was observed in the MS-AsIII than in the MS-AsV enrichment cultures. Eight genera were specifically found in MS-AsIII and seven were shared between the culture systems (Figure 1).

Dominant genera in MS-AsV were *Bacillus* (26%), *Pseudoxanthomonas* (18%), and *Brevundimonas* (16%). The predominant population in MS-AsIII was *Bacillus* (30%), followed by *Pseudoxanthomonas* (25%). The other bacteria related to MS-AsIII and MS-AsV are listed in Table 3. Although the *Proteobacteria* phylum was the most diverse and dominant, the *Bacillus* (29%) genus was the most abundant and diverse among the genera because it harbored eight identified species.

Characterization of As-reducing and Oxidizing Isolates and Identification of their Genes Involved in As Metabolism

The MICs for the MS-AsIII and MS-AsV isolates were determined. The highest MIC was found for AsV in which 94% of the isolates exhibited values ≥256 mM, whereas 90% of the isolates displayed MICs ranging from 32 mM to 64 mM for the most toxic AsIII.

The As-transformation ability of the isolates was determined with a qualitative test that revealed that 72% of the isolates were AsV-reducing, whereas 20% were AsIII-oxidizing. Of those, 8% presented AsV-reducing as well as AsIII-oxidizing activities. Among the 20 genera identified in both MS-AsIII and MS-AsV enrichment cultures, *Acidovorax* and *Achromobacter* presented only AsIII-oxidizing activity. No As-transformation activity was found in 8% of the total of MS-AsIII and MS-AsV isolates (123) (Table 3).

The molecular analysis of the MS-AsIII and MS-AsV isolates unveiled that the *arsC* gene was the most frequent (85%), followed by *aioA* (20%) and *arrA* (7%) (Table 3). Of those, *Bacillus* was the

Table 1. Metal concentration from sediment and water of Mina Stream and limits permitted by law.

Metals	Sediment (mg kg^{-1})	CONAMA* (mg kg^{-1})	Water (mg l^{-1})	CONAMA** (mg l^{-1})
Fe	492.8	NE	0.52	15
Ni	9.0	18	<0.1	2
Mn	1284.5	NE	1.45	1
Cu	387.7	35.7	0.19	1
Pb	8.7	35	NE	0.5
Cd	<2.5	0.6	<0.1	0.2
Zn	180.9	123	0.2	5
Al	2343.2	NE	<0.5	NE
As	297.1	5.9	<0.1	0.5
Cr	17.3	37.3	<0.1	1
Hg	<2.5	0.17	<0.1	0.01

NE – Not established.
*CONAMA resolution 344/04.
**CONAMA resolution 430/11.
doi:10.1371/journal.pone.0095655.t001

only genus harboring all three genes, and *Shewanella* was the only genus which did not harbor the most common gene (*arsC*) in the isolate analyzed (MS-AsIII-61). *Achromobacter* and *Acidovorax* both harbored the *aioA* gene, confirming the phenotypic data. *Thermomonas* and *Pannonibacter* also harbored As resistance genes.

General Features of Clone Libraries

To unveil the molecular diversity of genes involved in As metabolism in the Mina stream sediment, three clone libraries for *arsC*, *arrA*, and *aioA* genes were constructed. One hundred sixty-four sequences were analyzed after quality control and the removal of chimeric sequences. The coverage values of the three libraries (80%, 70% and 63%, respectively for *arsC*, *arrA*, and *aioA*) indicated that most of the diversity of these genes was detected.

Blastx analysis of *arsC*, *aioA*, and *arrA*-OTUs revealed high similarity with sequences from glutaredoxin-glutathione arsenate reductase (from 76 to 100%), molybdopterin-binding arsenite oxidase (from 71 to 96%), and respiratory arsenate reductase (from 64 to 98%) (Tables S1, S2, and S3 in Tables S1). The sequences corresponding to *arsC* were associated with *arsC* harboring different bacterial taxa from a variety of environments. The *aioA*-OTUs were closely related to uncultured and cultured clones from As contaminated environments. Furthermore, all *arrA*-OTUs were closely related to uncultured clones from rock biofilms of an ancient gold mine and Cache Valley Land Fill sediments, both arsenic-contaminated environments.

Table 2. Physicochemical parameters from water of Mina Stream.

Parameters	Water
pH	6.2
Conductivity (µs cm^{-1})	2151
Temperature (°C)	18.0
Dissolved Oxygen (mg l^{-1})	9.1
Redox (mV)	215
NO_3^- -N (µg l^{-1})	3103.8
NO_2^- -N (µg l^{-1})	161.3
NH_4^+ -N (µg l^{-1})	829.5
PO_4^{3-} -P (µg l^{-1})	2.3
Total P (µg l^{-1})	77.6
Total N (µg l^{-1})	2916.8

doi:10.1371/journal.pone.0095655.t002

Phylogenetic Analyses of 16S rRNA, *arsC*, *aioA*, and *arrA* Genes Sequences

In this study we have amplified, sequenced and reconstructed the evolutionary relationships of 16S rRNA and As metabolism genes encoded by As-resistant bacteria retrieved from a stream located at the Brazilian gold mining area and cultivated on As-enrichment sediment's culture, as well as As metabolism genes of clone libraries. The phylogeny of the AsIII-resistant bacteria (MS-AsIII) 16S rRNA gene sequences was reconstructed from an alignment containing 57 operational taxonomic units and 719 sites, which represent 99 sequences (Fig. 2). Therefore, 42 sequences were considered redundant by the Decrease Redundancy tool (www.expasy.org). The phylogenetic tree reconstructed by using the maximum likelihood method as implemented in PhyML [43], shows sequence's clear separation into three strongly supported clades, which have representatives of the *Firmicutes*, *Actinobacteria*, and *Proteobacteria* phyla (Fig. 2). Similar results were obtained for the AsV-resistant bacteria (MS-AsV) 16S rRNA phylogenetic analysis (Fig. 3). The evolutionary history was based on an alignment containing 40 OTUs and 721 sites, representing 79 sequences (Fig. 3). The Decrease Redundancy tool filtered about 50% of the initially selected sequences. The resulting phylogeny also exhibits the presence of three well-supported clades containing *Firmicutes*, *Actinobacteria*, and *Proteobacteria* phyla representatives (Fig. 3).

Concerning evolutionary histories of As metabolism genes, the phylogenetic tree of *arsC* sequences was reconstructed with 48 nucleotide sequences and 352 sites, which represent 142 sequences (Fig. 4). TrN+I+G+F was selected as best fit model. The resulting phylogeny supports the hypothesis that horizontal gene transfer (HGT) seems to have played a role in the widespread distribution of *arsC* coding gene in *Actinobacteria* and *Proteobacteria*. Similar findings were retrieved on the phylogeny reconstructed for *arrA*

Table 3. Phylogenetic distribution of the bacterial isolates and their As-metabolism phenotype and genotype.

Enriched culture	Phylum	Genus	N° of isolates*	Phenotype**	Genotype***
MS-AsV	*Proteobacteria*	*Acinetobacter*	1	reducer	*arsC aioA*
		Brevundimonas	9	reducer (3)	*arsC*[9] *aioA* [1]
		Diaphorobacter	1	reducer	*arsC*
		Pannonibacter	2	reducer (2)	*arsC*[2]
		Pseudomonas	6	reducer (2)	*arsC*[6] *aioA* [2]
		Pseudoxanthomonas	10	reducer (6)	*arsC*[7] *aioA* [1]
		Stenotrophomonas	5	reducer (5)	*arsC*[5] *aioA* [2]
		Thermomonas	1	reducer	*arsC*
	Firmicutes	*Bacillus*	14	reducer (9), oxidizer (8)	*arsC*[12] *aioA* [6] *arrA* [1]
		Exiguobacterium	1	reducer	*arsC aioA*
	Actinobacteria	*Microbacterium*	4	reducer (4) oxidizer (1)	*arsC*[3] *aioA* [1]
		Rhodococcus	1	reducer	*arsC*
MS-AsIII	*Proteobacteria*	*Achromobacter*	1	oxidizer	*arsC aioA*
		Acidovorax	2	oxidizer(2)	*arsC*[2] *aioA* [1]
		Acinetobacter	4	reducer (3)	*arsC*[2] *aioA* [1]
		Comamonas	5	reducer (5)	*arsC*[5] *arrA* [2]
		Ochrobactrum	1	reducer	*arsC*
		Pseudomonas	3	reducer (3)	*arsC*[1] *aioA*[1]
		Pseudoxanthomonas	17	reducer (9) oxidizer (5)	*arsC*[14] *aioA*[2] *arrA* [1]
		Shewanella	1	reducer	*arrA*
		Stenotrophomonas	2	reducer (2) oxidizer (1)	*arsC*[1] *aioA* [1]
		Thermomonas	2	reducer (1)	*arsC*[2] *arrA* [1]
	Firmicutes	*Bacillus*	21	reducer (19) oxidizer (5)	*arsC*[21] *aioA*[2] *arrA* [3]
		Paenibacillus	1	reducer oxidizer (1)	*arsC*
		Staphylococcus	4	reducer (4) oxidizer (1)	*arsC*[3]
	Actinobacteria	*Micrococcus*	3	reducer (2) oxidizer (1)	*arsC*[1] *aioA*[1]
		Microbacterium	1	reducer	-

*The number represents the total of bacterial isolates identified.
**Values in parentheses indicate the number of As-redox isolates.
***Values in bracket indicate the number of isolates harboring As-metabolism genes.
doi:10.1371/journal.pone.0095655.t003

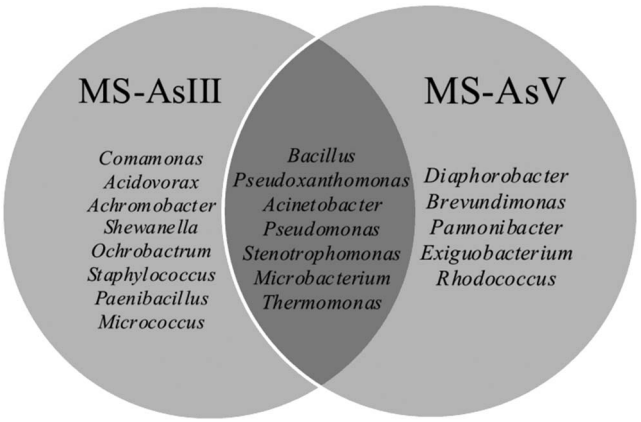

Figure 1. Venn diagram showing the exclusive and shared bacterial genera retrieved from MS-AsIII and MS-AsV enrichment cultures.
doi:10.1371/journal.pone.0095655.g001

sequences based on 47 nucleotide sequences and 242 sites where GTR+I+G+F was selected as best fit model (Fig. 5). On the other hand, the phylogenetic analysis of *aioA* sequences based on 72 nucleotide sequences and 543 sites shows two clades strongly supported: *alpha-* and *beta-proteobacteria* (Fig. 6) without clear evidence of HGT. For this analysis GTR+I+G+F was selected as best fit model. Interestingly, all putative *arrA* sequences obtained in this study (*arrA*- OTU) were more closely related to themselves or to sequences from uncultured bacteria, showing that more studies involving *arrA* sequences will be relevant to better understand the molecular diversity of those genes (Fig. 5).

Discussion

The environmental impact of gold mining is presently a major concern because its processes release toxic metals such as As in both soil and groundwater. Considering the relevance of bacteria in the speciation of As in aquatic environments, we bioprospected As-resistant bacteria and As-transforming genes originated from sediments impacted by long-term gold mining. Although some studies have focused in the identification of As-resistant bacterial communities in a long-term As-contaminated environment [9–

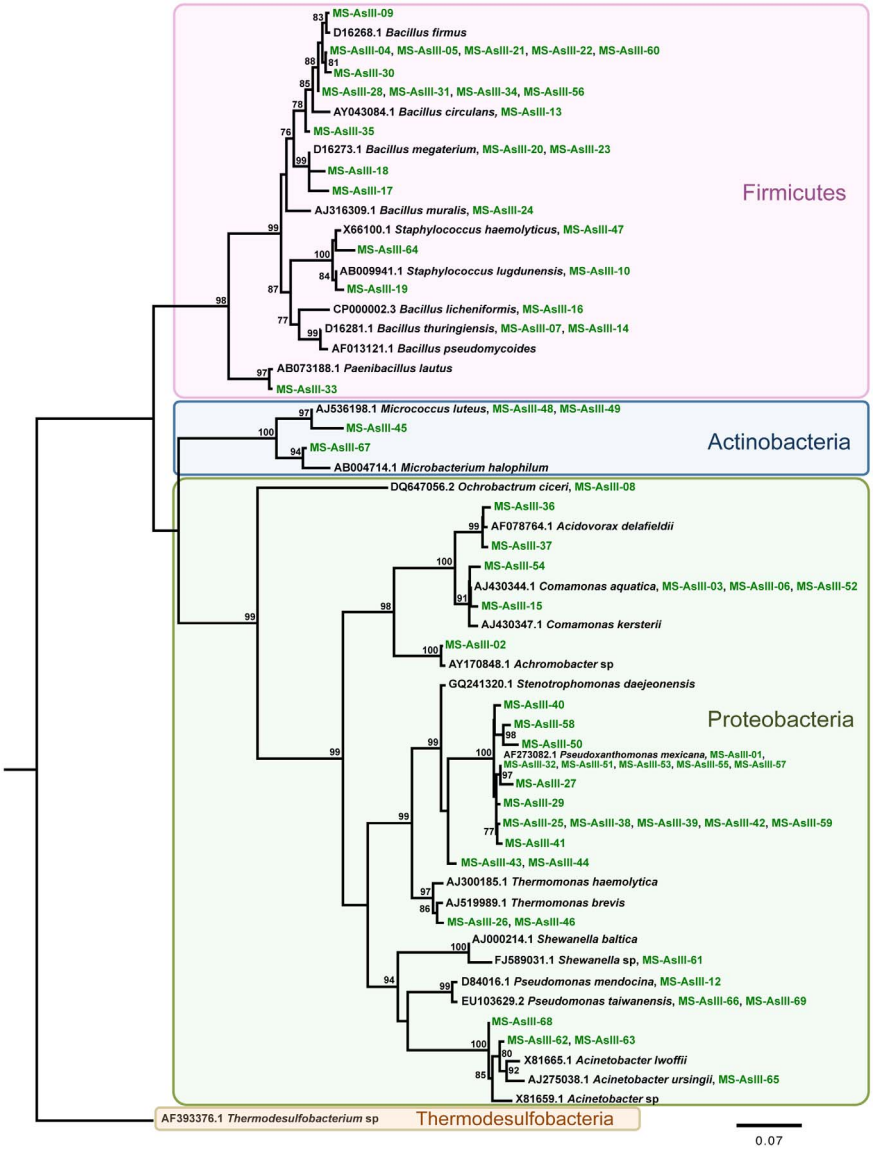

Figure 2. Evolutionary relationships of AsIII-resistant bacteria (MS-AsIII) 16S rRNA sequences. A total of 57 nucleotide sequences and 719 sites were analyzed. The phylogeny was reconstructed by maximum likelihoodand TIM3+I+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green. Different background colors highlight three well-supported clades: *Firmicutes*, *Actinobacteria*, and *Proteobacteria*. *Thermodesulfobacteria* was used as outgroup.
doi:10.1371/journal.pone.0095655.g002

13,49–52], the employment of combination of culture-based physiological and genomic approaches with metagenomic analysis in sediments collected from these areas are scarce [53,54]. In this study, we reveal a large number of phylogenetically distinct As-resistant bacterial genera retrieved from sediment collected from a stream in a long-term gold-mining area.

We found *Bacillus* as the dominant genus in both MS-AsIII and MS-AsV enrichment cultures. Members of *Bacillus* are often found in As-contaminated environments [9,11,16,54] being related to As-reduction and -oxidation, indicating that they are an essential component of As speciation in nature [54,55]. The observed abundance of *Bacillus* isolates harboring the *arsC* and *aioA* genes confirmed its ubiquity and high As-resistance in As-rich environments, as it is the case of Mina stream sediment. This suggests an important role for *Bacillus* in As speciation. It also points to a possible use of these natural isolates in future bioremediation projects.

A recent study of our group [56], using culture-independent approach to assess the prokaryotic diversity in Mina stream sediment, revealed the presence of the *Thermomonas*, *Acidovorax*, *Acinetobacter* and *Ochobactrum* genera also detected in the present study. Moreover, Bandyopadhyay *et al.* [57] have proposed a novel species of the *Pannonibacter* genus, *Pannonibacterindica*, which is able to grow in high concentrations of AsV. However, it should be noted that *Thermomonas* and *Panonnibacter* were not previously reported in the literature as As-transforming genera.

The phenotypic and genotypic characterization of the MS-AsIII and MS-AsV bacterial isolates revealed their ability to reduce and oxidize As. Most bacteria (85%) were AsV-resistant bacteria (ARB) harboring the *arsC* gene, responsible for the reduction of AsV to

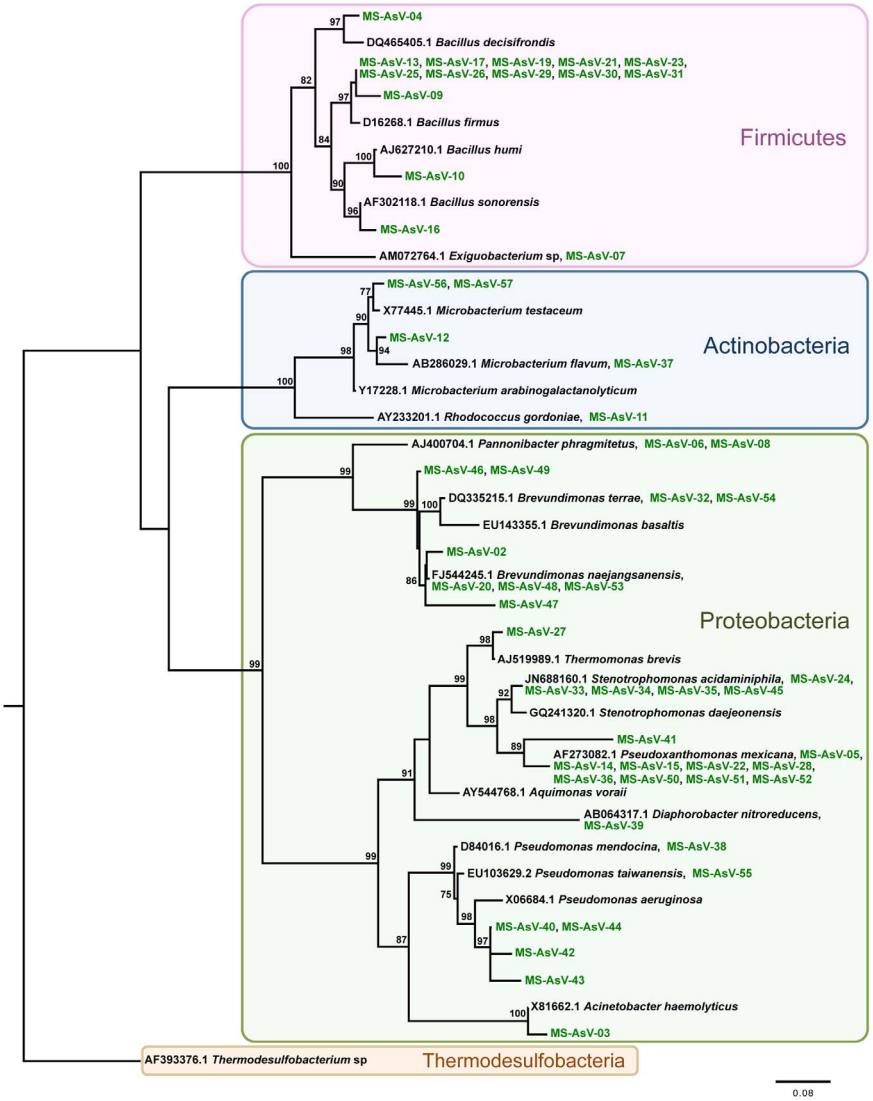

Figure 3. Evolutionary relationships of AsV-resistant bacteria (MS-AsV) 16S rRNA sequences. A total of 40 nucleotide sequences and 721 sites were analyzed. The phylogeny was reconstructed by maximum likelihood and TrN+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green. Different background colors highlight three well-supported clades: *Firmicutes*, *Actinobacteria*, and *Proteobacteria*. *Thermodesulfobacteria* was used as outgroup.
doi:10.1371/journal.pone.0095655.g003

AsIII, which is the most frequent detoxification reaction among bacteria in the environment [8]. Although the aerobic enrichment culture condition employed in this study could inhibit the growth of dissimilatory arsenate-reducing bacteria (DARB), it is likely that these bacteria were present because the *arrA* gene was detected. Several reports have evidenced DARB bioremediation potential of As-contaminated samples [2,54,58,59].

AsIII-oxidizing bacterial isolates were minority (20%). This result is in agreement with those reported by Silver & Phung [14], who suggest that most isolates from natural environments lack AsIII-oxidizing ability. In this study, all AsIII-oxidizing isolates were classified as heterotrophic AsIII oxidizers (HAO) spanning 11 genera. However, only isolates belonging to *Bacillus*, *Pseudoxanthomonas*, *Stenotrophomonas*, *Micrococcus*, *Achromobacter*, and *Acidovorax* genera co-presented the oxidizing phenotype and genotype. From an ecological perspective, oxidizing bacteria are important performers in As-contaminated environments because they promote transformation from AsIII into AsV.

In a few isolates (8%), oxidizing and reducing As-transformation activities were not observed in their phenotype and genotype. There are several possible explanations for this. First, the As-transformation gene expression observed in isolates grown in the laboratory is likely to be different from that encountered in these isolates in nature, because of the different conditions of these environments. Second, these differences may reflect mutations in the As-resistance genes studied. Third, alternative resistance genes may be expressed by these isolates [60].

The high diversity and adaptability of the bacterial community disclosed herein could be explained by the presence of multiple copies of As-resistance genes either on bacterial chromosomes or on plasmids as a consequence of pressure created by the long-term contamination that occurs in the Mina stream area. Nevertheless, further studies will be needed to establish this.

Figure 4. Evolutionary relationships of *arsC* sequences. A total of 48 nucleotide sequences and 352 sites were analyzed. The phylogeny was reconstructed by maximum likelihoodand TrN+I+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV)in green, and operational taxonomic unities (OTUs) from As gene clone librariesin blue. Different background colors highlight *Actinobacteria* and three *Proteobacteria* classes – *Gamma-*, *Beta*, and *alpha-proteobacteria*.
doi:10.1371/journal.pone.0095655.g004

Previous studies on As-resistance genes are associated with As-resistant cultivable isolates [10,16,59,61]. Considering that the vast majority of bacteria are uncultivable, this traditional approach has limited our understanding of the extreme functional diversity in natural bacterial communities. Therefore, a metagenomic approach to investigate the functional genes associated with As-transformation in nature is essential to further our current knowledge on this matter. The analysis of *arrA* sequences revealed that all of them exhibited similarity with those from uncultured organisms. This predominance of uncultured organisms indicates that *arrA* gene present in Mina stream sediment is expressed by unidentified DARB. The *arsC* sequences detected in the sediment were similar to those previously reported [37,62,63]. The primers used in this study amplified *aioA*-like sequences [20,64]. The *aioA* sequences were similar to several *aioA* genes of the *Proteobacteria* phylum. This finding is in agreement with Quéméneur *et al.* [65], who reported prevalence of AsIII-oxidizing *Proteobacteria* in mesophilic As-contaminated soils. However, it should be noted that *aioA* genes have been also detected in non-proteobacterial lineages [53,69].

Phylogenetic analyses' findings performed for MS-AsIII 16S rRNA, Ms-AsV 16S rRNA and As metabolism genes were consistent with findings obtained by similarity searches (blastx and blastn, respectively). Overall, the phylogenetic trees reconstructed

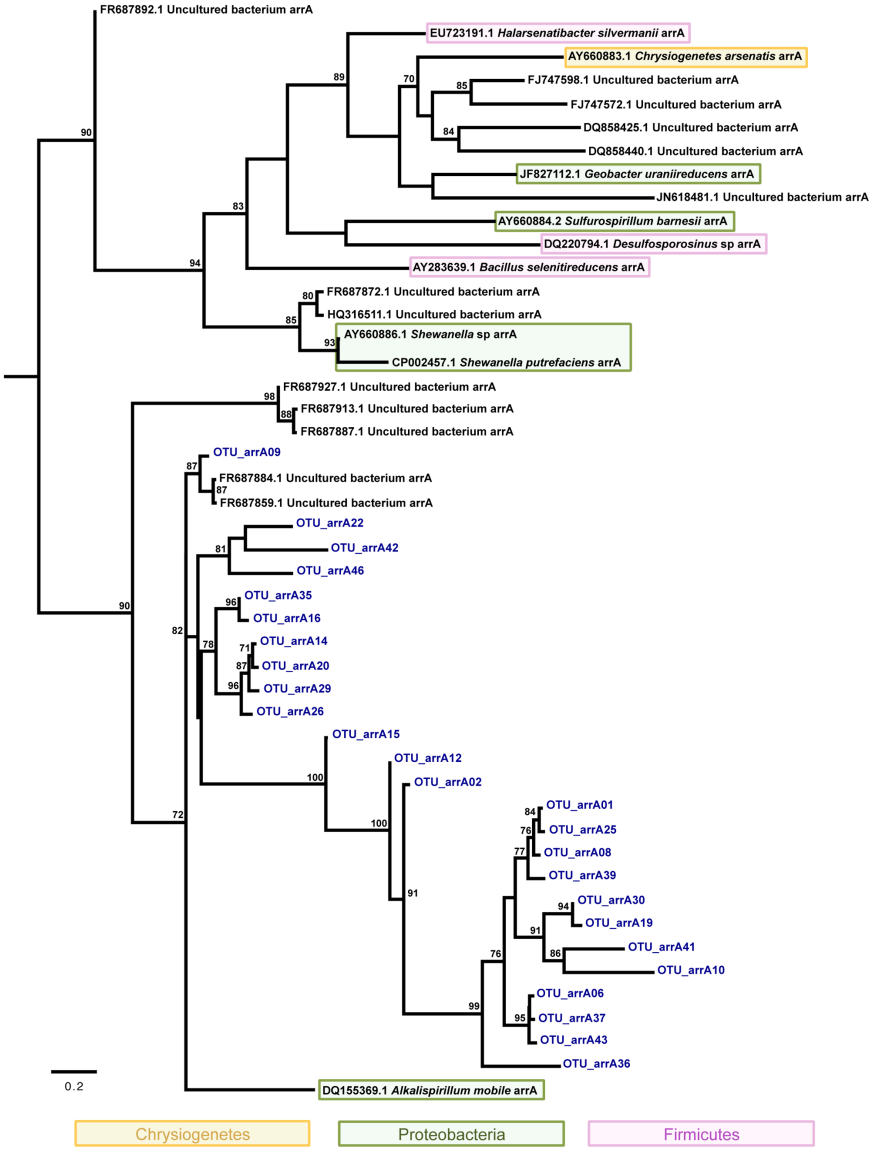

Figure 5. Evolutionary relationships of *arrA* sequences. A total of 47 nucleotide sequences and 242 sites were analyzed. The phylogeny was reconstructed by maximum likelihoodand GTR+I+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green, and operational taxonomic unities (OTUs) from As gene clone libraries in blue. Different background colors highlight three bacterial phyla - *Proteobacteria*, *Firmicutes*, and *Chrysiogenetes*.
doi:10.1371/journal.pone.0095655.g005

for MS-AsIII and MS-AsV 16S rRNA sequences show very similar evolutionary histories where the relationships among *Firmicutes*, *Actinobacteria*, *Proteobacteria*, and *Thermodesulfobacteria* phyla members reflect the current knowledge regarding their evolution [66].

As previously described on the literature, the evolutionary relationships of *arsC* and *arrA* homologs (Figs.4 and 5) support the role of horizontal gene transfer (HGT) on the evolution of arsenate oxidases e.g. [67,68]. The phylogeny reconstructed for *arsC* homologs (Fig. 4) clearly shows two *Ochrobactrum* sequences clustered in different well-supported clades suggesting that these two homologs were acquired by HGT from unrelated donors. Although it is known that due to HGT events *aioA* sequences are not a suitable marker for microbial diversity studies [53], it was not observed on the *aioA* phylogeny here presented (Fig. 6). Albeit *aioA* sequences have been detected in non-proteobacterial lineages [53,69], our findings show two strongly supported clades clustering *alpha-* and *beta-proteobacteria* homologs. Such results probably reflect the bias existing on GenBank databases where most *aioA*sequences available are from proteobacterial lineages.

Overall, evolutionary analyses revealed high genetic similarity between some *arsC* and *aioA* sequences obtained from isolates and clone libraries, suggesting that those isolates may represent environmentally important bacteria acting in As speciation. In addition, some *arsC, aioA, and arrA* sequences were found to be closely related to homologs from uncultured bacteria. Thus, it may be hypothesized that these divergent sequences could represent novel variants of the As-resistance genes or other genes with related function. In addition, our findings show that the diversity of *arrA* genes is wider than earlier described, once none *arrA*-OTUs were affiliated with known reference strains. Therefore, the

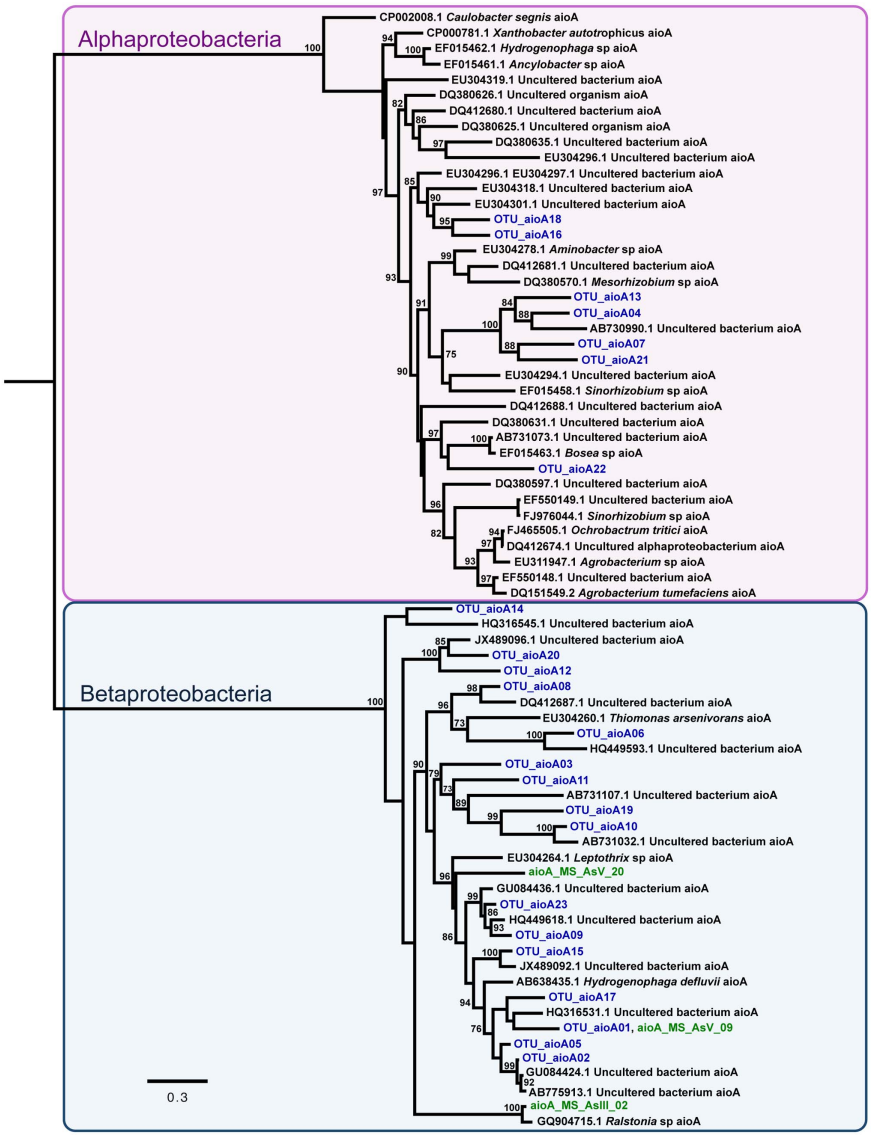

Figure 6. Evolutionary relationships of *aioA* sequences. A total of 72 nucleotide sequences and 543 sites were analyzed. The phylogeny was reconstructed by maximum likelihood and GTR+I+G+F was selected as best fit model. Support values for each node were estimated using the Akaike Likelihood Ratio Test (aLRT). Only support values higher than 70% are shown. Reference sequences retrieved from the non-redundant database of the NCBI are shown in black, bacterial isolates (MS-AsIII and MS-AsV) in green, and operational taxonomic unities (OTUs) from As gene clone libraries in blue. Different background colors highlight two Proteobacteria classes – *beta-* and *alpha-proteobacteria*.
doi:10.1371/journal.pone.0095655.g006

molecular diversity of arrA genes is far from being fully explored deserving further attention.

Altogether, this study is a bioprospection of AsIII-oxidizing and AsV-reducing bacteria and As-transforming genes in sediments impacted by long-term gold mining. Our culture efforts successfully identified a large number of phylogenetically distinct arsenic-resistant bacterial genera and revealed two novel As-transformation genera, *Thermomonas* and *Pannonibacter*. Our heterotrophic arsenite oxidizers and DARB isolates open new opportunities for their use in bioremediation of long-term gold-mining impacted areas. Furthermore, metagenomic analysis of As functional genes revealed a predominance of previously unidentified DARB.

Supporting Information

Figure S1 Map showing the sampling site. Crosshatch, red and yellow areas represent mining, urban, and sampling areas, respectively.
(TIF)

Tables S1 This file includes Table S1, S2 and S3. Table S1. Phylogenetic affiliation of *aioA* OTUs based on blastx protein database. Table S2. Phylogenetic affiliation of *arsC* OTUs based on blastx protein database. Table S3. Phylogenetic affiliation of *arrA* OTUs based on blastx protein database.
(DOCX)

Acknowledgments

We appreciate the technical supportf from Laboratório de Análises Químicas/DEMET/UFMG do Instituto Nacional de Ciência e Tecnologia em Recursos Minerais, Água e Biodiversidade – INCT –ACQUA in the chemical analyses. The authors acknowledge the use of the computing resources of the Center for Excellence in Bioinformatics (CEBio// CPqRR/Fiocruz, Brazil).

Author Contributions

Conceived and designed the experiments: PSC ECS AMAN. Performed the experiments: PSC MPR PLO LBI MLSS FARB. Analyzed the data: PSC LLSS MPR AVC ECS AMAN. Contributed reagents/materials/ analysis tools: FARB AMAN. Wrote the paper: PSC LLSS MPR ECS AMAN.

References

1. Lièvremont D, Bertin PN, Lett MC (2009) Arsenic in contaminated waters: Biogeochemical cycle, microbial metabolism and biotreatment processes. Biochimie 91: 1229–1237.
2. Páez-Espino D, Tamames J, de Lorenzo V, Cánovas D (2009) Microbial responses toenvironmentalarsenic. Biometals 22: 117–130.
3. Neubauer O (1947) Arsenical cancer - a review. Br J Cancer 1: 192–251.
4. Bhattacharya P, Welch AH, Stollenwerk KG, McLaughlin MJ, Bundschuh J, et al. (2007) Arsenic in the environment: Biology and Chemistry. Sci Total Environ 379: 109–20.
5. McClintock TR, Chen Y, Bundschuh J, Oliver JT, Navoni J, et al. (2012) Arsenic exposure in Latin America: biomarkers, risk assessments and related health effects. Sci Total Environ 429: 76–91.
6. Kruger MC, Bertin PN, Heipieper HJ, Arsène-Ploetze F (2013) Bacterial metabolism of environmental arsenic-mechanisms and biotechnological applications. Appl Microbiol Biotechnol 97: 3827–3841.
7. Nordstrom DK (2002) Public health-worldwide occurrences of arsenic in ground water. Science 296: 2143–2145.
8. Tsai SL, Singh S, Chen W (2009) Arsenic metabolism by microbes in nature and the impact on arsenic remediation. Curr Opin Biotechnol 20: 659–67.
9. Anderson CR, Cook GM (2004) Isolation and characterization of arsenate-reducing bacteria from arsenic-contaminated sites in New Zealand. Curr Microbiol 48: 341–347.
10. Chang JS, Yoon IH, Lee JH, Kim KR, An J, et al. (2008) Arsenic detoxification potential of *aox* genes in arsenite oxidizing bacteria isolated from natural and constructed wetlands in the Republic of Korea. Environ Geochem Health 32: 95–105.
11. Drewniak L, Styczek A, Majder-Lopatka M, Sklodowska A (2008) Bacteria, hypertolerant to arsenic in the rocks of an ancient gold mine, and their potential role in dissemination of arsenic pollution. Environ Pollut 156: 1069–74.
12. Cai L, Liu G, Rensing C, Wang G (2009) Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. BMC Microbiol 9: 4.
13. Cavalca L, Zanchi R, Corsini A, Colombo M, Romagnoli C, et al. (2010) Arsenic-resistant bacteria associated with roots of the wild *Cirsiumarvense* (L.) plant from an arsenic polluted soil, and screening of potential plant growth-promoting characteristics. Syst Appl Microbiol 33: 154–64.
14. Silver S, Phung LT (2005) Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. Appl Environ Microbiol 71: 599–608.
15. Kaur S, Kamli MR, Ali A (2009) Diversity of arsenate reductase genes (*arsC* genes) from arsenic-resistant environmental isolates of *E. coli*. CurrMicrobiol 59: 288–94.
16. Liao VHC, Chu YJ, Su YC, Hsiao SY, Wei CC, et al. (2011) Arsenite-oxidizing and arsenate-reducing bacteria associated with arsenic-rich groundwater in Taiwan. J Contam Hydrol 123: 20–9.
17. Malasarn D, Saltikov CW, Campbell KM, Santini JM, Hering JG, et al. (2004) *arrA* is a reliable marker for As(V) respiration. Science 36: 455.
18. Zargar K, Conrad A, Bernick DL, Lowe TM, Stolc V, et al. (2012) ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. Environ Microbiol 14: 1635–45.
19. Muller D, Lièvremont D, Simeonova DD, Hubert JC, Lett MC (2003) Arsenite oxidase *aox* genes from a metal-resistant *beta-proteobacterium*. J Bacteriol 185: 135–41.
20. Hamamura N, Macur RE, Korf S, Ackerman G, Taylor WP, et al. (2009) Linking microbial oxidation of arsenic with detection and phylogenetic analysis of arsenite oxidase genes in diverse geothermal environments. Environ Microbiol 11: 421–31.
21. Stolz JF, Basu P, Oremland RS (2010) Microbial arsenic metabolism: new twists on an old poison. Microbe 5: 53–59.
22. Slyemi D, Bonnefoy V (2012) How prokaryotes deal with arsenic. Environ Microbiol Reports 4: 571–586.
23. Lett MC, Muller D, Lièvremont D, Silver S, Santini J (2012) Unified nomenclature for genes involved in prokaryotic aerobic arsenite oxidation. J Bacteriol 194: 207–8.
24. Oremland RS, Saltikov CW, Wolfe-Simon F, Stolz JF (2009) Arsenic in the evolution of earth and extraterrestrial ecosystems. Geomicrobiol J 26: 522–536.
25. Sun W, Sierra-Alvarez R, Milner L, Field JA (2010) Anaerobic oxidation of arsenite linked to chlorate reduction. Appl Environ Microbiol 76: 6804–6811.
26. Kulp TR, Hoeft SE, Asao M, Madigan MT, Hollibaugh JT, et al. (2008) Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. Science 321: 967–970.
27. Borba RP, Figueiredo BR, Rawlins BG, Matchullat J (2000) Arsenic in water and sediment in the Iron Quadrangle, Minas Gerais state, Brasil. Revista Brasileira de Geociências 30: 554–557.
28. Instituto Mineiro de Gestão das Águas (IGAM) (2004) Camargos LMM. Plano diretor de recursos hídricos da bacia hidrográfica do rio das Velhas: resumo executivo - Belo Horizonte, MG. Instituto Mineiro de Gestão das Águas, Comitê da Bacia Hidrográfica do Rio das Velhas.
29. Salomons W, de Rooij NM, Kerdijk H, Bril J (1987) Sediments as a source for contaminants? Hydrobiologia149: 13–30.
30. Rasmussen H, Jørgensen BB (1992) Microelectrode studies of seasonal oxygen uptake in a coastal sediment: role of molecular diffusion. Marine Ecology Progress Series 81: 289–303.
31. Marchand C, Lallier-Verges E, Allenbach M (2011) Redox conditions and heavy metals distribution in mangrove forests receiving shrimp farm effluents (Teremba bay, New Caledonia). J Soils Sediments 11: 529–541.
32. Mackereth FJH, Heron J, Talling JF (1978) Water analysis: some revised methods for limnologists. Freshwater Biological Association Scientific Publication, Wareham.
33. Golterman HL, Clymo RS, Ohnstad MAM (1978) Methods for chemical analysis of fresh waters. Blackwell Scientific Publications, Philadelphia, PA.
34. Dramsi S, Biswas I, Maguin E, Braun L, Mastroeni P, et al. (1995) Entry of *Listeriamonocytogenes* into hepatocytes requires expression of *inIB*, a surface protein of the internalin multigene family. Mol Microbiol 16: 251–261.
35. Freitas DB, Lima-Bittencourt CI, Reis MP, Costa PS, Assis PS, et al. (2008) Molecular characterization of early colonizer bacteria from wastes in a steel plant. Lett Appl Microbiol 47: 241–249.
36. Lane DJ (1991) 16S/23S rRNA sequencing. John Wiley and Sons, New York.
37. Sun Y, Polishchuk EA, Radoja U, Cullen WR (2004) Identification and quantification of *arsC* genes in environmental samples by using real-time PCR. J Microbiol Methods 58: 335–49.
38. Sarkar A, Kazy SK, Sar P (2013) Characterization of arsenic resistant bacteria from arsenic rich groundwater of West Bengal, India. Ecotoxicology 2: 363–376.
39. Schloss PD, Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol 71: 1501–1506.
40. Good IJ (1953) The population frequencies of species and the estimation of population parameters. Biometrika 40: 237–262.
41. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30: 772–780.
42. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench Bioinformatics 25: 1189–1191.
43. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol. 59: 307–21.
44. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9: 772.
45. Salmassi TM, Venkateswaren K, Satomi M, Nealson KH, Newman DK, et al. (2002) Oxidation of arsenite by *Agrobacteriumalbertimagni*, AOL15, spnov., isolated from Hot Creek, California. Geomicrobiol J 19: 53–66.
46. Conselho Nacional do Meio Ambiente (CONAMA) (2011) Resolução N° 430, de 13 de maio de 201. URL: http://www.mma.gov.br/port/conama/estr.cfm. Accessed 20 June 2011.
47. Canadian Council of Ministers of the Environment (CCME) Canadian Environmental Quality Guidelines. URL: http://www.ccme.ca/. Accessed 02 April 2011.
48. Salas HJ, Martino P (1991) A simplified phosphorus trophic state model for warm-water tropical lakes. Water Res 25: 341–350.
49. Santini JM, Sly LI, Schnagl RD, Macy JM (2000) A new chemolithoautotrophicarsenite-oxidizing bacterium isolated from a gold mine: phylogenetic, physiological, and preliminary biochemical studies. Appl Environ Microbiol 66: 92–97.
50. Santini JM, Sly LI, Wen A, Comrie D, Wulf-Durand P, et al. (2002) New arsenite-oxidizing bacteria isolated from australian gold mining environments-phylogenetic relationships. Geomicrobiol J 19: 67–76.
51. Santini JM, vanden Hoven RN (2004) Molybdenum-containing arsenite oxidase of the chemolithoautotrophicarsenite oxidizer NT-26. J Bacteriol 186: 1614–1619.

52. Oliveira A, Pampulha ME, Neto MM, Almeida AC (2009) Enumeration and characterization of arsenic-tolerant diazotrophic bacteria in a long-term heavy-metal-contaminated soil.Water Air Soil Pollut 200: 237–243.
53. Heinrich-Salmeron A, Cordi A, Brochier-Armanet C, Halter D, Pagnout C, et al. (2011) Unsuspected Diversity of Arsenite-Oxidizing Bacteria as Revealed by Widespread Distribution of the aoxB Gene in Prokaryotes. Appl Environ Microbiol 77: 4685–92.
54. Yamamura S, Ike M, Fujita M (2003) Dissimilatory arsenate reduction by a facultative anaerobe, Bacillus sp. strain SF-1. J BiosciBioeng 96: 454–60.
55. Chang JS, Kim IS (2010) Arsenite oxidation by *Bacillus* sp. strain SeaH-As22w isolated from coastal seawater in Yeosu Bay. Environ Eng Res 15: 15–21.
56. Reis MP, Barbosa FA, Chartone-Souza E, Nascimento AMA (2013) The prokaryotic community of a historically mining-impacted tropical stream sediment is as diverse as that from a pristine stream sediment. Extremophiles 17: 301–309.
57. Bandyopadhyay S, Schumann P, Das SK (2013) *Pannonibacter indica* sp. nov., a highly arsenate-tolerant bacterium isolated from a hot spring in India. Arch Microbiol 195: 1–8.
58. Dowdle PR, Laverman AM, Oremland RS (1996) Bacterial dissimilatory reduction of arsenic(V) to arsenic(III) in anoxic sediments. Appl Environ Microbiol 62: 1664–9.
59. Chang YC, Nawata A, Jung K, Kikuchi S (2012) Isolation and characterization of an arsenate-reducing bacterium and its application for arsenic extraction from contaminated soil. J Ind Microbiol Biotechnol 39: 37–44.
60. Achour AR, Bauda P, Billard P (2007) Diversity of arsenite transporter genes from arsenic-resistant soil bacteria. Res Microbiol 158: 128–37.
61. Sri LSM, Prashant S, Bramha CPV, Nageswara RS, Balaravi P, et al. (2012) Molecular identification of arsenic-resistant estuarine bacteria and characterization of their *ars* genotype. Ecotoxicology21: 202–12.
62. Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, et al. (2008) The pangenome structure of *Escherichiacoli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates,JBacteriol 190: 6881–93.
63. Gootz TD, Lescoe MK, Dib-Hajj F, Dougherty BA, He W, et al. (2009) Genetic organization of transposase regions surrounding *blaKPC* carbapenemase genes on plasmids from *Klebsiella* strains isolated in a New York City hospital. Antimicrob Agents Chemother 53: 1998–2004.
64. Inskeep WP, Macur RE, Hamamura N, Warelow TP, Ward SA, et al. (2007) Detection, diversity and expression of aerobic bacterial arsenite oxidase genes. Environ Microbiol 9: 934–43.
65. Quéméneur M, Heinrich-Salmeron A, Muller D, Lièvremont D, Jauzein M, et al. (2008) Diversity surveys and evolutionary relationships of *aoxB*genes in aerobic arsenite-oxidizing bacteria. Appl Environ Microbiol 74: 4567–73.
66. Olsen GJ, Woese CR, Overbeek R (1994) The winds of (Evolutionary) change: breathing new life into microbiology. J Bacteriol 176: 1–6.
67. Jackson CR, Dugas SL (2003) Phylogenetic analysis of bacterial and archaeal *arsC* gene sequences suggests an ancient, common origin for arsenate reductase BMC Evol Biol.3: 18.
68. Duval S, Ducluzeau AL, Nitschke W, Schoepp-Cothenet B (2008) Enzyme phylogenies as markers for the oxidation state of the environment: the case of respiratory arsenate reductase and related enzymes. BMC Evol Biol 8: 206.
69. Andres J, Arsène-Ploetze F, Barbe V, Brochier-Armanet C, Cleiss-Arnold J, et al. (2013) Life in an arsenic-containing gold mine: genome and physiology of the autotrophic arsenite-oxidizing bacterium rhizobium sp. NT-26. Genome Biol Evol 5: 934–953.

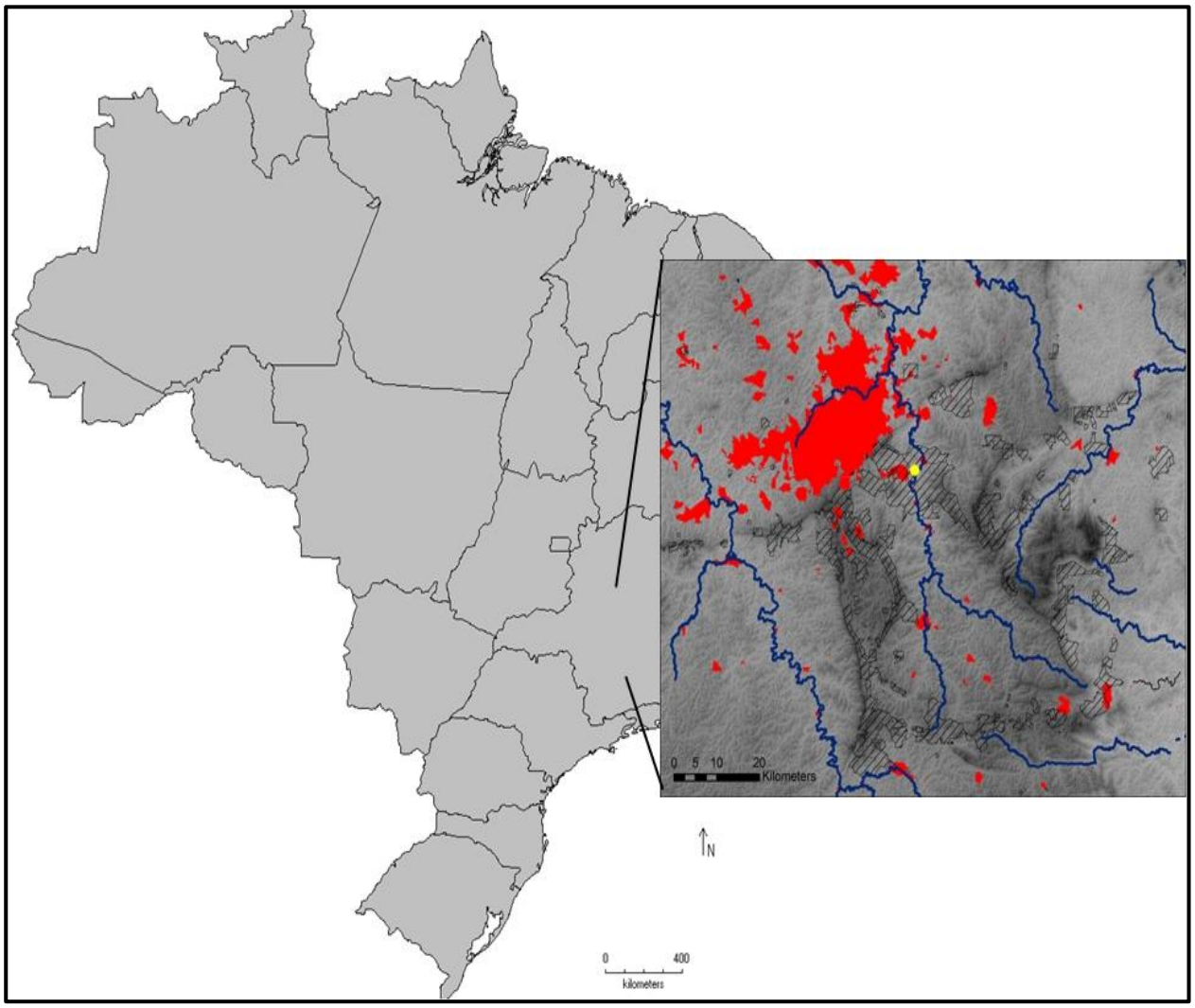

Figure S1 - Map showing the sampling site. Crosshatch, red and yellow areas represent mining, urban, and sampling areas, respectively.

Table S1 – Phylogenetic affiliation of *aioA* OTUs based on blastx protein database

OTU	Blastx Best Results	Organism (access number)	Id (%)	Source
AIOA1	arsenite oxidase large subunit	*Hydrogenophaga defluvii* (BAK39656.1)	89	Groundwater with arsenic
AIOA2	arsenite oxidase large subunit	*Hydrogenophaga defluvii* (BAK39656.1)	94	Groundwater with arsenic
AIOA3	arsenite oxidase large subunit	*Pseudomonas* sp. (AEL22192.1)	78	As-contaminated sediment from Sainte Marie aux Mines
AIOA4	arsenite oxidase large subunit	uncultured bacterium (ADE33043.1)	86	arsenic contaminated water
AIOA5	arsenite oxidase large subunit	*Hydrogenophaga defluvii* (BAK39656.1)	94	Groundwater with arsenic
AIOA6	arsenite oxidase large subunit	uncultured bacterium (AEL22131.1)	89	As-contaminated sediment from Sainte Marie aux Mines
AIOA7	arsenite oxidase large subunit	uncultured bacterium (AEC32846.1)	92	microbial mats from gold mine
AIOA8	arsenite oxidase large subunit	*Thiomonas arsenivorans* (ABY19316.1)	90	As-contaminated soil from a disused gold mine
AIOA9	arsenite oxidase large subunit	uncultured bacterium (AEL22137.1)	93	As-contaminated sediment from Sainte Marie aux Mines
AIOA10	arsenite oxidase large subunit	uncultured bacterium (BAM24613.1)	96	aquatic sediment
AIOA11	arsenite oxidase large subunit	*Thiomonas* sp. (ACA28597.1)	71	arsenic-rich acidic waters draining the Carnoules mine tailings
AIOA 12	arsenite oxidase alpha subunit	*Burkholderia* sp. (ADF47196.1)	77	soil enrichment culture with arsenic
AIOA 13	arsenite oxidase large subunit	uncultured bacterium (ADE33043.1)	86	arsenic contaminated water
AIOA 14	arsenite oxidase alpha subunit	*Alcaligenes* sp. (ADF47197.1)	76	soil enrichment culture with arsenic
AIOA 15	arsenite oxidase	uncultured bacterium (ABE02193.1)	96	arsenic contaminated sediment
AIOA 16	arsenite oxidase large subunit	*Aminobacter* sp. (ABY19334.1)	77	As-contaminated mining site
AIOA 17	arsenite oxidase large subunit	*Hydrogenophaga defluvii* (BAK39656.1)	91	Groundwater with arsenic
AIOA 18	arsenite oxidase large subunit	*Aminobacter* sp. (ABY19334.1)	80	As-contaminated mining site
AIOA 19	arsenite oxidase large subunit	uncultured bacterium (ABY19352.1)	72	As-contaminated soil
AIOA 20	arsenite oxidase	uncultured bacterium (ABE02209.1)	88	arsenic contaminated sediment
AIOA 21	arsenite oxidase large subunit	uncultured bacterium (AEC32846.1)	84	microbial mats from gold mine
AIOA 22	arsenite oxidase alpha subunit	*Aminobacter* sp. (ABY19334.1)	90	As-contaminated mining site
AIOA 23	arsenite oxidase large subunit	uncultured bacterium (ADE33054.1)	96	arsenic contaminated water

Table S2 - Phylogenetic affiliation of *arsC* OTUs based on blastx protein database

OTU	blastx best results	Organism (access number)	Id (%)	Source
ARSC1	arsenate reductase	*Klebsiella pneumoniae* (YP_002287000.1)	100	hospital
ARSC2	arsenate reductase	*Cronobacter turicensis* (YP_003212785.1)	100	neonatal infections
ARSC3	arsenate reductase	*Vibrio* sp. (ABO29820.1)	92	marine environments
ARSC4	arsenate reductase	*Mesorhizobium alhagi* (ZP_09294313.1)	82	Symbiotic root nodules
ARSC5	arsenate reductase	*Escherichia coli* (YP_001464969.1)	100	-
ARSC6	arsenate reductase	*Methylocystis* sp. (ZP_08072529.1)	76	-
ARSC7	arsenate reductase	*Vibrio* sp. (ABO29820.1)	100	marine environments
ARSC8	arsenate reductase	*Cronobacter turicensis* (YP_003212785.1)	100	neonatal infections
ARSC9	arsenate reductase	*Agrobacterium* sp. (ZP_08528114.1)	95	industrial strain
ARSC10	arsenate reductase	*Novosphingobium* sp. (YP_004534179.1)	92	marine environments
ARSC11	arsenate reductase	*Escherichia hermannii* (ZP_09808969.1)	97	-
ARSC12	arsenate reductase	*Klebsiella oxytoca* (EHS95433.1)	100	-
ARSC13	arsenate reductase	*Oligotropha carboxidovorans* (YP_002287046.1)	83	wastewater
ARSC14	arsenate reductase	*Klebsiella oxytoca* (EHS95433.1)	97	-
ARSC15	arsenate reductase	*Oligotropha carboxidovorans* (YP_002287046.1)	83	wastewater

Table S3 – Phylogenetic affiliation of *arrA* OTUs based on blastx protein database

OTU	blastx best results	Organism (access number)	Id (%)	Source
ARRA1	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	83	rock biofilms from an ancient gold mine
ARRA2	arsenate respiratory reductase	uncultured bacterium (CBW77477.1)	78	rock biofilms from an ancient gold mine
ARRA6	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	83	rock biofilms from an ancient gold mine
ARRA8	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	79	rock biofilms from an ancient gold mine
ARRA9	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	98	rock biofilms from an ancient gold mine
ARRA10	arsenate respiratory reductase	uncultured bacterium (CBW77457.1)	65	rock biofilms from an ancient gold mine
ARRA12	arsenate respiratory reductase	uncultured bacterium (CBW77477.1)	75	rock biofilms from an ancient gold mine
ARRA14	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	78	rock biofilms from an ancient gold mine
ARRA15	arsenate respiratory reductase	uncultured bacterium (CBW77477.1)	74	rock biofilms from an ancient gold mine
ARRA16	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	81	rock biofilms from an ancient gold mine
ARRA19	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	83	rock biofilms from an ancient gold mine
ARRA20	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	78	rock biofilms from an ancient gold mine
ARRA22	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	78	rock biofilms from an ancient gold mine
ARRA25	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	83	rock biofilms from an ancient gold mine
ARRA26	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	81	rock biofilms from an ancient gold mine
ARRA29	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	83	rock biofilms from an ancient gold mine
ARRA30	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	84	rock biofilms from an ancient gold mine
ARRA35	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	82	rock biofilms from an ancient gold mine
ARRA36	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	78	rock biofilms from an ancient gold mine
ARRA37	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	84	rock biofilms from an ancient gold mine
ARRA39	arsenate respiratory reductase	uncultured bacterium (AEX97846.1)	64	Cache Valley LandFill sediments
ARRA41	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	72	rock biofilms from an ancient gold mine
ARRA42	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	69	rock biofilms from an ancient gold mine
ARRA43	arsenate respiratory reductase	uncultured bacterium (CBW77469.1)	81	rock biofilms from an ancient gold mine
ARRA46	arsenate respiratory reductase	uncultured bacterium (CBW77477.1)	72	rock biofilms from an ancient gold mine

4.2 - Capítulo 2

Metagenome of a microbial community inhabiting a metal-rich tropical stream sediment

Metagenome of a microbial community inhabiting a metal-rich tropical stream sediment

Patrícia S. Costa[1], Mariana P. Reis[1], Marcelo P. Ávila[1], Laura R. Leite[2], Flávio M.G. de Araújo[2], Anna C.M. Salim[2], Guilherme Oliveira[2], Francisco Barbosa[1], Edmar Chartone-Souza[1], Andréa M.A. Nascimento[1*].

[1]Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais; Av. Antônio Carlos 6627 Belo Horizonte, Minas Gerais, Brazil, CEP: 31270-901.

[2]Centro de Pesquisas René Rachou Fundação Oswaldo Cruz; Av. Augusto de Lima 1715 Belo Horizonte, Minas Gerais, Brazil, CEP: 30190-002.

*Corresponding author. Phone: 55-31-3409-2588, Fax: 55-31-3409-2567.

E-mail address: amaral@ufmg.br

Running title: Metagenome of a metal-rich sediment

Keywords: metagenome, prokaryote, metal, sediment, metabolic fingerprint, nitrogen cycle genes

Abstract

Here, we describe the metagenome and functional composition of a microbial community in a historically metal-contaminated tropical freshwater stream sediment. The sediment was collected from the Mina Stream located in the Iron Quadrangle (Brazil), one of the world's largest mining regions. Environmental DNA was extracted and was sequenced using SOLiD technology, and a total of 7.9 Gbp was produced. A taxonomic profile that was obtained by comparison to the Greengenes database revealed a complex microbial community with a dominance of *Proteobacteria* and *Parvarcheota*. Contigs were recruited by bacterial and archaeal genomes, especially *Candidatus Nitrospira defluvii* and *Nitrosopumilus maritimus*, and their presence implicated them in the process of N cycling in the Mina Stream sediment (MSS). Functional reconstruction revealed a large, diverse set of genes for ammonium assimilation and ammonification. These processes have been implicated in the maintenance of the N cycle and the health of the sediment. SEED subsystems functional annotation unveiled a high degree of diversity of metal resistance genes, suggesting that the prokaryotic community is adapted to metal contamination. Furthermore, a high metabolic diversity was detected in the MSS, suggesting that the historical arsenic contamination is no longer affecting the prokaryotic community. These results expand the current knowledge of the microbial taxonomic and functional composition of tropical metal-contaminated freshwater sediments.

Introduction

Prokaryotic species exhibit broad distribution, having been researched across a wide range of natural environments such as soil, marine and freshwater, as well as in plants, animals and humans. Many of these species have been revealed to be important for the health and/or ecological balance of various environments. Indeed, a link between the set of microbial species and the host or environment-associated biological processes and health has been extensively reported [1, 2]. Because of their essential roles in life and in ecosystem functioning, ambitious multidisciplinary efforts across the globe are ongoing to characterize microbial communities [3].

Sediment has been recognized as a special realm in aquatic ecosystems because its species richness is higher than that of the water community and is comparable to soil microbial diversity [4, 5]. In mining-contaminated regions, sediments of water bodies play an important role in the transport and storing of contaminants. Indeed, sediment characteristics determine the ecological balance and biodiversity of the aquatic ecosystem [6].

There is a consensus in the literature that metal-contaminated freshwater sediment exhibits an extremely complex and well-adapted community [7-9]. These studies revealed that *Proteobacteria*, especially *Beta-proteobacteria*, and *Bacteroidetes* are the main contributors to the composition of these environments. It should be noted that sediment communities play an important role in biogeochemical cycling and are involved in the transformation of nutrients such as N and C [9].

Although previous studies of microbial communities in metal-contaminated freshwater sediment have been performed [5, 8, 10, 11], none of them assessed the microbial community of a metal-contaminated tropical sediment through taxonomic and functional diversity evaluation. Moreover, all of the studies, except Reis *et al.* [8], focused their analysis on sediments of temperate streams. However, due to the restricted power of the methodology employed by Reis *et al.* [8], these authors did not cover all of the taxonomic richness present in the tropical stream

studied here. Thus, much is still unknown about the functional and taxonomic microbial diversity of tropical metal-contaminated streams. Considering that microorganisms play an essential role in environmental biogeochemical cycling, and may influence the speciation and bioavailability of metals, it is relevant to obtain a more comprehensive knowledge of the taxonomic and functional diversity of the prokaryotic community in metal-contaminated freshwater sediments.

One powerful strategy to assess both the functional and taxonomic microbial diversity is a metagenomic approach. Indeed, over the last 20 years, new sequencing technologies, together with metagenomic and computational tools, have transformed microbial ecology research. Metagenomics provides insight into the interactions of microbial communities with the environment and offers an extraordinary opportunity to comprehensively examine the ecosystem's response to environmental changes [12]. However, metagenomic surveys that thoroughly assess the microbial diversity in freshwater sediments with extreme geochemical conditions involving high concentrations of As, Fe, and Mn are still lacking.

In this study, we applied a shotgun metagenomic approach and a metabolic analysis to examine the taxonomic and functional composition of the prokaryotic community of a historically metal-contaminated tropical stream sediment. The stream studied herein, the Mina Stream, is located in the Iron Quadrangle (IQ, Brazil), one of the world's largest mining regions, which has been undergone to mining activities since the late 17th century. Accordingly, the IQ presents a historical metal contamination of waters and sediments from streams and rivers, including the Mina Stream [8, 13-16]. We also performed comparative metagenomic analysis between our metagenome and a rich arsenic well metagenomic dataset from Bangladesh [17].

Material and Methods

Ethics Statement

For sampling in the Mina Stream, no specific permit was required for the described field study. The study location is not privately owned or protected in any way, and we confirmed that the field study did not involve endangered or protected species.

Study area

The Mina stream (19º58'46.80"S and 43º49'17.07"W) is located in one of the world's largest mining regions and is extremely rich in iron and gold ores (Iron Quadrangle, Minas Gerais state, Brazil). Collections of Mina stream sediment have been previously described by our group [11]. This stream was chosen because it has suffered stress by metal pollution exceeding the maximum allowable concentrations established by Brazilian environmental regulations, such as Cu 387.7 mg kg^1, Zn 180.9 mg kg^1 and As 297.1 mg kg^1, which were presented in an earlier study [11].

The sediment sample in this study was taken from the upper part (oxic zone) during the dry season and was named according to the location from which it was retrieved, i.e., Mina Stream sediment (MSS). For metabolic analysis, the anaerobic environment of the sediment sample was maintained by substituting the O_2 for CO_2 using a CO_2 pump, and the tube was hermetically closed. Two hours after collection, the sediment sample was introduced into an anaerobic chamber where subsequent experiments were performed.

Microbial metabolic diversity

The capability of aerobic and anaerobic sediment microbial communities to utilize different carbon sources was assessed using Biolog Ecoplate^{TM} (Biolog.Inc, Hayward, CA, USA). This system contained 31 carbon sources, in triplicate, divided into amines, amino acids, carbohydrates, carboxylic acids, and polymers, among others. In addition to the specific carbon

source, each well contained tetrazolium violet redox dye as a color indicator for the utilization of the carbon sources by the microorganisms [18] (Table S1). Sediment sample was filtered (10 g wet weight; pore size 0.45μm) and diluted in sterile saline. Then, 120 μL from the 10^{-2} dilutions was inoculated into each well and subsequently incubated aerobically and anaerobically in the dark at 28 °C. Color development was measured at OD_{590} every 24 h for 4 d using an ELISA plate reader (BIO-RAD Model 3550 Microplate Reader). For the anaerobic BIOLOG assay, four plates were used, one for each day of reading. This procedure was performed by taking into account the loss of anaerobic conditions when the plate was withdrawn from the anaerobic chamber. For aerobic conditions, one plate was used. The detected value of the absorbance for the blank (water) reading was subtracted from all wells.

Ecoplate data analysis

The data generated by 96 h readings were statistically analyzed. Because raw OD_{590} values were corrected, the microbial activity for each microplate was expressed as the average well-color development (AWCD) and was calculated as follows: AWCD = ΣODi/31 where ODi is the optical density value for each well. The richness (number of carbon substrates consumed) and the Shannon-Weaver index were calculated using a cutoff line of OD = 0.25 for a positive microbial response [19]. The Shannon-Weaver index was calculated as follows: H'=-Σpi (ln pi), where pi is the ratio between the microbial activity of each substrate (ODi) and the sum of microbial activities of all substrates (ΣODi). The Evenness index was calculated with the formula E = H'/ ln R, where H' is the value of the Shannon index, and R is the richness of substrates.

DNA extraction and shotgun metagenomic sequencing

Total DNA was extracted from the sediment sample (10 g wet weight) using the PowerSoil DNA Extraction kit (MoBio Laboratories, USA) according to the manufacturer's instructions. Quantification and quality of total DNA were determined using the Agilent 2100 Bioanalyzer equipment according to the manufacturer's instructions.

Sediment sample was subjected to shotgun sequencing using the high-throughput sequencer Applied Biosystems SOLiD™ v.4 following the manufacturer's protocol. Briefly, 10 μg of total DNA was randomly fragmented using the Covaris™ S2 System. A DNA fragment library from 200 to 250 bp long was constructed for sequencing. Then, emulsion PCR was performed to clonally amplify fragments on sequencing beads, followed by enrichment and preparation for deposition in plate for sequencing according to the manufacturer's instructions (http://tools.lifetechnologies.com/content/sfs/manuals/SOLiD4_Library_Preparation_man.pdf). After sequencing, 50 bp reads were generated for further analysis.

16S rRNA gene amplification and sequencing

Amplification of the V3-V4 hypervariable region of the 16S rRNA gene was performed using region-specific bacterial/archaeal primers S-D-Bact-0341-b-S-17 forward 5'-CCTACGGGNGGCWGCAG-3' and S-D-Bact-0785-a-A-21 reverse 5'-GACTACHVGGGTATCTAATCC-3' [20], with Illumina adapters added. Barcoded amplicons were generated using KAPA HiFi HotStart ReadyMix (KAPA, Woburn, MA, USA) and were purified using AMPure XP beads (Agencourt Bioscience, Beverley, MA, USA). Sequencing was performed using the MiSeq platform (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's instructions (http://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf).

Bioinformatics analysis

V3-V4 region from 16S rRNA gene

16S rRNA microbiota primary data analysis was performed with PRINSEQ (stand alone lite version, http://prinseq.sourceforge.net/), where quality-based trimming was performed. Reads with N's or an overall mean Q-score of < 25 were discarded. The resulting fasta file was also screened for ambiguous bases and homopolymers using MOTHUR v.1.33.0 (http://www.mothur.org). Furthermore, chimeras were detected using the UCHIME algorithm (http://drive5.com/uchime).

Operational taxonomic units (OTUs) and taxonomic classification were determined using the MOTHUR pipeline [20, 21] and the Greengenes reference database (http://greengenes.secondgenome.com/downloads/database/13_5, from May 2013) to obtain the microbial composition of the MSS microbiota. OTUs were determined using similarity levels between sequences of at least 97% for classifying a microorganism at the species level, as proposed by Drancourt *et al.* [22]. Good's coverage [23] was calculated for OTUs with an evolutionary distance of 0.03. Rarefaction curves were calculated for OTUs with an evolutionary distance of 0.03, 0.05 and 0.10. The nucleotide sequences were submitted to Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra/) with the accession number of SRR1573431.

Shotgun metagenome data

Metagenomic primary data analysis was performed with SOLiD ™ Accuracy Enhancement Tool (SAET) software (http://solidsoftwaretools.com/gf), a spectral alignment algorithm that screens for errors inherent to the sequencing platform and the encodeFasta.py program (http://gnome.googlecode.com/svn/trunk/pyGenotypeLearning/src/pytools/encodeFasta.py), that converts the sequences represented in color space to letter space format. Then, the assembly of the

metagenome data was performed to generate contigs using the Metavelvet software [24] with parameters according to the recommendations of the authors (kmer 27, -exp_covauto) [24].

A Fasta file with contig sequences was deposited into the Metagenomics RAST Server (MG-RAST v3.3) [25]. Prior to annotation, MG-RAST provides a quality control of sequences that consists of artificially removing duplicate sequences and screening based on quality and size of sequences. Functional analysis was performed using the SEED subsystem and KEEG available on MG-RAST with the following cutoff parameters: $1x10^{-5}$ e-value and 60% of identity percentage [26]. The data from this study are available via MG-RAST with the ID 4519449.3.

A recruitment plot was used to identify abundant species genomes in the MSS metagenome. In this representation, MSS metagenome contigs were compared to individual bacterial genomes. Fragment recruitment of the MSS contigs was performed using BLASTN against bacterial and archaeal complete genomes. Data were plotted using R (http://cran.r-project.org), and the criteria for counting a hit were a minimum identity of 90%, e-value cutoff 0.001 and minimum alignment of 50 bp.

Comparative metagenomic analysis

Comparative metagenomic analysis was performed using the Statistical Analysis of Metagenomic Profiles (STAMP) program [27] to determine statistically significant functional composition differences in any two metagenomes using two-sided Fisher exact tests [27]. The most important metabolic categories were selected by using a p-value >0.05. To accomplish that, the MG-RAST functional matches at all levels were compared using the SEED database (http://www.theseed.org). The statistical comparison was conducted with the data from a rich arsenic well metagenome (4461675.3) [17] due to similarity between the two environments, i.e., the high As contamination.

Real-time PCR (qPCR)

Quantitative real-time PCR was performed to estimate the absolute number of copies of bacterial and archaeal 16S rRNA genes in the MSS. To accomplish this outcome, total DNA sample was added to a 20 μl reaction containing a SYBR Green master mix and the bacterial and archeal primer set: 338F (5'-TACGGGAGGCAGCAG-3') and 344F (5'-ACGGGGCGCAGCAGGCGCGA-3'), respectively [28] and 518R (5'-ATTACCGCGGCTGCTGG-3') for both [29]. Standard curves were generated from the 16S rRNA gene amplicons obtained using conventional PCR from *Halococcus morrhuaea* ATCC 17082 and *Escherichia coli* ATCC 25922 as previously described by Cardinali-Rezende *et al*. [30]. The procedure was performed using the ABIPRISM 7900HT sequence detection system (Applied Biosystems, Foster City, CA). The conditions used to amplify the 16S rRNA gene from bacteria and archaea were according to Cardinali-Rezende *et al*. [30].

Results

Taxonomic composition of the prokaryotic community

The MSS microbiota resulted in 273,710 high-quality reads with an average read length of 450 bp. Of a total of 31,656 OTUs, 678 OTUs were not classified within the Bacteria and Archaea domains. Thus, a total of 30,978 OTUs remained for downstream analysis. Of these OTUs, 22,184 were singletons and 2,242 were doubletons composed of only a few reads (27,077). Bacteria were by far the most abundant prokaryotic domain, constituting 98.2% (30,738 OTUs), whereas archaeal reads showed a relative paucity (1.8%, 240 OTUs). The Good's coverage value (89%) and rarefaction curve (Fig. S1) obtained with an evolutionary distance of 0.03 indicated that most of the prokaryotic diversity was detected in the sample.

Bacterial and archaeal phyla diversity are shown in Figure 1. A total of 30,738 OTUs were assigned to 52 known bacterial phyla. Nevertheless, most OTUs were affiliated with four phyla:

Proteobacteria (45%), *Bacteroidetes* (18%), and an equal proportion (4%) of *Acidobacteria* and OD1. The group "other bacteria" comprised minor bacterial phyla such as *Gemmatimonadetes*, *Cyanobacteria*, *OP3*, *OP11*, *Spirochaetes*, and *TM7*, among others, representing 8% of the OTUs. Furthermore, 2,157 OTUs were considered to be unclassified at the phylum level and, thus, may represent new bacterial taxa (Fig. 1).

All proteobacterial classes were represented, with the *Beta-*, *Gamma-*, and *Deltaproteobacteria* classes being the most abundant (81%). *Bacteroidetes* were identified primarily as members of the *Sphingobacteria* (53%), *Flavobacteria* (17%), *Bacteroidia* (17%) and incertae sedis (13%) classes. The *Acidobacteria* phylum was represented by 19 classes, with Gp6, Gp17, Gp3 and *Holophagae* accounting for 69% of representation.

Only 8,430 OTUs (26.6%) were classified at the genus level. The predominant genera observed were, *Sediminibacterium* (*Bacteroidetes*, 520 OTUs), *Flavobacterium* (*Bacteroidetes*, 392), *Prevotella* (*Bacteroidetes*, 371), *Geobacter* (*Proteobacteria*, 323), *Nitrospira* (*Nitrospirae*, 303), *Haliscomenobacter* (*Bacteroidetes*, 250), *Thermomonas* (*Proteobacteria*, 245), *Thiobacillus* (*Proteobacteria*, 240), *Acinetobacter* (*Proteobacteria*, 222) and *Acidovorax* (*Proteobacteria*, 200).

Recruitment of MSS metagenome contigs by bacterial genomes are illustrated in Figures 2 and S2A-G. The bacterial genome that recruited the majority of the contigs was Candidatus *Nitrospira defluvii* (Fig. 2A and B). Other bacterial species were also reasonably well recruited, such as *Anaeromyxobacter dehalogenans*, *Chitinophaga pinensis* DSM2588, *Geobacter metallireducens*, *Leptothrix chlolodnii*, *Sideroxydans lithotrophicus*, *Thiobacillus denitrificans*, and *Thiomonas* 3As (Fig. S2).

The taxonomic affiliation of the Archaea domain revealed that most of the OTUs belonged to the *Parvarchaeota* phylum (83%) represented by the *Parvarchaea* (83%) and *Micrarchaea* (17%) classes. The *Crenarchaeota* phylum (1%) was also represented by three OTUs related to the Miscellaneous Crenarchaeotal Group (MCG). Although members of the *Thaumarchaeota* phylum were not identified in the MSS microbiota, it was possible to recruit the partial genome of

three *Thaumarchaeota* species: *Nitrosopumilus maritimus* SCM1, an ammonia oxidizing archaea belonging to the *Nitrosopumilaceae* family that was originally isolated from a marine fish tank [31] (Fig. 2C and D); *Cenarchaeum symbiosum*, a psychrophilic archaea species that belongs to *Cenarchaeaceae* family and inhabits a marine sponge; and Candidatus *Nitrososphaera gargensis*, an ammonia oxidizing species from *Nitrososphaeraceae* family (Figs. S2H and I).

Abundance of the Bacteria and Archaea domains

The absolute quantification of bacterial and archaeal communities by qPCR was accomplished and generated R^2 values of 0.99 for both curves and slopes of -3.23 and -3.35, respectively (Supplementary Figures S3A-D). According to qPCR analysis, the bacterial 16S rRNA gene copy number (7.7×10^6 gene copies g^{-1}) was two orders of magnitude higher than the archaeal, with 5.3×10^4 gene copies g^{-1} in the sediment sample (Figs. S4A and B).

Overview of metagenomic data

Random shotgun metagenome sequencing from MSS resulted in 158,882,631 reads (50 bp per read) totaling a ~7.9 Gbp dataset. Assembly of reads by Metavelvet resulted in 378,588 contigs ranging from 60 to 2911 bp. After being trimmed by MG-RAST based on quality, size, and artificial removal of duplicate reads, a total of 350,111 clean contigs were used for further analysis. The contig dataset was used to determine the functional analysis. The MSS metagenome exhibited a wide range of GC content from 15% to 80%. Most of the contigs were grouped and ranged from 40 to 60% GC content, with an average GC content of 45 ± 8%.

SEED and KEEG analyses with MG-RAST

Of the 350,111 contigs analyzed for the functional annotation based on the SEED subsystem classification (MG-RAST), 135,632 contigs (39%) could be assigned to functional categories, i.e., predicted proteins with known functions. Nevertheless, most of the contigs (53%) were related to predicted proteins with unknown function, whereas the remaining contigs (8%) presented no match with the SEED database.

Twenty-eight functional subsystems were identified in the MSS metagenome. Protein metabolism, clustering-based subsystems, miscellaneous, carbohydrates, and RNA metabolism presented the largest number of annotated contigs. Other subsystems were related to mobile elements (phages, transposons, integrons, plasmids, and pathogenicity islands) (4%) and stress response (3%), both of which are involved in the fast response and adaptation of the microbial community to changes in the environment (Fig. 3).

Functional analysis with the KEGG Mapper tool of the MG-RAST allows an integrated view of the environmental global metabolism. Assignment of the MSS contigs revealed that most of the metabolic pathways were detected (Supplementary Figure S6). The metabolic pathways identified in the KEGG database as the most abundant were carbohydrate, amino acids, and energy metabolic pathways, indicating that microbial communities inhabiting the MSS are well adapted to degrade carbon substrates such as soluble carbohydrates or polysaccharides and amino acid and derivatives.

Among the genes detected in the MSS, we focused our SEED and KEGG analyses on metal resistance and nitrogen metabolism, which might have particular importance for this environment.

Nitrogen metabolism analysis

The Mina Stream is a eutrophic water body presenting high nitrogen concentration and of its inorganic forms [11]. Therefore, nitrogen metabolism was analyzed, and revealed the presence of enzymes that play a role in ammonia assimilation (49%), nitrate and nitrite ammonification

(33%), allantoin utilization (7%), nitrogen fixation (5%), nitric oxide synthase (3%), and cyanate hydrolysis (3%). Relevant genes involved in these six processes revealed by SEED and KEGG databases are displayed in Table 1 and Supplementary Figure S5.

Metal resistance analysis

The genes associated with heavy metals were highly diverse, with cobalt-zinc-cadmium (47%) and copper resistance (30%) being the most abundant, followed by the arsenic resistance genes accounting for 6% (Fig. 5 and Table 2). Interestingly, the presence of the *arsC* resistance gene was not detected in the MSS metagenome even though this gene is the most widespread arsenic resistance gene in the environment [32].

Statistical comparison of As-contaminated environment

Statistical comparison of the SEED subsystem resemblances between two or more environments can reveal enriched subsystems for a particular environment. To determine biologically significant differences, the functional subsystems detected in the MSS metagenome were statistically compared with the RAW metagenome, as described by Mailloux *et al*. [17]. SEED subsystem comparison revealed a high degree of similarity between the MSS and RAW metagenomes (Fig. 4). However, some differences were observed with significantly over abundant reads in the MSS, which were assigned to mobile elements, regulation and cell signaling, phosphorus metabolism, virulence and defense subsystems, among others. By contrast, the RAW metagenome identified more reads in the amino acid and derivative, clustering-based, carbohydrates, and subsystems related to cell maintenance (Fig. 4). The two metagenomes, MSS and RAW, statistically differed in the enrichment of contigs related to respiratory arsenate reductase (ArrA and ArrB proteins) and multicopper oxidase, which were more frequent in the MSS. By contrast, the RAW metagenome overrepresented arsenate reductase (ArsC) and copper homeostasis (CutE) proteins in the dataset (Fig. 5).

Metabolic diversity and community-level physiological profiles (CLPP)

The metabolic profile of the microbial community of the MSS was assessed using Biolog Ecoplate™ (Biolog, Inc.). Substrate utilization patterns from microbial communities are shown in Table S1. The highest metabolic diversity was observed under anaerobic conditions (30 carbon sources consumed), whereas the community under aerobic conditions consumed 26 carbon sources. 2-hydroxy benzoic acid was the only carbon source not consumed by either microbial community. The substrates α-ketobutyric acid, L-threonine, glycogen, and α-D-lactose were not consumed by the aerobic microbial community.

AWCD reflects the carbon source utilization ability of the microbial community over time. AWCD analysis showed that the microbial community under anaerobic conditions reached the maximum carbon source utilization at just 72 h, after which time the activity reached a plateau (as demonstrated by the maximum color development). By contrast, the microbial community under aerobic conditions did not reach the maximum color development, showing slower growth and consumption of the carbon source (Fig. 6).

The Shannon and Simpson diversity indices of the microbial community metabolic profile were calculated, revealing moderate diversity in both communities (Table S1). Although the anaerobic community presented greater diversity, the differences were not statistically significant ($P \leq 0.05$) (Table S1). In addition, the Simpson's index of microbial community response showed that a few dominant microbial species were responsible for the metabolic profile of both communities.

Discussion

The microbial community plays an important role in the freshwater environment, especially in stream ecosystems where they are responsible for most of the organic matter decomposition [33]. The dataset presented in this study is the first to taxonomically and functionally characterize the microbial community of a metal-contaminated sediment from a tropical freshwater stream using a combination of approaches such as metabolic fingerprinting, qPCR, and shotgun metagenomic sequencing.

Taxonomic analyses revealed that a highly complex bacterial community was present in the MSS. Taxonomic data indicated *Proteobacteria* (especially *Beta-proteobacteria*) was the most abundant phylum followed by *Bacteroidetes*. A previous investigation on the prokaryotic diversity in the MSS also showed the predominance of the *Proteobacteria*, but with its classes presenting different tendencies, and *Bacteroidetes* phyla [8]. However, the present study revealed that the bacterial and archaeal 16S rRNA gene copy number was lower in the dry season, in contrast to the increase detected in the rainy season by Reis *et al.* [8]. The observed increase, up to 10 times that of metal concentrations (mainly Zn and As), in the dry season may have affected the cell abundance of the microbial communities present in the MSS, a finding that was reflected in the abundance of the 16S rRNA gene copy number. The eutrophic environment and the presence of high concentrations of metals in the MSS could explain the predominance of *Beta-proteobacteria* and *Bacteroidetes*. Indeed, according to Brümmer *et al*. [33], the predominance of *Beta-proteobacteria* is associated with the presence of high concentrations of ammonia and metals in contaminated water.

Our freshwater tropical sediment results differ from those reported recently for temperate sediments showing that *Proteobacteria* (especially *Deltaproteobacteria*) and *Acidobacteria* were the most abundant phylas [34]. In addition, *Bacteroidetes* were found to be in low proportion in freshwater sediment, albeit enriched when interdital wetland sediments were analyzed [34]. It should be noted that our data presented a taxonomic similarity with a previous investigation in tropical pristine sediment [8].

Several bacterial species that play an important role in metal contaminated environments were found to inhabit the MSS, as supported by the recruitment plots (Figs. 2 and S2A-E). The *Beta-proteobacteria* class harbors chemolithoautotrophic members as ferrous iron oxidizing bacteria (FeOB), which were broadly represented in our data [35]. The *Gallionellaceae* family was represented by *Sideroxydans lithotrophicus*, a neutrophilic FeOB that prefers low oxygen and iron-rich environments [7]. *S. lithotrophicus* may play an important role in the removal of As from the MSS environment as FeIII binds with arsenate (AsV), which facilitates its precipitation and decreases its bioavailability in the environment [7]. *Leptothrix chlolodnii,* which is often found in eutrophic freshwater environments, was detected in our analysis. This bacterium oxidizes MnII into manganese oxide (MnIII and MnIV) [36, 37]. The *Betaproteobacteria* found in our sample included, among others, *Thiobacillus denitrificans* and *Thiomonas cuprina*. The former oxidizes various reduced inorganic sulfur compounds, such as ferrous sulfide (FeS), coupling with the reduction of nitrate [38, 39]. The latter is an AsIII-oxidizing bacterium that is ubiquitous in arsenic-contaminated environments and is capable of gaining energy from the oxidation of reduced inorganic sulfur compounds (e.g., able to perform the dissimilatory oxidation of iron) [39, 40]. Three FeIII-reducing members of the *Deltaproteobacteria* class were detected. One of them, *Anaeromyxobacter dehalogenans*, is a dissimilatory FeIII-reducing bacterium known to gain energy with Fe reduction [41], a contrasting role to that performed by *Thiomonas cuprina*. The two other members belonged to the *Geobacter* genus and showed the highest abundance among the *Deltaproteobacteria* of the MSS metagenome. Members of this genus were the most recovered in enrichment cultures by FeIII reduction [42]. Altogether, the presence of these taxa may reflect the high concentrations of metals such as Fe, Mn, Cu, As, and Zn found in the MSS. Moreover, the genome of *Chitinophaga pinensis* was well represented in the fragment recruitment plots [43]. This species is associated with organic carbon cycling in both anaerobic and aerobic sediments through the breakdown of simple carbohydrates to organic acids and degradation of a wide range of biopolymers [44, 45].

Members of *Actinobacteria*, *Firmicutes*, and *Nitrospirae* are generally recovered in large proportions from freshwater environments [46, 34], which is in contrast to the present observation for the MSS. Studies suggest that the abundance of the *Actinobacteria* and *Firmicutes* phyla is significantly correlated with metal-contaminated environments, particularly resistance to As and Hg [7, 8, 47, 48]. However, the metal contamination found in MSS does not appear to favor their abundance. Future research will be needed to ascertain the reason for the observed decrease of the abundance of these bacteria in this freshwater sediment and to find whether it is a widespread phenomenon.

The members of the Archaea domain from the MSS belonged to the *Parvarchaeota* and *Crenarchaeota* phyla. The *Crenarchaeota* phylum has been previously described in metal-contaminated environments [7, 49-51]. Our data contrasted with previous studies on archaeal diversity in metal-impacted environments that usually find a predominance of *Crenarchaeota* [8, 52]. The *Parvarchaeota* phylum was recently proposed by Rinke et al. [53] from single-cell genome sequencing of an uncultured archaea. *Thaumarcheota* were represented, only in the metagenomic shotgun sequencing data, by the following ammonium oxidizer species: *Cenarchaeum symbiosum*, *Nitrosopumilus maritimus*, and Candidatus *Nitrosphaera gargensis* [54-56]. Previous investigation on water columns of the Amazon River also detected *Cenarchaeum symbiosum* and *Nitrosopumilus maritimus*, indicating the importance of these species in the nitrogen (N) cycle of sediment from freshwater environments [50, 57]. It should be noted that the species *Nitrosopumilus maritimus*, detected in the MSS metagenome, showed the highest genome coverage of archaeal reads, indicating that this chemolithoautotrophic nitrifier is globally distributed and is essential for the nitrification mechanisms in this environment.

The presence of various metal resistance genes detected in the MSS metagenome was expected, because the MSS exhibited high concentrations of As, Mn, Zn, and Cu. Despite absence of Co and Cd in MSS, resistance genes associated with cobalt-zinc-cadmium resistance were the most abundant. Resistance determinants to these metals are usually organized as an operon

harboring the genes *czcC*, *czcB*, and *czcA*, which are responsible for expression of an efflux pump that transports the ions Co^{+2}, Zn^{+2}, and Cd^{+2} out of the bacterial cell [58, 59]. A previous study investigated the expression of this operon in the presence of these metals separately and found that the expression was more efficient in the presence of high concentrations of Zn [60]. Thus, the high concentration of Zn in the MSS could explain the abundance of these genes in this environment. Moreover, genes that confer resistance to Hg were also found in MSS, despite the low concentration of this metal (<2.5 mg/Kg and <0.1 mg/L for sediment and water, respectively). This finding could be due to the fact that the Hg resistance genes are co-selected as they are usually located on plasmids and transposons that harbor other resistance genes, such as resistance to betalactamic antibiotics, kanamycin, tetracycline, and others [61-63].

The Cu resistance gene, the second most abundant in the MSS, may be related to bacterial cell protection mechanisms against high concentrations of this metal found in this environment. Cu is an essential metal for the metabolism of the cell, because it is required as a cofactor for several enzymes [64]. Nevertheless, high concentrations of this metal may be toxic for the bacterial cells that have developed homeostasis mechanisms to ensure appropriate internal concentrations of Cu [65].

The As resistance mechanism most widespread in the environment, performed by the ArsC enzyme, was not detected in the MSS. Interestingly, the other genes of the *ars* operon (*arsA*, *arsB*, *arsD*, *arsH*, *arsR*) were found. The bacterial respiratory arsenate reductase enzymes encoded by the *arrA* and *arrB* genes was abundant in the MSS metagenome. A previous study from our group [11] investigating As resistance genes in the MSS using a metagenomic approach also found that the *arrA* gene was the most diverse As gene in the sample, indicating that this dissimilatory arsenate reduction is the most frequent activity. This microbial reduction is one of the main pathways involved in As mobilization in anoxic environments because release of the most toxic and soluble form of As, AsIII, by reducing Fe- or Mn-oxides may increase the contamination of water bodies [66].

Microbial community physiological profile analysis based on the ability to use different carbon sources has been successfully used to characterize microbial diversity in different environments [67-69]. Xiong *et al*. observed that soil uncontaminated by As showed greater metabolic diversity (C sources consumed) than soil newly contaminated with this metalloid, indicating that the microbial community was affected by this contamination. By contrast, our data showed a high metabolic diversity in the MSS, suggesting that As contamination is most likely not affecting the microbial diversity. Furthermore, other studies also reported that high nutrient concentrations in metal-contaminated sediments promote prokaryotic diversity [5, 71].

In freshwater ecosystems, phosphorus (P) and N are limiting nutrients, i.e., variation of these nutrient concentrations limits biological productivity. These nutrients were previously found in various organic and inorganic forms, and their bioavailability to higher trophic levels occurred through microbial transformations, because the organisms used them for growth and, in some cases, as an energy source [72].

The major transformations of N are N fixation, nitrification, denitrification, anammox, and ammonification, all highly dependent on the activities of a diverse assemblage of microorganisms such as bacteria, archaea, and fungi [73, 74]. In addition to metal contamination, Mina Stream is considered to be a eutrophic water body containing high concentrations of total N and its inorganic forms, nitrate (NO_3^{2-}N, 3103.8 µg l^{-1}) and ammonium (NH_4^+-N, 829.5 µg l^{-1}). Thus, it is likely that several bacterial and archaeal species related to the N cycle, such as *Thiobacillus denitrificans*, Candidatus *Nitrospira defluvii*, *Cenarchaeum symbiosum*, *Nitrosopumilus maritimus*, and Candidatus *Nitrosphaera gargensis*, among others, may play important roles in the N metabolism of the MSS. *Candidatus Nitrospira defluvii* was highly abundant in the MSS metagenome, being the bacterial genome with the highest coverage. These bacterial species are the dominant nitrite-oxidizing species in wastewater treatment plants and have already been found in metal-contaminated sediments [7, 75].

Analysis of N cycling genes from the MSS metagenome unveiled ammonium assimilation and ammonification as the two most abundant N cycle processes. Indeed, genes responsible for ammonium assimilation such as glutamate synthase (EC 1.4.1.13 and EC 1.4.1.14) and glutamine synthetase type I and type III (EC 6.3.1.2) were detected in our samples. Ammonium assimilation performed by the microbial community can retain N and make the sediment act as a temporary buffer in aquatic environments [76, 77]. The ammonification process is performed by saprophytic bacteria and is based on the decomposition of organic molecules containing N, e.g., amino acids and DNA that are released into the environment when an organism excretes waste or dies. N is required for the survival of all organisms, because it is an essential component of DNA, RNA, and protein, and thus, is essential for the maintenance of the aquatic microbial community. As most N exists in the form of organic molecules, the availability of N to higher trophic levels depends on microbial transformation.

In conclusion, our data reveal that the microbial communities from the MSS have significantly different features than those presented by other metal-contaminated environments. The data recovered agree with the expected assemblage of organisms thriving in metal-rich and eutrophic environments. This study provides important insights into the structure of the prokaryotic community of a tropical freshwater sediment, indicating a possible role for this community in the N and C cycles and in the transformation of Fe and As. Functional annotation unveiled a high degree of diversity of several metal resistance genes, indicating that this microbial community is well adapted to environments containing metal contamination. Finally, the results reported here expand the current knowledge of the microbial taxonomic and functional composition of tropical, metal-contaminated, freshwater sediments. Our data, together with those revealed by many other research efforts across the globe, may be an indirect and yet relevant contribution to the enormous endeavor being championed by the Earth microbiome project.

Acknowledgments

This study was supported by grants from CNPq (Instituto Nacional de Ciência e Tecnologia em Recursos Minerais, Água e Biodiversidade – INCT – ACQUA nº 15206-7, and Universal nº 472411/2012-8), FAPEMIG (nº APQ-00801-12), and Vale. We appreciate the technical support provided by the Laboratório de Análises Químicas/DEMET/UFMG do Instituto Nacional de Ciência e Tecnologia em Recursos Minerais, Água e Biodiversidade – INCT – ACQUA in the chemical analyses. AMAN, FB, and GO are recipients of CNPq fellowships.

References

1 - Kasting JF, Siefert JL (2002) Life and the evolution of Earth's atmosphere. Science 296: 1066-1068.

2- Xu J (2011) Microbial ecology in the age of metagenomics. In: Bruijn J, Willey J Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches.

3 - Earth Microbiome Project - http://www.earthmicrobiome.org. Accessed 15 May 2013

4 – Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. PNAS 104: 11436-11440.

5 – Bouskill NJ, Barker-Finkel J, Galloway TS, Handy RD, Ford TE (2010) Temporal bacterial diversity associated with metal-contaminated river sediments. Ecotoxicol 19: 317-328.

6 - Rezende PS, Moura PAS, Durão Jr WA, Nascentes CC, Windmöller CC, Costa LM (2011) Arsenic and mercury mobility in Brazilian sediments from the São Francisco River Basin. J Braz Chem Soc 22: 910-918.

7 - Halter D, Cordi A, Gribaldo S, Gallien S, Goulhen-Chollet F, Heinrich- Salmeron A *et al*. (2011) Taxonomic and functional prokaryote diversity in mildly arsenic contamined sediments. Res Microbiol 162: 1-11.

8 - Reis MP, Barbosa FA, Chartone-Souza E, Nascimento AMA (2013) The prokaryotic community of a historically mining-impacted tropical stream sediment is as diverse as that from a pristine stream sediment. Extremophiles 17: 301-319.

9 - Gillan DC, Roosa S, Kunath B, Billon G, Wattiez R (2014) The long-term adaptation of bacterial communities in metal-contaminated sediments: a metaproteogenomic study. Environ Microbiol 1-15.

10 - Gough HL, Stahl DA (2011) Microbial community structures in anoxic freshwater lake sediment along a metal contamination gradient. ISME J 5: 543-58.

11 - Costa PS, Scholte LL, Reis MP, Chaves AV, Oliveira PL, Itabayana LB, Suhadolnik ML, Barbosa FA, Chartone-Souza E, Nascimento AMA (2014) Bacteria and genes involved in arsenic speciation in sediment impacted by long-term gold mining. PLoS One 9:e95655.

12 - Gilbert J.A., Dupont C.L. (2011) Microbial metagenomics: beyond the genome. *Ann Rev Mar Sci* **3**: 347-71.

13 - Borba RP, Figueiredo BR, Rawlins BG, Matchullat J (2000) Arsenic in water and sediment inthe Iron Quadrangle, Minas Gerais state, Brasil. Rev Brasil Geocienc 30: 554-557.

14 - Deschamps E, Ciminelli VST, Raue B, Lange FT, Schmidt H (2002) Soil and sediment geochemistry of the Iron Quadrangle, Brazil. J Soil Sediment 2: 216-222.

15 - Borba RP, Figueiredo BR, Cavalcanti JA (2004) Arsênio na água subterrânea em Ouro Preto e Mariana, Quadrilátero Ferrífero (MG). R Esc Minas 57: 45-51.

16 - Bundschuh J, Litter MI, Parvez F, Román-Ross G, Nicolli HB, Jean JS *et al*. (2012) One century of arsenic exposure in Latin America: a review of history and occurrence from 14 countries. Sci Total Environ 429: 2-35.

17 - Mailloux BJ, Trembath-Reichert E, Cheung J, Watson M, Stute M, Freyer GA, Ferguson AS, Ahmed KM, Alam MJ, Buchholz BA, Thomas J, Layton AC, Zheng Y, Bostick BC, van Geen A (2013) Advection of surface-derived organic carbon fuels microbial reduction in Bangladesh groundwater. PNAS 110:5331-5.

18 - Garland JL, Mills AL (1991) Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. Appl Environ Microbiol 57:2351-2359.

19 - Garland JL (1996) Analysis and interpretation of community-level physiological profiles in microbial ecology. FEMS Microbiol Ecol 24: 289-300

20 - Klindworth A, Pruesse E, Schweer T, Peplles J, Quast C, et al. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41:1-11.

21 - Mothur Pipeline MiSeq Sop - http://www.mothur.org/wiki/MiSeq_SOP-20-07-2014. Accessed 20 April 2014.

22 - Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral J-P, Raoult D (2000) 16S Ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 34:3623–3630.

23 - Good IJ (1953) The population frequencies of species and the estimation of population parameters. Biometrika 40:237–262.

24 - Namiki T, Hachiya T, Tanaka H, Sakakibara Y (2012) MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res 1:e155.

25 - Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al*. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 19:386.

26 - Mitra S, Rupek P, Richter DC, Urich T, Gilbert JA, Meyer F *et al*. (2011) Functional analysis of metagenomes and metatranscriptomes using SEED and KEGG. BMC Bioinformatics 12:1-8.

27 - Parks DH, Beiko RG (2010) Identifying biologically relevant differences between metagenomic communities. Bioinformatics 26: 715-721.

28 - Raskin L, Stromley JM, Rittmmann BE, Stahl DA (1994) Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. Appl Environ Microbiol 60: 1232-1240.

29 - Muyzer G, De Waal EC, Uitterlinden AG (1993) Profiling of complexmicrobial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for16S rRNA. Appl Environ Microbiol 59: 695-700.

30 - Cardinali-Rezende J, Colturato LF, Colturato TD, Chartone-Souza E, Nascimento AMA, Sanz JL (2012) Prokaryotic diversity and dynamics in a full-scale municipal solid waste anaerobic reactor from start-up to steady-state conditions. Bioresour Technol 119: 373-83.

31 - Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature 437:543-6.

32 - Liao VH, Chu Y, Su Y, Hsiao S, Wei C, Liu C, et al. (2011) Arsenite-oxidizing and arsenate-reducing bacteria associated with arsenic-rich groundwater in Taiwan. J Contam Hydrol 123: 20–9.

33 - Brümmer IH, Fehr W, Wagner-Dobler I (2000) Biofilm community structure in polluted rivers: abundance of dominant phylogenetic groups over a complete annual cycle. Appl Environ Microbiol 66: 3078–3082.

34 - Wang Y, Sheng HF He Y, Wu JY, Jiang YX, Tam NF *et al*. (2012) Comparison of the levels of bacterial diversity in freshwater, intertidal wetland, and marine sediments by using millions of illumina tags. Appl Environ Microbiol 78:8264-8271.

35 - Hedrich S, Schlömann M, Johnson DB (2011) The iron-oxidizing proteobacteria. Microbiology 157:1551–1564.

36 - Tebo BM, Johnson HA, McCarthy JK, Templeton AS (2005) Geomicrobiology of manganese(II) oxidation. Trends Microbiol 13: 421-8.

37 - Takeda M, Kawasaki Y, Umezu T, Shimura S, Hasegawa M, Koizumi J (2012) Patterns of sheath elongation, cell proliferation, and manganese(II) oxidation in *Leptothrix cholodnii*. Arch Microbiol 194: 667-73.

38 - Jørgensen CJ, Jacobsen OS, Elberling B, Aamand J (2009) Microbial oxidation of pyrite coupled to nitrate reduction in anoxic groundwater sediment. Environ Sci Technol 43: 4851-4857.

39 - Hedrich S, Schlömann M, Johnson DB (2011) The iron-oxidizing proteobacteria. Microbiology 157: 1551-1564.

40 - Bryan CG, Marchal M, Battaglia-Brunet F, Kugler V, Lemaitre-Guillier C, Lièvremont D, *et al*. (2009) Carbon and arsenic metabolism in *Thiomonas* strains: differences revealed diverse adaptation processes. BMC Microbiol 23: 127.

41 - Truede N, Rosencrantz D, Liesack W, Schnell S (2003) Strain Fac12, a dissimilatory iron-reducing member of the *Anaeromyxobacter* subgroup of Myxococcales. FEMS Microbiol Lett 44: 261-269.

42 - Snoeyenbos-West O, Nevin KP, Anderson RT, Lovley DR (2000) Enrichment of *Geobacter* species in response to stimulation of Fe(III) reduction in sandy aquifer sediments. Microbial Ecol 39: 153-167.

43 - Del Rio TG, Abt B, Spring S, Lapidus A, Nolan M, Tice H et al. (2010) Complete genome sequence of Chitinophaga pinensis type strain (UQM 2034T). Stand Genomic Sci 2: 87–95.

44 - Pankratov TA, Kulichevskaya IS, Liesack W, Dedysh SN (2006) Isolation of aerobic, gliding, xylanolytic and laminarinolytic bacteria from acidic Sphagnum peatlands and emended description of *Chitinophaga arvensicola* Kampfer et al. 2006. Int J Syst Evol Microbiol 56: 2761-4.

45 - Briée C, Moreira D, López-García P (2007) Archaeal and bacterial community composition of sediment and plankton from a suboxic freshwater pond. Res Microbiol 158: 213-27.

46 - Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S (2011) A guide to the natural history of freshwater lake bacteria. Microbiol Mol Biol Rev 75: 14-49.

47 - Gremion F, Chatzinotas A, Harms H (2003) Comparative 16S rDNA and 16S rRNA sequence analysis indicates that Actinobacteria might be a dominant part of the metabolically active bacteria in heavy metal-contaminated bulk and rhizosphere soil. Environ Microbiol 5: 896-907.

48 - Vishnivetskaya TA, Mosher JJ, Palumbo AV, Yang ZK, Podar M, Brown SD *et al*. (2011) Mercury and other heavy metals influence bacterial community structure in contaminated Tennessee streams. Appl Environ Microbiol 77: 302-311.

49 - Bruneel O, Pascault N, Egal M, Bancon-Montigny C, Goni-Urriza MS, Elbaz-Poulichet F *et al*. (2008) Archaeal diversity in a Fe-As rich acid mine drainage at Carnoules (France). Extremophiles 12: 563-571.

50 - Porat I, Vishnivetskaya TA, Mosher JJ, Brandt CC, Yang ZK, Brooks SC *et al*. (2010) Characterization of archaeal community in contaminated and uncontaminated surface stream sediments. Microb Ecol 60: 784-95.

51 - Haller L, Tonolla M, Zopfi J, Peduzzi R, Wildi W, Poté J (2011) Composition of bacterial and archaeal communities in freshwater sediments with different contamination levels (Lake Geneva, Switzerland). Water Res 45: 1213-28.

52 - Gupta RS, Shami A (2011) Molecular signatures for the Crenarchaeota and the Thaumarchaeota. Antonie van Leeuwenhoek 99: 133-157.

53 - Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T (2013) Insights into the phylogeny and coding potential of microbial dark matter. Nature 499:431-7.

54 - Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. Nat Rev Microbiol 6: 245-52.

55 - Spang A, Hatzenpichler R, Brochier-Armanet C, Rattei T, Tischler P, Spieck E *et al*. (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. Trends Microbiol 18: 331-40.

56 - Tourna M, Stieglmeier M, Spang A, Könneke M, Schintlmeister A, Urich T *et al*. (2011) *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *PNAS* 108: 8420-5.

57 - Ghai R, Rodriguez-Valera F, McMahon KD, Toyama D, Rinke R, Oliveira SC *et al*. (2011) Metagenomics of the water column in the pristine upper course of the Amazon river. PLoS One 6: e23785.

58 - Nies DH, Brown N (1997) Two-component systems in the regulation of heavy metal resistance. In: Silver S, Walden W (ed.). Metal ions in gene regulation. Chapman & Hall, London, England.

59 - Nies DH, Rehbein G, Hoffmann T, Baumann C, Grosse C (2006) Paralogs of genes encoding metal resistance proteins in *Cupriavidus metallidurans* strain CH34. J Mol Microbiol Biotechnol 11:82-93.

60 - Legatzki A, Grass G, Anton A, Rensing C, Nies DH (2003) Interplay of the Czc system and two P-type ATPases in conferring metal resistance to *Ralstonia metallidurans*. J Bacteriol 185: 4354-4361.

61 - Silver S, Phung LT (2005) Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. Appl Environ Microbiol 71: 599-608.

62 - Soge OO, Beck NK, White TM, No DB, Roberts MC (2008) A novel transposon, Tn6009, composed of a Tn916 element linked with a *Staphylococcus aureus* mer operon. J Antimicrob Chemother 62: 674-80.

63 - Novais A, Baquero F, Machado E, Cantón R, Peixe L, Coque TM (2010) International spread and persistence of TEM-24 is caused by the confluence of highly penetrating enterobacteriaceae clones and an IncA/C2 plasmid containing Tn1696::Tn1 and IS5075-Tn21. Antimicrob Agents Chemother 54: 825-34.

64 - Krupanidhi S, Sreekumar A, Sanjeevi CB (2008) Copper & biological health. Indian J Med Res 128: 448-61.

65 - Solioz M, Abicht HK, Mermod M, Mancini S (2010) Response of gram-positive bacteria to copper stress. J Biol Inorg Chem 15: 3-14.

66 - Song B, Chyun E, Jaffé PR, Ward BB (2009) Molecular methods to detect and monitor dissimilatory arsenate-respiring bacteria (DARB) in sediments. *FEMS Microbiol Ecol* 68:108-17.

67 - Mondini C, Insam H (2003) Community level physiological profiling as a tool to evaluate compost maturity: a kinetic approach. Eur J Soil Biol 39:141-148.

68 - Cardinali-Rezende J, Moraes AMM, Colturato LFDB, Carneiro EV, Marriel IE, Chartone-Souza E *et al*. (2011) Phylogenetic and physiological characterization of organic waste-degrading bacterial communities. World J Microbiol Biotechnol 27: 245-252.

69 - Leon C, Campos V, Urrutia R, Mondaca MA (2012). Metabolic and molecular characterization of bacterial community associated to Patagonian Chilean oligotrophic-lakes of quaternary glacial origin. World J Microbiol Biotechnol 28: 1511-21.

70 - Xiong J, Wu L, Tu S, Van Nostrand JD, He Z, Zhou J *et al*. (2010) Microbial communities and functional genes associated with soil arsenic contamination and the rhizosphere of the arsenic-hyperaccumulating plant *Pteris vittata* L. Appl Environ Microbiol 76: 7277-7284.

71 - Zhou J, Kia B, Treves DS, Wu L-Y, Marsh TL, O'Neill RV *et al.* (2002) Spatial and resource factors influencing high microbial diversity in soil. Appl Environ Microbiol 68: 326–334.

72 - Barlett MA, Leff LG (2010) The effects of N:P ratio and nitrogen form on four major freshwater bacterial taxa in biofilms. Can J Microbiol 56: 32-43.

73 - Purkhold U, Pommerening-Roser A, Juretschko S, Schmid MC, Koops HP, Wagner M (2000) Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and *amoA* sequence analysis: implications for molecular diversity surveys. Appl Environ Microbiol 66: 5368-5382.

74 - Wuchter C, Abbas B, Coolen MJ, Herfort L, van Bleijswijk J, Timmers *et al*. (2006) Archaeal nitrification in the ocean. PNAS 103: 12317-12322.

75 - Maixner F, Wagner M, Lücker S, Pelletier E, Schmitz-Esser S, Hace K *et al*. (2008) Environmental genomics reveals a functional chlorite dismutase in the nitrite-oxidizing bacterium *'Candidatus Nitrospira defluvii'*. Environ Microbiol 10: 3043-56.

76 - Dähnke K, Emeis K, Johannsen A, Nagel B (2010) Stable isotope composition and turnover of nitrate in the German Bight. Mar Ecol Prog Ser 408: 7-18.

77 - Dähnke K, Moneta A, Veuger B, Soetaert K, Middelburg JJ (2012) Nitrogen turnover in a tidal flat sediment: assimilation and dissimilation by bacteria and benthic microalgae. Biogeosciences Discuss 9: 6987-7019.

Figures

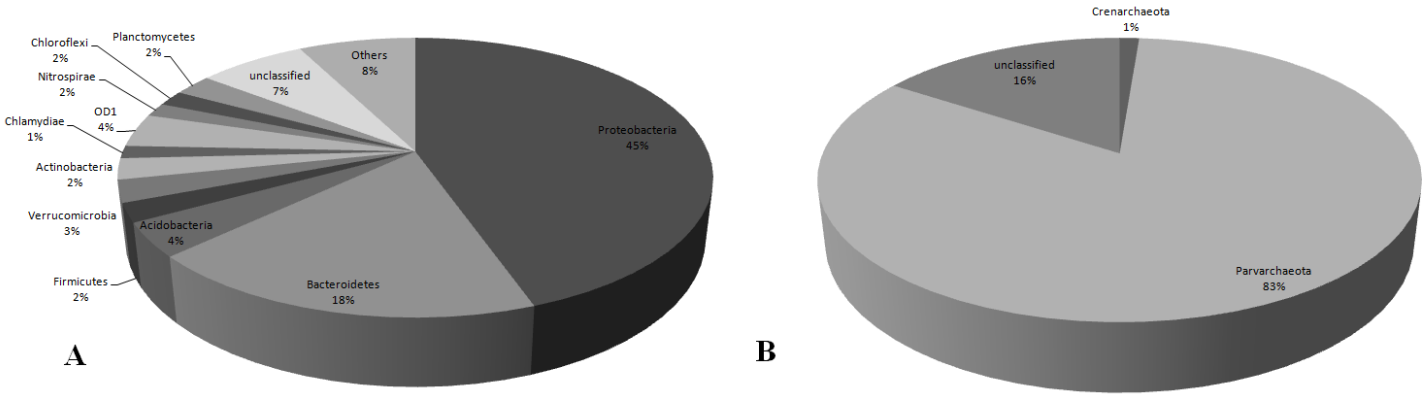

Figure 1 - Taxonomic composition of bacterial (A) and archaeal (B) taxa from MSS microbiota based on the Greengenes database. Other bacteria: *Gemmatimonadetes*, *Cyanobacteria*, *OP3*, *OP11*, *Spirochaetes*, *TM7*, *Chlorobi*, *WS3*, *Elusimicrobia*, *GN04*, *TM6*, *GN02*, *Tenericutes*, *Armatimonadetes*, *BRC1*, *NC10*, *WPS*-2, *Fibrobacteres*, *Fusobacteria*, *H-178*, *FCPU426*, *Kazan*-3B-28, *WS5*, *NKB19*, *Thermi*, *AC1*, *TPD*-58, *WS6*, *Synergistetes*, *OP8*, *WS2*, *ZB3*, *SC4*, *OP1*, *SBR1093*, *SR1*, *Lentisphaerae*, *GAL15*, *PAUC34f*, *LCP*-89 and *MVS*-104.

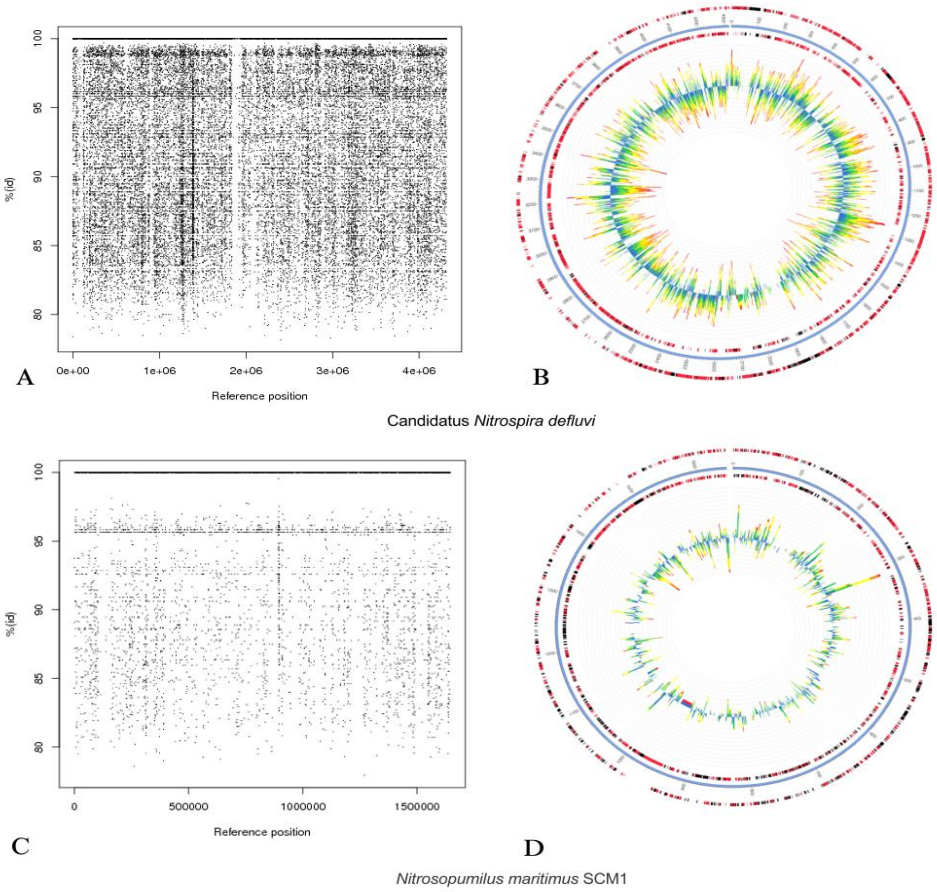

Figure 2 - Fragment recruitment plots of the MSS contigs by Candidatus *Nitrospira defluvi* (A and B - FP929003.1) and *Nitrosopumilus maritimus* (C and D - CP000866.1) genomes. The comparison was made using BLASTn. Vertical axis showed the % identity of the metagenomic contigs to the respective bacterial or archaeal genome. A and C – recruitment by R software; B and D – recruitment by MG-RAST.

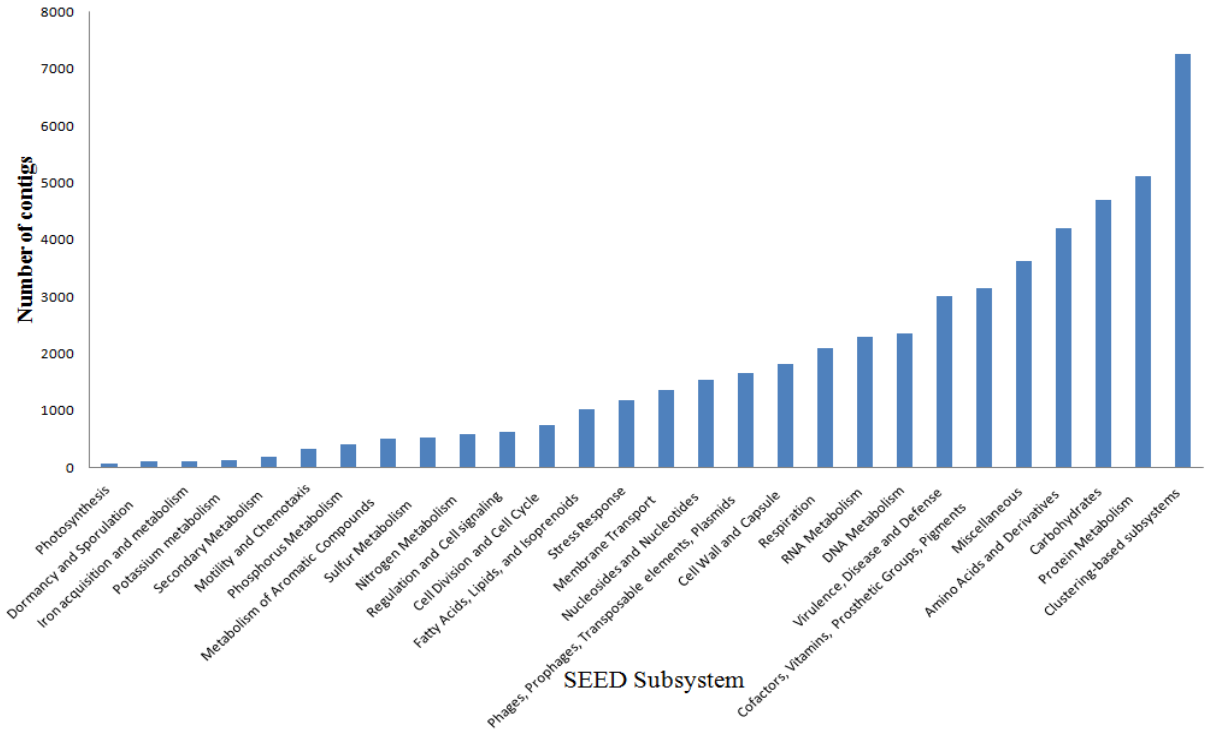

Figure 3 – SEED subsystems distribution of the MSS metagenome based on MG-RAST annotation. The cutoff parameters were e-value $1x10^{-5}$ and 60% of identity.

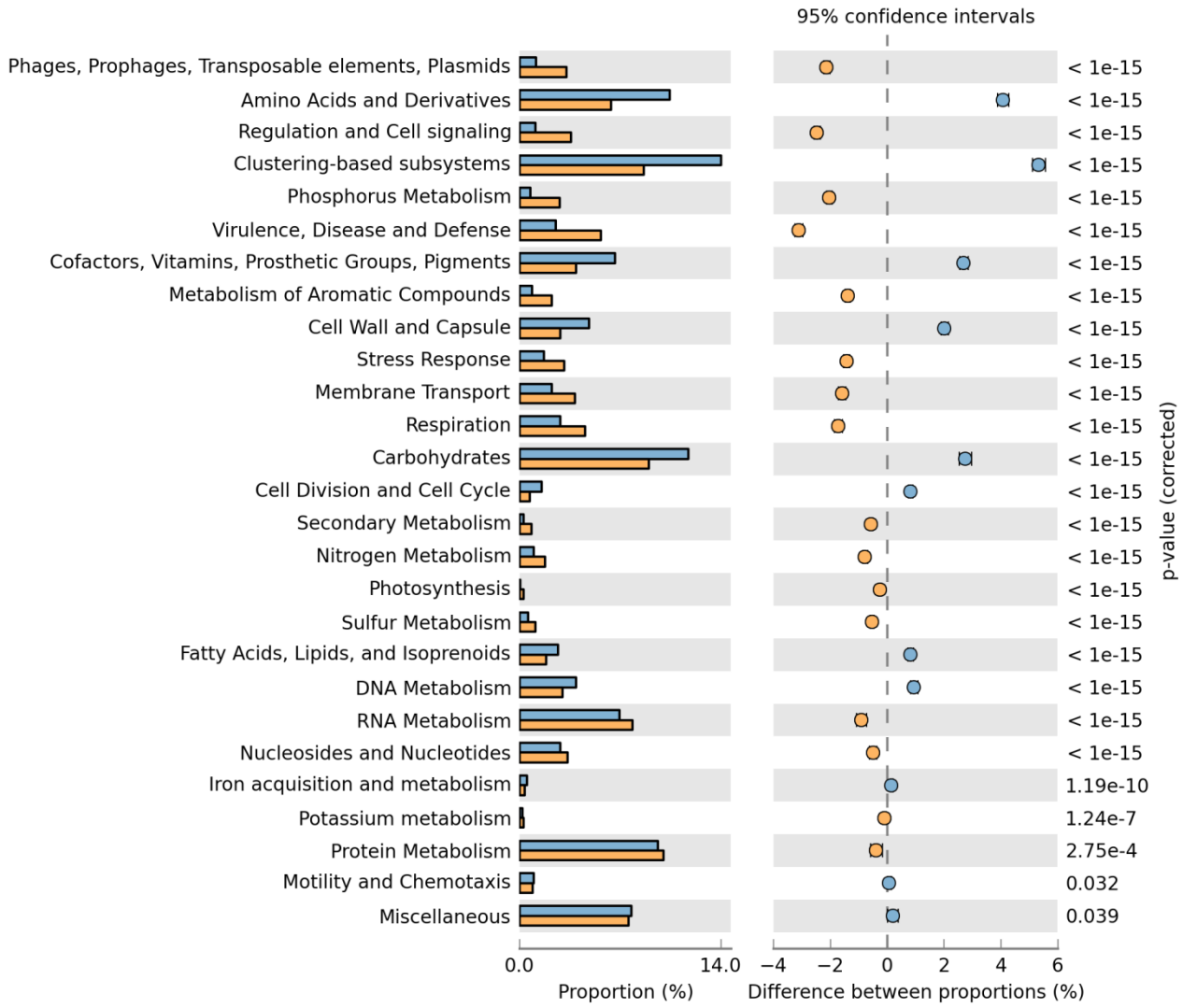

Figure 4 - Significant SEED subsystem differences as a result of a Fisher exact test between the MSS and RAW metagenomes conducted with the STAMP program. Enrichment of SEED subsystem in the RAW metagenome has a positive difference between proportions (blue circles), whereas enrichment of SEED subsystem in the MSS metagenome has a negative difference between proportions (orange circles). Bars on the left represent the proportion of each subsystem in the data. Subsystems difference with a *p value* of >0.05 were considered to be significant.

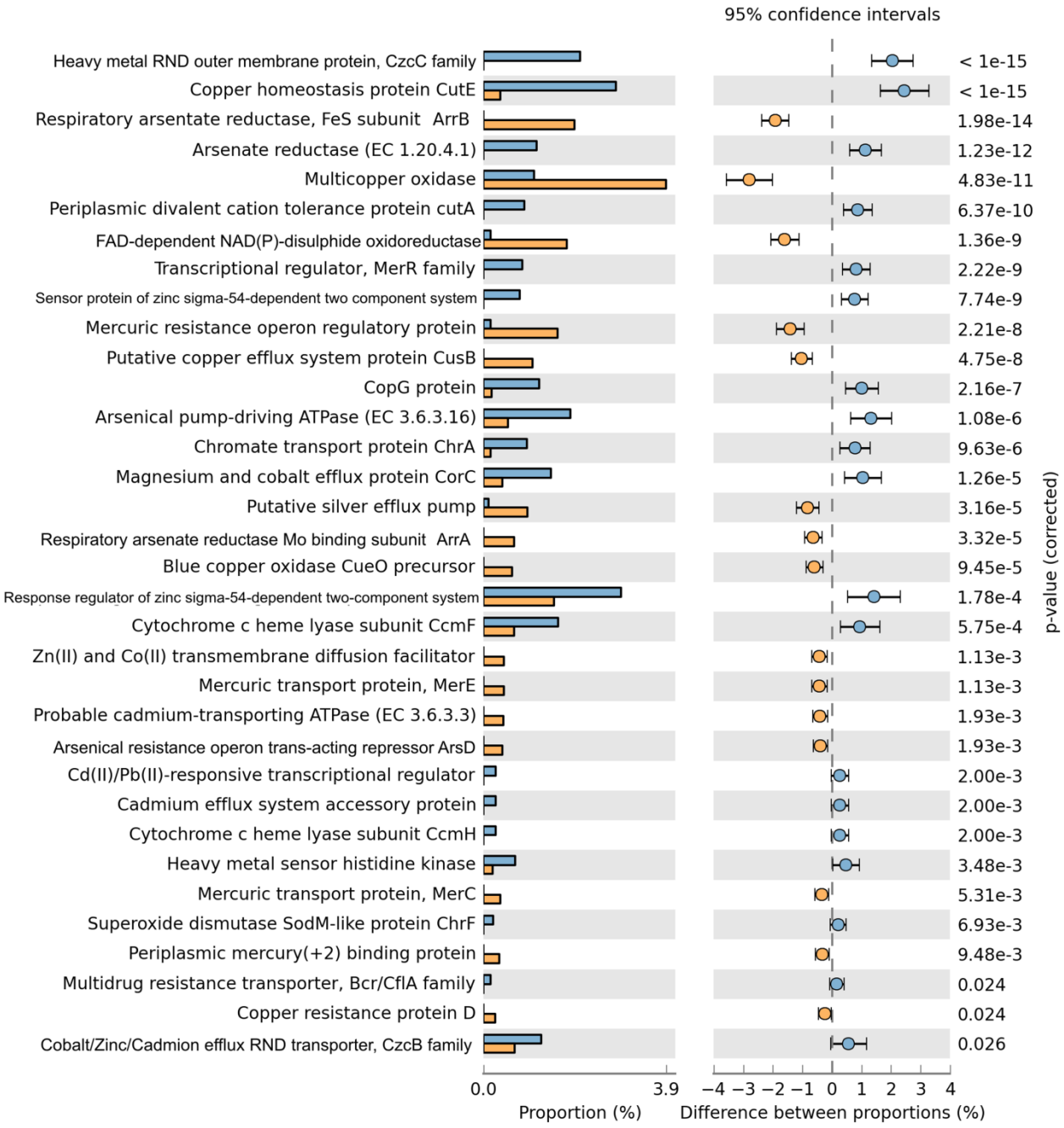

Figure 5 - Significant metal resistance genes differences as a result of a Fisher exact test between the MSS and RAW metagenomes conducted with the STAMP program. Enrichment of metal resistance genes in the RAW metagenome has a positive difference between proportions (blue circles), whereas enrichment of metal resistance genes in the MSS metagenome has a negative difference between proportions (orange circles). Barson the left represent the proportion of each metal resistance protein in the data. Metal resistance difference with a *p value* of >0.05 were considered to be significant.

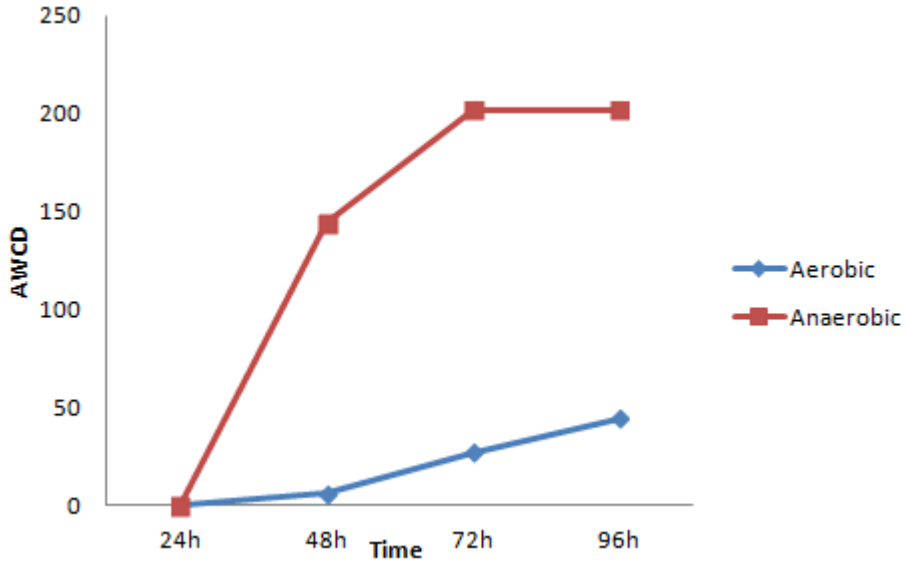

Figure 6 – Average well-color development (AWCD) calculated from the consumption of carbon sources of anaerobic and aerobic microbial communities.

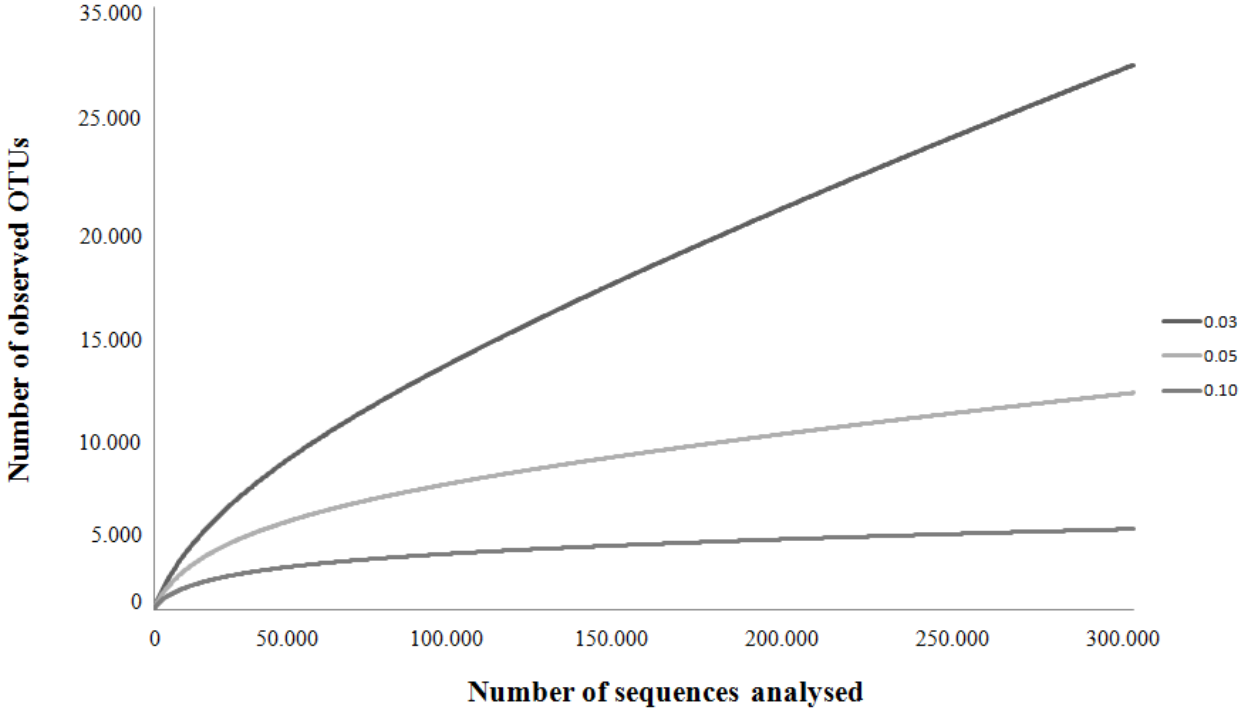

Figure S1 - Rarefaction curve of number of OTUs observed with an evolutionary distance of 0.03, 0.05 and 0.10.

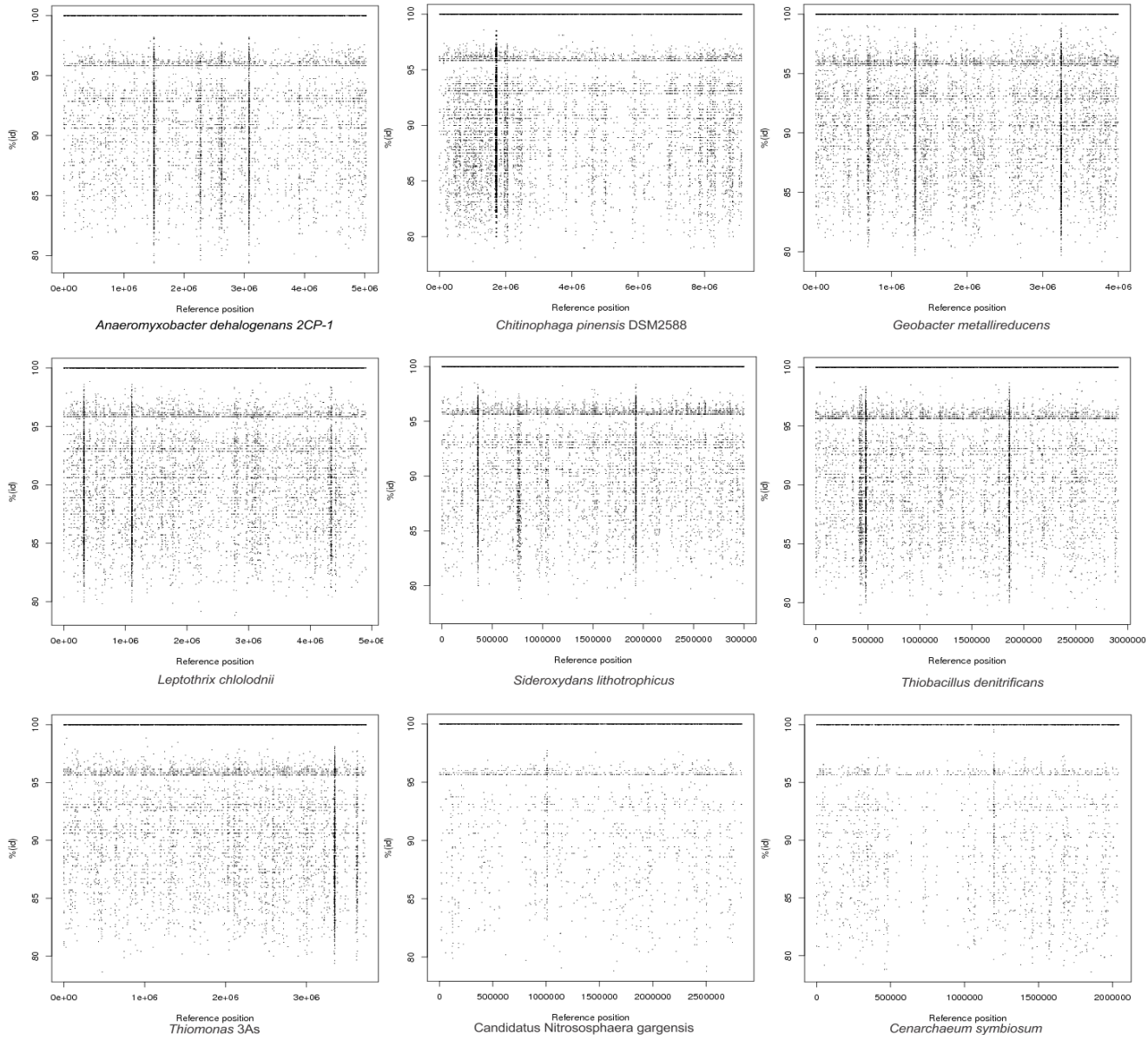

Figure S2 – Fragment recruitment plots of the MSS contigs by bacterial and archaeal genomes. The comparison was made using BLASTn. Vertical axis showed the % identity of the metagenomic contigs to the respective bacterial or archaeal genome. A - *Anaeromyxobacter dehalogenans* 2CP-1 (CP000251.1); B - *Chitinophaga pinensis* DSM2588 (CP001699.1); C - *Geobacter metallireducens* (CP000148.1); D - *Leptothrix chlolodnii* (CP001013.1); E - *Sideroxydans lithotrophicus* (CP001965.1); F - *Thiobacillus denitrificans* ATCC25259 (CP000116.1); G - *Thiomonas arsenitoxydans* 3As (FP475956.1); H - Candidatus *Nitrososphaera gargensis* (CP002408.1); I - *Cenarchaeum symbiosum* (DP000238.1).

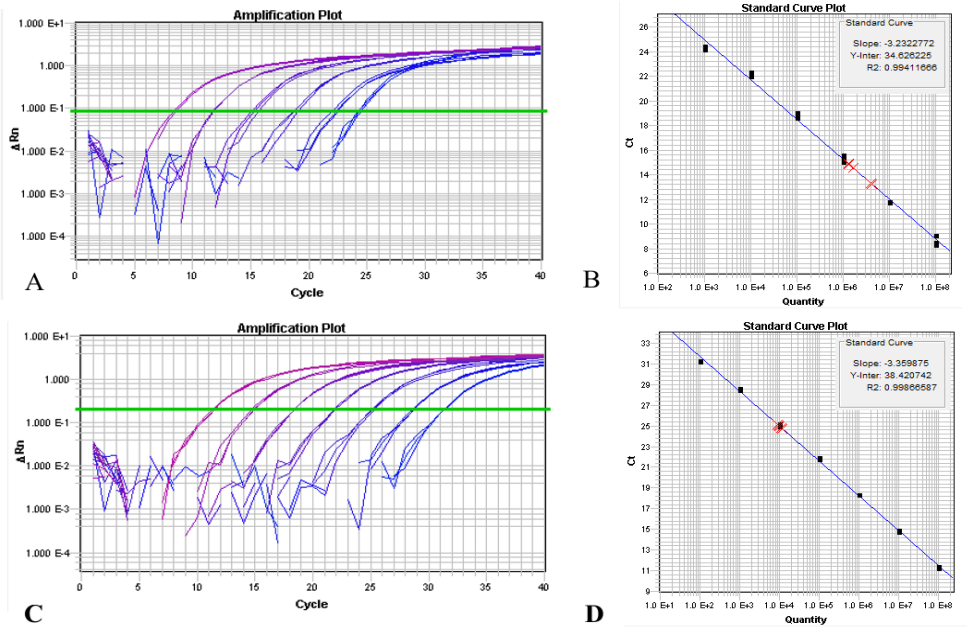

Figure S3 - The bacterial standard curve (A and B) and the archaeal qPCR standard curve (C and D).

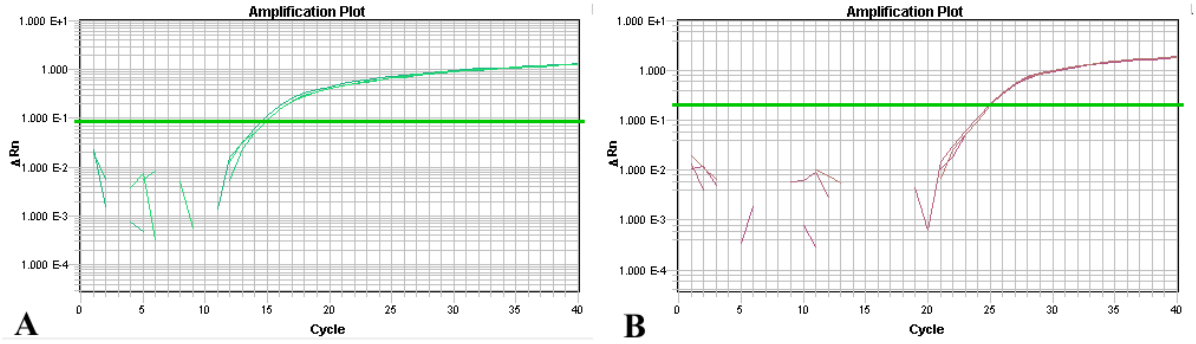

Figure S4 - The C_t values from the 16S rRNA gene amplifications. A and B represent bacterial and archaeal amplifications, respectively.

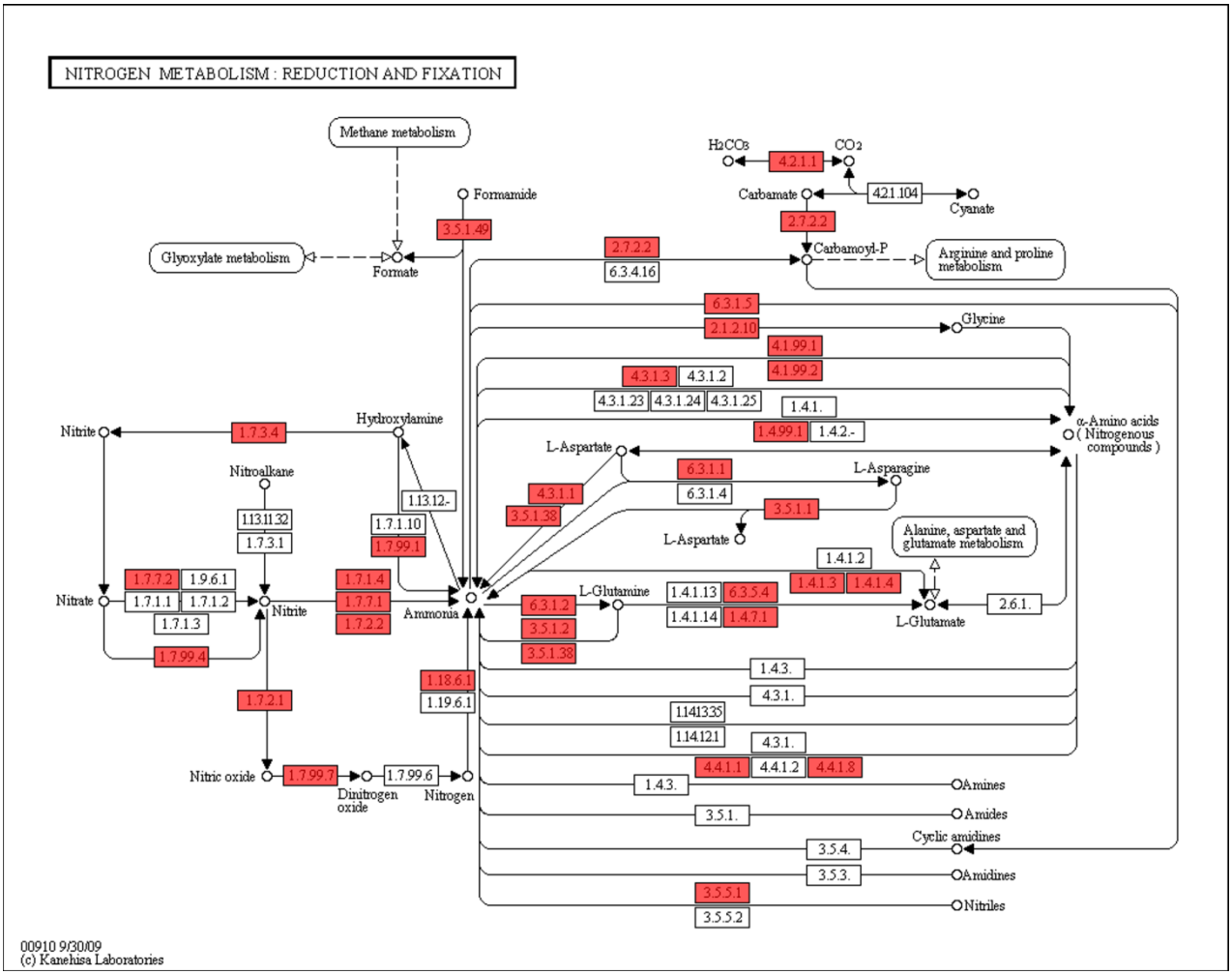

Figure S5 – Nitrogen cycle representation obtained in the Keeg Mapper analysis of MG RAST web server based on SEED database. The red square represents the presence of enzyme sequence in the MSS metagenome.

Table 1 – The most frequent nitrogen metabolism genes in the MSS metagenome obtained using the MG-RAST web server based on SEED database.

Nitrogen Metabolism	Protein and gene	Number of contigs
Allantoin Utilization	2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)	43
	Allantoate amidohydrolase (EC 3.5.3.9)	50
	Ureidoglycolate dehydrogenase (EC 1.1.1.154)	53
Ammonia assimilation	Ammonium transporter	64
	Ferredoxin-dependent glutamate synthase (EC 1.4.7.1)	525
	Glutamate synthase [NADH] (EC 1.4.1.14)	43
	Glutamate synthase [NADPH] large and small chain (EC 1.4.1.13)	123
	Glutamate-ammonia-ligase adenylyltransferase - GlnE (EC 2.7.7.42)	44
	Glutamine synthetase type I and type III (EC 6.3.1.2)	130
Cyanate hydrolysis	Carbonic anhydrase - CynT (EC 4.2.1.1)	44
Nitrate and nitrite ammonification	Assimilatory nitrate reductase large subunit (EC:1.7.99.4)	77
	Nitrite reductase [NAD(P)H] small subunit (EC 1.7.1.4)	30
	Nitrite reductase probable electron transfer 4Fe-S subunit (EC 1.7.1.4)	210
	NrfC protein	51
	Polyferredoxin NapH (periplasmic nitrate reductase)	59
	Putative thiol:disulfide oxidoreductase, nitrite reductase complex assembly	76
	Respiratory nitrate reductase delta chain (EC 1.7.99.4)	42
	Respiratory nitrate reductase subunit, conjectural (EC 1.7.99.4)	78
Nitric oxide synthase	Manganese superoxide dismutase (EC 1.15.1.1)	42
Nitrogen fixation	AnfO protein, required for Mo- and V-independent nitrogenase	51
	Nitrogenase (vanadium-iron) beta chain (EC 1.18.6.1)	47

Table 2 – The most frequent metal resistance genes in the MSS metagenome obtained using the MG-RAST web server based on SEED database

Metal resistance	Proteins	Number of contigs
Arsenic resistance	Arsenic resistance operon (ArsB, ArsH, ArsA, ArsR, ArsD)	67
	Arsenical-resistance protein ACR3	112
	Respiratory arsenate reductase, Mo binding subunit and FeS subunit (ArrA and ArrB)	124
Cobalt-zinc-cadmium resistance	Cadmium-transporting ATPase - CRA (EC 3.6.3.3)	24
	Probable cadmium-transporting ATPase - PCT (EC 3.6.3.3)	20
	Cation efflux system protein (CusA,CusR, CusC, CusB)	990
	Cobalt-zinc-cadmium resistance protein (CzcA, CzcD, CzcB, CzrR, CzrB)	2007
	Probable Co/Zn/Cd efflux system membrane fusion protein (CusB/CzsB)	160
	Putative silver efflux pump	45
Copper homeostasis	Copper-translocating P-type ATPase (EC 3.6.3.4)	504
	Cytochrome c heme lyase subunit CcmF	31
	Multicopper oxidase	187
Mercuric reductase	FAD-dependent NAD(P)-disulphide oxidoreductase	85
Mercury resistance operon	Mercuric resistance proteins (MerC, MerE, MerT, MerD, MerR, MerP, MerA)	200
Resistance to chromium compounds	Chromate resistance proteins (ChrI, ChrA, ChrC)	17
Zinc resistance	Response regulator of zinc sigma-54-dependent two-component system (ZraR)	72

Table S1 – Carbon sources utilization by aerobic and anaerobic bacterial communities and diversity index in sediment of the Mina stream.

Carbon source	Bacterial community	
	Aerobic	Anaerobic
β-Methyl-D-Glucoside	+	+
D-Galactonic Acid γ-Lactone	+	+
L-Arginine	+	+
Pyruvic Acid Methyl Ester	+	+
D-Xylose	+	+
D-Galacturonic Acid	+	+
L-Asparagine	+	+
Tween 40	+	+
i-Erythritol	+	+
2-Hydroxy Benzoic Acid	-	-
L-Phenylalanine	+	+
Tween 80	+	+
D-Mannitol	+	+
4-Hydroxy Benzoic Acid	+	+
L-Serine	+	+
α-Cyclodextrin	+	+
N-Acetyl-D-Glucosamine	+	+
γ- Hydroxybutyric Acid	+	+
L-Threonine	-	+
Glycogen	-	+
D-Glucosaminic Acid	+	+
Itaconic Acid	+	+
Glycyl-L-Glutamic Acid	+	+
D-Cellobiose	+	+
Glucose-1-Phosphate	+	+
α-Ketobutyric Acid	-	+
Phenylethylamine	+	+
α-D-Lactose	-	+
D,L-α-Glycerolphosphate	+	+
D-Malic acid	+	+
Putrescine	+	+
Richeness	26	30
Shannon- Weaver index	2.96	3.37
Simpson index	0.9	0.99

5 – Discussão Geral

A água é um elemento natural essencial à preservação da vida. Atualmente, a escassez de água potável e sua degradação por atividades antrópicas são de grande preocupação em todo o mundo, já que a poluição de ambientes aquáticos superficiais e subterrâneos é um problema global (Lord, 2001). O Córrego da Mina é um retrato desta realidade, sendo considerado eutrófico com altas concentrações de nitrogênio e suas formas inorgânicas. Além disso, a concentração de metais deste córrego excedeu o máximo permitido pela legislação brasileira (CONAMA, resolução de 2011). Como esperado, o sedimento exibiu maior concentração de metais que a água. Esta maior concentração pode ser explicada pela precipitação ou adsorção dos metais que são depositados no sedimento, com baixa concentração dos metais dissolvidos na água. Além disso, os sedimentos possuem papel importante no transporte e armazenamento de contaminantes, fazendo com que a sua qualidade determine a saúde ecológica do ambiente (Rao et al., 2011). Considerando a relevância que a comunidade microbiana desempenha nos ciclos biogeoquímicos de metais no sedimento, investigou-se neste estudo a taxonomia e as funções decorrentes da diversidade procariótica, utilizando duas abordagens distintas: abordagem clássica de cultivo, com ênfase em bactérias resistentes ao arsênio e os genes responsáveis pela sua transformação; e metagenoma, usando plataforma de sequenciamento de nova geração.

As análises dos isolados obtidos das culturas enriquecidas com AsIII e AsV, e do microbioma revelaram perfis taxonômicos distintos. O filo Proteobacteria predominou em ambas abordagens apesar de distribuição distinta de suas classes. A seleção das culturas enriquecidas evidenciou a presença da classe Gammaproteobacteria e dos filos Firmicutes e Actinobacteria, já que esses foram originalmente encontrados em menor proporção no microbioma. Os filos Proteobacteria, Firmicutes e Actinobacteria são frequentemente descritos em estudos de bactérias resistentes ao arsênio, sendo os gêneros *Bacillus*, *Pseudoxanthomonas* e *Pseudomonas* os mais comuns, independente do ambiente estudado, indicando claro favorecimento das culturas enriquecidas por membros desses táxons (Pepi et al., 2007; Fan et al., 2008; Bachate et al., 2009; Escalante et al., 2009; Liao et al., 2011). Além disso, estudos sugerem uma correlação positiva

entre resistência ao As e a abundância dos filos Firmicutes e Actinobacteria, o que poderia explicar o favorecimento desses filos nas culturas enriquecidas (Gremion et al., 2003; Halter et al., 2011; Vishnivetskaya et al., 2011; Reis et al., 2013).

Embora Bacteroidetes tenha sido o segundo filo mais abundante no microbioma, ele não foi detectado dentre os isolados. Uma possível explicação para este resultado é o fato de que membros desse filo serem raramente encontrados em ambientes contaminados por arsênio (Halter et al., 2011). Em contraste, este filo é frequentemente encontrado por outros pesquisadores em ambientes eutrofizados, ricos em matéria orgânica, como o do sedimento do Córrego da Mina (Sinkko et al., 2013).

A maioria dos gêneros detectados dentre os isolados, como, principalmente, *Bacillus* e *Pseudoxanthomonas*, apresentou baixa abundância no microbioma, indicando que gêneros raros na comunidade bacteriana do Córrego da Mina foram favorecidos pela cultura enriquecida. Por outro lado, os gêneros *Acinetobacter*, *Acidovorax* e *Thermomonas* foram abundantes no microbioma. Assim, as culturas enriquecidas não parecem refletir as condições ambientais do sedimento (nutrientes, concentrações do As e presença de outros metais). Portanto, era esperado que a cultura enriquecida favorecesse táxons normalmente cultivados em laboratório, com crescimento rápido e melhor adaptação às condições da cultura (Bachate et al., 2009). Essa seleção não revelou a maioria da comunidade bacteriana do Córrego da Mina, incluindo os gêneros dominantes que necessitavam de condições especiais de cultivo, ausentes na cultura enriquecida. Isso condiz com o conhecimento atual das comunidades bacterianas ambientais, ou seja: apenas <1% dos microrganismos são cultiváveis em laboratório (Hugenholtz et al, 1998).

Essa discrepância entre os isolados bacterianos e o microbioma também foi observado na análise molecular dos genes de resistência ao As. Os genes *arsC*, *arrA* e *aioA* foram avaliados tanto nos isolados como nas bibliotecas de clones. O gene de oxidação de AsIII, *aioA*, não foi detectado no metagenoma do sedimento e não foi muito diverso nos isolados (presente em apenas 20%). Apesar de ter sido detectado na biblioteca de clones, esse gene apresentou baixa diversidade

sendo as sequências afiliadas apenas às classes Alfa e Betaproteobacteria. Resultado similar foi encontrado também por outros pesquisadores (Quéméneur et al., 2008; Tomczyk-Żak et al., 2013). O AsIII é bastante solúvel sendo pouco encontrado em sedimentos, ao contrário do AsV, que por sua facilidade em precipitar junto a outros metais é mais estável no sedimento (Oremland & Stoltz, 2005). Portanto, nossos dados sugerem que a oxidação de AsIII é menor no sedimento do Córrego da Mina.

Apesar de *arsC* não ter sido detectado no metagenoma, os outros genes do operon *ars* foram identificados, indicando que a ausência do gene *arsC* poderia estar relacionada com falha na montagem das sequências em *contigs* do metagenoma. Os *contigs* apresentam vantagens para análise funcional, pois são maiores, contendo mais informações para a designação funcional, mas excluem as sequências únicas que apresentam potencial para trazer novas informações sobre a comunidade (Desai et al, 2012).

Dos três genes, apenas *arrA* foi detectado e abundante no microbioma, indicando que a redução dissimilatória do AsV é a transformação prevalente do As no sedimento do Córrego da Mina. Esta conclusão pode ser reforçada pela análise metabólica da comunidade pelo Biolog que detectou uma maior adaptabilidade e diversidade em anaerobiose. Segundo Fennel et al. (2009), o sedimento possui duas camadas, sendo a primeira aeróbica localizada diretamente abaixo da interface sedimento-água e a segunda camada anaeróbica localizada logo abaixo da primeira. Esse ambiente de anaerobiose e altas concentrações de AsV presentes no sedimento, devido ao fato desse elemento se precipitar mais facilmente, favorecem bactérias capazes de realizar a redução dissimilatória de AsV (Oremland & Stoltz, 2005). Além disso, Lear et al. (2007) identificaram associação da entrada de carbono no ambiente, o que acontece em ambientes eutrofizados, como o Córrego da Mina, com a prevalência de microrganimos que efetuam a redução dissimilatória de AsV para AsIII. O gene *arrA* foi detectado em poucos isolados provavelmente devido às condições aeróbicas da cultura enriquecida, uma vez que esse gene age na respiração anaeróbica. Além disso, as sequências do gene *arrA* das bibliotecas de clones foram relacionadas a bactérias

não cultivadas, confirmando que esse gene é expresso por bactérias resistentes ao As, ainda não identificadas, como também descrito por Song et al. (2009).

6 – Conclusão Geral

Neste estudo relatou-se uma grande diversidade de gêneros bacterianos redutores de AsV e oxidadores de AsIII, sendo *Thermomonas* e *Pannonibacter* descritos pela primeira vez como gêneros transformadores de arsênio. Esses resultados podem abrir oportunidades para o uso dessas bactérias na biorremediação de ambiente impactado por este metaloide. A análise taxonômica dos isolados e do microbioma revelou perfis taxonômicos distintos. O sedimento do Córrego da Mina abriga uma comunidade complexa com predominância dos filos Proteobacteria e Bacteroidetes. Análise filogenética molecular revelou alta similaridade entre algumas sequências de *arsC* and *aioA* obtidas de isloados e bibiotecas de clones, sugerindo que estes isolados podem representar bactérias ambientalmente importantes na especiação de As. O gene *arrA* foi detectado apenas no microbioma, indicando predominância no processo de redução de AsV, conhecido como redução dissimilatória. Destaca-se, ainda, que o gene *arrA* exibiu maior diversidade molecular. A análise funcional do microbioma revelou alta diversidade de subsistemas, incluindo aqueles relacionados com elementos móveis e resposta ao estresse, bem como de diversos genes de resistência a metais, evidenciando uma comunidade metabolicamente diversa com alta capacidade de adaptação.

7 – Referências Bibliográficas

Ahmann D, Roberts AL, Krumholz LR & Morel FM (1994) Microbe grows by reducing arsenic. Nature 371:750.

Anderson CR & Cook GM (2004) Isolation and characterization of arsenate-reducing bacteria from arsenic-contaminated sites in New Zealand. Curr Microbiol 48:341-347.

Ávila JLT & Monte-Mór RLM (2007) Urbanização e Impactos ambientais: Uma análise da relação entre as características dos espaços urbanos e a população hídrica na região do médio Rio Doce (MG). VII Encontro da Sociedade Brasileira de Economia Ecológica.

Bachate SP, Cavalca L & Andreoni V (2009) Arsenic-resistant bacteria isolated from agricultural soils of Bangladesh and characterization of arsenate-reducing strains. J Appl Microbiol 107:145-56.

Barragner-Bigot P (2004) Contribución al estudio de cinco zonas contaminadas naturalmente por Arsénico en Nicaragua. *In:* Morega O, UNICEF report. UNICEF Publication, Managua, Nicaragua.

Bentley R & Chasteen TG (2002) Microbial methylation of metalloids: arsenic, antimony, and bismuth. Microbiol Mol Biol Rev 66:250-71.

Berg M, Tran HC, Nguyen TC, Pham HV, Schertenleib R & Giger W (2001) Arsenic contamination of groundwater and drinkingwater in Vietnam: a human health threat. Environ Sci Technol 35: 2621–6.

Berg M, Stengel C, Pham TK, Pham HV, Sampson ML, Leng M, Samreth S & Fredericks D (2007) Magnitude of arsenic pollution in the Mekong and Red River Deltas - Cambodia and Vietnam. Sci Total Environ 372: 413-25. 84

Bhattacharya P, Tandukar N, Neku A, Valero AA, Mukherjee AB & Jacks G (2003) Geogenic arsenic in groundwaters from Terai alluvial plain of Nepal. J Physique IV 107:173-6.

Bhattacharya P, Claesson M, Bundschuh J, Sracek O, Fagerberg J, Jacks G, Martin RA, Storniolo Adel R & Thir JM (2007) Distribution and mobility of arsenic in the Rio Dulce alluvial aquifers in Santiago del Estero Province, Argentina. Sci Total Environ 358:97-120.

Bhattacharya P, Welch AH, Stollenwerk KG, Mclaughlin MJ, Bundschuh J & Panaullah G (2007) Arsenic in the environment : Biology and Chemistry. Sci Total Environ 379:109–120.

Borba RP, Figueiredo BR, Rawlins BG & Matchullat J (2000) Arsenic in water and sediment in the Iron Quadrangle, Minas Gerais state, Brasil. Revista Brasileira de Geociências 30**:**554-557.

Borba RP, Figueiredo BR & Matschullat J (2003) Geochemical distribution of arsenic in waters, sediments and weathered gold mineralized rocks from Iron Quadrangle, Brazil, Environ. Geol 44: 39–52.

Borba RP, Figueiredo BR & Cavalcanti JA (2004) Arsênio na água subterrânea em Ouro Preto e Mariana, Quadrilátero Ferrífero (MG). R. Esc. Minas 57: 45-51.

Bundschuh J, Litter MI, Parvez F, Román-Ross G, Nicolli HB, Jean JS, Liu CW, López, Armienta MA, Guilherme LR, Cuevas AG, Cornejo L, Cumbal L & Toujaguez R (2012) One century of arsenic exposure in Latin America: a review of history and occurrence from 14 countries. Sci Total Environ. 1;429:2-35.

Cai L, Liu G, Rensing C & Wang G (2009) Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. BMC Microbiol 9:4.

Cavalca L, Zanchi R, Corsini A, Colombo M, Romagnoli C, Canzi E & Andreoni V (2010) Arsenic-resistant bacteria associated with roots of the wild *Cirsium arvense* (L.) plant from an arsenic polluted soil, and screening of potential plant growth-promoting characteristics. Syst Appl Microbiol 33:154-64.

Chang JS, Yoon IH, Lee JH, Kim KR, An J & Kim KW (2008) Arsenic detoxification potential of aox genes in arsenite oxidizing bacteria isolated from natural and constructed wetlands in the Republic of Korea. Environ Geochem Health 32:95-105.

Chistoserdova L (2010) Recent progress and new challenges in metagenomics for biotechnology. Biotechnol Lett 32:1351-9.

Conselho Nacional do Meio Ambiente (CONAMA) (2011) Resolução No 430, de 13 de maio de 2011. Disponível em: http://www.mma.gov.br/port/conama/estr.cfm.

Costa AT, Nalini HA, Lena JC, Friese K & Mages M (2003) Surface water quality and sediment geochemistry in the Gualaxo do Norte basin, eastern Quadrilátero Ferrífero, Minas Gerais, Brazil. Environ Geol 45:226–235. 86

Daus B, Wennrich R, Morgenstern P, Weiß H, Palmieri HEL, Nalini HA, Leonel LV, Monteiro RPG & Moreira RM (2005) Arsenic Speciation in Plant Samples from the Iron Quadrangle, Minas Gerais, Brazil. Microchim Acta 151: 175–180.

Desai N, Antonopoulos D, Gilbert JA, Glass EM & Meyer F (2012) From genomics to metagenomics. Curr Opin Biotechnol 23:72-6.

Deschamps E, Ciminelli VST, Raue B, Lange FT & Schmidt H (2002) Soil and Sediment Geochemistry of the Iron Quadrangle, Brazil. Journal of Soils & Sediments 2:216-222.

Drewniak L, Styczek A, Majder-Lopatka M & Sklodowska A (2008) Bacteria, hypertolerant to arsenic in the rocks of an ancient gold mine, and their potential role in dissemination of arsenic pollution. Environ Pollut 156:1069-74.

Edwards RA & Dinsdale EA (2007) Marine environmental genomics: unlocking the ocean's secrets. Oceanography 20:26–61

Escalante G, Campos VL, Valenzuela C, Yañez J, Zaror C & Mondaca MA (2009) Arsenic resistant bacteria isolated from arsenic contaminated river in the Atacama Desert (Chile). Bull Environ Contam Toxicol 83:657-61.

Fan H, Su C, Wang Y, Yao J, Zhao K, Wang Y & Wang G (2008) Sedimentary arsenite-oxidizing and arsenate-reducing bacteria associated with high arsenic groundwater from Shanyin, Northwestern China. J Appl Microbiol 105:529-39.

Fennel K, Brady D, DiToro D, Fulweiler RW, Gardner WS, Giblin A, et al. (2008) Modeling denitrification in aquatic sediments. Biogeochemistry 93:159–178.

Foley ME, Sigler V & Gruden CL (2008) A multiphasic characterization of the impact of the herbicide acetochlor on freshwater bacterial communities. ISME Journal 2:56–66.

Garelick H & Jones H (2008) Mitigating Arsenic Pollution: Bridging the Gap Between Knowledge and Practice. Disponível em: www.iupac.org/publications/ci/2008/3004/2_garelick.html.

Garland JL (1996) Analysis and interpretation of community-level physiological profiles in microbial ecology. FEMS Microbiol Ecol 24:289-300

Gilbert JA & Dupont CL (2011) Microbial metagenomics: beyond the genome. Ann Rev Mar Sci 3:347-71.

Gremion F, Chatzinotas A & Harms H (2003) Comparative 16S rDNA and 16S rRNA sequence analysis indicates that Actinobacteria might be a dominant part of the metabolically active bacteria in heavy metal-contaminated bulk and rhizosphere soil. Environ Microbiol 5: 896-907.

Gupta P & Vakhlu J (2011) Metagenomics: a quantum jump from bacterial genomics. Indian J Microbiol 51:539-41.

Halter D, Cordi A, Gribaldo S, Goulhen-chollet F, Heinrich-salmeron A, Carapito C, et al. (2011) Taxonomic and functional prokaryote diversity in mildly arsenic-contaminated sediments. Res Microbiol 162:877-887.

Hamamura N, Macur RE, Korf S, Ackerman G, Taylor WP, Kozubal M, Reysenbach AL & Inskeep WP (2009) Linking microbial oxidation of arsenic with detection and phylogenetic analysis of arsenite oxidase genes in diverse geothermal environments. Environ Microbiol 11:421-31. 88

Handelsman J, Rondon MR, Brady SF, Clardy J & Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 5:245-249.

Handelsman J (2004) Metagenomics: Application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev 68:669–685.

Hugenholtz P, Goebel BM & Pace NR (1998) Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. J Bacteriol 180: 4765-4774.

Instituto Mineiro de Gestão das Águas (IGAM) (2012) Monitoramento da qualidade das águas superficiais no estado de Minas Gerais - Relatório trimestral. Instituto Mineiro de Gestão das Águas, Gerência de Monitoramento Hidrometeorológico.

Jones W (2010) High-Throughput Sequencing and Metagenomics. Estuaries and Coasts Volume 33: 944-952.

Jünemann S, Prior K, Szczepanowski R, Harks I, Ehmke B, Goesmann A, Stoye J & Harmsen D (2012) Bacterial community shift in treated periodontitis patients revealed by ion torrent 16S rRNA gene amplicon sequencing. PLoS One 7:e41606.

Kaur S, Kamli MR & Ali A (2009) Diversity of arsenate reductase genes (*arsC* Genes) from arsenic-resistant environmental isolates of *E. coli*. Curr Microbiol 59:288-94. 89

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M & Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41:e1.

Koch R (1880) Investigations into the etiology of traumatic infective diseases. New Sydenham Society, London.

Kruger MC, Bertin PN, Heipieper HJ & Arsène-Ploetze F (2013) Bacterial metabolism of environmental arsenic-mechanisms and biotechnological applications. App Microbiol Biotechnol 97: 3827-41.

Lear G, Song B, Gault G, Polya D & Lloyd JR (2007) Molecular analysis of arsenate-reducing bacteria within Cambodian sediments following amendment with acetate. Appl Environ Microbiol 73:1041–8.

Lett MC, Muller D, Liévremont D, Silver S & Santini J (2012) Unified nomenclature for genes involved in prokaryotic aerobic arsenite oxidation. J Bacteriol 194:207–8.

Liao VHC, Chu YJ, Su YC, Hsiao SY, Wei CC, Liu CW, Liao CM, Shen WC & Chang FJ (2011) Arsenite-oxidizing and arsenate-reducing bacteria associated with arsenic-rich groundwater in Taiwan. J Contam Hydrol 123:20-9.

Lièvremont D, Bertin PN & Lett MC (2009) Arsenic in contaminated waters: Biogeochemical cycle, microbial metabolism and biotreatment processes. Biochimie 91:1229–1237.

Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. (2012) Comparison of next-generation sequencing systems. Journal of biomedicine & biotechnology 2012: 1-11.

Lord S (2001) A Ética do Uso da Água Doce: um levantamento. Brasília: UNESCO, 80p.

Madigan MT, Martinko JM & Parker J (2004) Microbiologia de Brock. São Paulo: Prentice Hall, 10ª ed.

Malasarn D, Saltikov CW, Campbell KM, Santini JM, Hering JG & Newman DK (2004) *arrA* Is a Reliable Marker for As(V) Respiration. Science 36:455.

Mardis ER (2008) Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 9:387-402.

Matschullat J (2000) Arsenic in the geosphere-a review. Sci Total Environ 249:297-312.

Matschullat J, Birmann K, Borba RP, Ciminelli V, Deschamps EM, Figueiredo BR, Gabrio T, Haßler S, Hilscher A, Junghänel I, de Oliveira N, Raßbach K, Schmidt H, Schwenk M, Vilhena MJO & Weidner U (2007) Long-term environmental impact of arsenic-dispersion in Minas Gerais, Brazil. Trace Metals and other Contaminants in the Environment 9:365-382

McClintock TR, Chen Y, Bundschuh J, Oliver JT, Navoni J, Olmos V, et al. (2012) Arsenic exposure in Latin America: biomarkers, risk assessments and related health effects. Sci Total Environ 429:76–91.

Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S & Rappuoli R (2008) Microbiology in the post-genomic era. Nat Rev Microbiol 6:419-30.

Mello JWV, Dias LE, Daniel AM, Abrahão WAP, Deschamps E & Schaefer CEGR (2006) Preliminary evaluation of acid mine drainage in Minas Gerais State, Brazil. R. Bras. Ci. Solo 30:365-375.

Moraes DSL & Jordão BQ (2002) Water resources deterioration and its impact on human health. Rev Saúde Pública 36:370-374.

Mullis KB & Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. Methods in Enzymology 155:335-350.

Nascimento AMA (2011) Use of the rRNA operon and genomic repetitive sequences for the identification of bacteria. *In:* Bruijn J, Willey J Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches.

Nordstrom DK (2002) Public health.Worldwide occurrences of arsenic in ground water. Science 296:2143–5.

Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, et al. (2011). Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Res 39:3204–23.

Oliveira A, Pampulha ME, Neto MM & Almeida AC (2009) Enumeration and Characterization of Arsenic-Tolerant Diazotrophic Bacteria in a Long-Term Heavy-Metal-Contaminated Soil. Water Air Soil Pollut 200:237-243.

Oremland RS & Stolz JF (2003) The ecology of arsenic. Science 300:939-944.

Oremland RS & Stolz JF (2005) Arsenic, microbes and contaminated aquifers. TRENDS Microbiol 13:45-49.

Oremland RS, Saltikov CW, Wolfe-simon F & Stolz JF (2009) Arsenic in the evolution of earth and extraterrestrial ecosystems 26:522–536.

Pace NR, Stahl DA, Lane DJ & Olsen GJ (1985) The analysis of natural microbial populations by ribosomal RNA sequences. Adv Microb Ecol 9:1-55.

Páez-Espino D, Tamames J, de Lorenzo V & Cánovas D (2009) Microbial responses to environmental arsenic. Biometals 22:117–130. 92

Peixoto BM (2011) Bioinformática aplicada a um projeto de metagenômica. Disponível em: http://www.ic.unicamp.br/~zanoni/orientacoes/mestrado/bruno/PropostaBruno.pdf.

Pepi M, Volterrani M, Renzi M, Marvasi M, Gasperini S, Franchi E & Focardi SE (2007) Arsenic-resistant bacteria isolated from contaminated sediments of the Orbetello Lagoon, Italy, and their characterization. J Appl Microbiol 103:2299-308.

Pompêo MLM, Moschini-Carlos V, Alexandre NZ & Santo E (2004) Qualidade da água em região alterada pela mineração de carvão na microbacia do rio Fiorita (Siderópolis, Estado de Santa Catarina, Brasil). Acta Scientiarum. Biological Sciences 26: 125-136.

Pontes DS, Lima-Bittencourt CI, Chartone-Souza E. & Nascimento AMA (2007) Molecular approaches: advantages and artifacts in assessing bacterial diversity. J Ind Microbiol Biotechnol 120: 95-99.

Quéméneur M, Heinrich-Salmeron A, Muller D, Lièvremont D, Jauzein M, et al. (2008) Diversity surveys and evolutionary relationships of aoxB genes in aerobic arsenite-oxidizing bacteria. Appl Environ Microbiol 74: 4567–73.

Rabadjieva D, Tepavitcharova S, Todorov T, Dassenakis M, Paraskevopoulou V & Petrov M (2008) Chemical speciation in mining affected waters: the case study of Asarel-Medet mine. Environ Monit Assess 159:353-66

Rappé MS & Giovannoni SJ (2003) The uncultured microbial majority. Annu Rev Microbiol 57:369-394.

Reis MP, Barbosa FA, Chartone-Souza E & Nascimento AMA (2013) The prokaryotic community of a historically mining-impacted tropical stream sediment is as diverse as that from a pristine stream sediment. Extremophiles 17: 301–309.

Reyes C, Lloyd JR & Saltikov CW (2008) Geomicrobiology of iron and arsenic in anoxic sediments. In: Arsenic contamination of groundwater: Mechanism, analysis, and remediation. Editado por Ahuja S. John Wiley & Sons, Inc.

Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, et al. (2013) Insights into the phylogeny and coding potential of microbial dark matter. Nature 499:431–7.

Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne J, Clardy MS, Handelsman J & Goodman RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl Environ Microbiol 66:2541-2547.

Rudi K, Zimonja M, Trosvik P & Naes T (2007) Use of multivariate statistics for 16S rRNA gene analysis of microbial communities. Int J Food Microbiol 120: 95−99

Sanger F, Nicklen S & Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 74:5463-5467.

Schwarzenbach RP, Egli T, Hofstetter TB, von Gunten U & Wehrli B (2010) Global water pollution and human health. Ann Rev Environ Resour 35:109-136.

Shapiro JA & Dworkin M (1997) Bacteria as multicellular organisms. Oxford University Press, New York, USA.

Silver S & Phung LT (2005) Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. Appl Environ Microbiol 71:599-608.

Sinkko H, Lukkari K, Sihvonen LM, Sivonen K, Leivuori M, Rantanen M, et al. (2013). Bacteria contribute to sediment nutrient release and reflect progressed eutrophication-driven hypoxia in an organic-rich continental sea. PloS one 8:e67061.

Slyemi D & Bonnefoy V (2012) How prokaryotes deal with arsenic. Environ Microbiol Rep 4: 571–86.

Smedley PL & Kinniburgh DG (2002) A review of the source, behavior and distribution of arsenic in natural waters. Appl Geochem 17:517-68.

Smedley PL, Nicolli HB, Macdonald DMJ, Barros AJ & Tullio JO (2002) Hydrochemistry of arsenic and other inorganic constituents in groundwater from La Pampa, Argentina. Appl Geochem 17:259-84.

Smedley PL, Kinniburgh DG, Macdonald DMJ, Nicolli HB, Barros AJ, Tullio JO, Pearce JM & Alonso MS (2005) Arsenic associations in sediments from the loess aquifer of La Pampa, Argentina. Appl Geochem 20: 989-1016.

Smith AH, Marshall G, Yuan Y, Ferreccio C, Liaw J, von Ehrenstein O, Steinmaus C, Bates MN & Selvin S (2006) Increased mortality from lung cancer and bronchiectasis in young adults after exposure to arsenic in utero and in early childhood. Environ Health Perspect 114:1293–6.

Song B, Chyun E, Jaffé PR & Ward BB (2009) Molecular methods to detect and monitor dissimilatory arsenate-respiring bacteria (DARB) in sediments. FEMS microbiology ecology 68:108–17.

Stolz JF, Basu P & Oremland RS (2002) Microbial transformation of elements: the case of arsenic and selenium. Int Microbiol 5:201-7.

Stolz JF, Basu P & Oremland RS (2010) Microbial Arsenic Metabolism: New Twists on an Old Poison. Microbe 5:53-59.

Suenaga H (2012) Targeted metagenomics: a high-resolution metagenomics approach for specific gene clusters in complex microbial communities. Environ Microbiol 14:13-22.

Tomczyk-Żak K, Kaczanowski S, Drewniak L, Dmoch L, Sklodowska A & Zielenkiewicz U (2013) Bacteria diversity and arsenic mobilization in rock biofilm from an ancient gold and arsenic mine. Sci Total Environ, 461-462, 330–40.

Tourova TP (2003) Copy Number of Ribosomal Operons in Prokaryotes and Its Effect on Phylogenetic Analyses. Microbiology 72:389–402. 95

Torsvik V, Ovreas L & Thingstad TF (2002) Prokaryotic diversity-magnitude, dynamics, and controlling factors. Science 296:1064–1066.

Tringe SG & Hugenholtz P (2008) A renaissance for the pioneering 16S rRNA gene. Curr Opin Microbiol 11: 442–446.

Tsai SL, Singh S & Chen W (2009) Arsenic metabolism by microbes in nature and the impact on arsenic remediation. Curr Opin Biotechnol 20:659-67.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH & Smith HO (2004) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304:66–74.

Whitman WB, Coleman DC & Wiebe WJ (1998) Prokaryotes: the unseen majority. Proc Natl Acad Sci USA 95:6578-83.

Vishnivetskaya TA, Mosher JJ, Palumbo AV, Yang ZK, Podar M, Brown SD, et al. (2011) Mercury and other heavy metals influence bacterial community structure in contaminated Tennessee streams. Appl Environ Microbiol 77: 302-311.

Woese CR (1987) Bacterial Evolution. Microbiol Rev 51: 221-271.

Xu J (2011) Microbial ecology in the age of metagenomics. *In:* Bruijn J, Willey J Handbook of Molecular Microbial Ecology I: Metagenomics and Complementary Approaches.

Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, et al. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microb 12:635–645.

Zargar K, Hoeft S, Oremland R & Saltikov CW (2010) Identification of a novel arsenite oxidase gene, *arxA*, in the haloalkaliphilic, arsenite-oxidizing bacterium *Alkalilimnicola ehrlichii* strain MLHE-1. J Bacteriol 192: 3755–62.

Zargar K, Conrad A, Bernick DL, Lowe TM, Stolc V, et al. (2012) ArxA, a new clade of arsenite oxidase within the DMSO reductase family of molybdenum oxidoreductases. Environ Microbiol 14: 1635–45.