

Thèse présentée pour obtenir le grade de
Docteur de l'Université de Strasbourg

Discipline Sciences du Vivant
Aspects moléculaires et cellulaires de la biologie

Diversité des mécanismes d'oxydo-réduction de l'arsenic chez les bactéries : Des souches cultivables aux communautés

par Audrey Heinrich-Salmeron

Soutenue publiquement le 6 Novembre 2009

Membres du jury

M. Michael DuBow (Université Paris Sud 11) - Rapporteur externe

M. Pierre Peyret (Université Blaise Pascal Clermont-Ferrand) - Rapporteur externe

Mme Valérie Geoffroy (Université de Strasbourg) - Rapporteur interne

Mme Marie-Claire Lett (Université de Strasbourg) - Directrice de thèse

M. Philippe Bertin (Université de Strasbourg) - Invité

Thèse présentée pour obtenir le grade de
Docteur de l'Université de Strasbourg

Discipline Sciences du Vivant
Aspects moléculaires et cellulaires de la biologie

Diversité des mécanismes d'oxydo-réduction de l'arsenic chez les bactéries : Des souches cultivables aux communautés

par Audrey Heinrich-Salmeron

Soutenue publiquement le 6 Novembre 2009

Membres du jury

M. Michael DuBow (Université Paris Sud 11) - Rapporteur externe

M. Pierre Peyret (Université Blaise Pascal Clermont-Ferrand) - Rapporteur externe

Mme Valérie Geoffroy (Université de Strasbourg) - Rapporteur interne

Mme Marie-Claire Lett (Université de Strasbourg) - Directrice de thèse

M. Philippe Bertin (Université de Strasbourg) - Invité

« La science consiste à passer d'un étonnement à un autre »
Aristote

à Arnaud,

REMERCIEMENTS

Ce travail de thèse a été réalisé au sein du laboratoire de Génétique Moléculaire, Génomique et Microbiologie (UMR7156-UdS/CNRS) dirigé par Monsieur le Professeur Serge Potier. Je le remercie de m'avoir accueillie au sein de son unité.

Je tiens à exprimer ma reconnaissance envers Messieurs M. BuDow et P. Peyret et Madame V. Geoffroy qui me font l'honneur de juger ce travail.

Cette thèse a été financée par une allocation du Ministère de l'Enseignement et de la Recherche et par un contrat de Monitorat de l'Université de Strasbourg.

Je tiens à remercier Marie-Claire Lett pour son encadrement durant ma thèse. Pour m'avoir guidée et encouragée dans mon travail, mais aussi pour m'avoir soutenue lors de mon monitorat.

Je voudrais adresser ma gratitude à Philippe Bertin. Merci pour tes conseils et pour la part importante que tu as prise dans ce travail. Merci pour ta confiance, pour la génomique et pour tout le chocolat...

Merci à Didier de m'avoir accompagnée lors de mon stage de Master deuxième année, merci pour ta disponibilité et tes conseils.

Merci aux Flo pour votre aide, votre implication et votre disponibilité. Merci de m'avoir initiée à la protéomique. Merci à FloC pour ton amitié, rendez-vous dans 10 ans...

Un grand merci aux Arseniquettes (Jessica et Marie) et aux Arseniqueurs (David et François). Un merci particulier à la plus tactile des Arseniquettes : merci à toi Sandrine pour tous tes conseils, ton aide précieuse, ta disponibilité, ta gentillesse, ta bonne humeur... Merci de votre amitié, pour toutes nos conversations et pour toutes les heures d'ouverture du bureau des renseignements et des ragots ! Merci pour tous ces bons moments d'amitié et d'émulation.

Je n'oublie pas toutes les personnes du laboratoire qui ont su m'aider dans la réussite de ce travail et tous ceux qui ont rendu ces quatre années conviviales et agréables. Merci à Emilie (longue vie à Aloïs, à nos projets futurs, à notre belle amitié), Nic, Paul, Christelle,... ainsi qu'à tous mes stagiaires : Annie, Jamie, Olivier, Mathieu et François.

Je souhaite aussi exprimer toute ma reconnaissance à Pascaline Ullmann pour avoir été mon tuteur pédagogique lors de mon monitorat. Merci à toi Pascaline de m'avoir transmis ton goût pour l'enseignement et merci pour tous les très bons moments passés en TP.

Merci à toutes les personnes qui ont fructueusement collaboré à ce travail. Merci en particulier à Claudine Médigue, à toute l'équipe AGC et au Génoscope.

Une pensée émue à mes parents, merci de votre soutien inconditionnel. Merci à toi papa de m'avoir transmis la curiosité et l'éblouissement face à la magie des choses. Merci à toi maman de m'avoir donné le goût d'apprendre et de ne jamais abandonner.

Merci à ma famille, à ma grand-mère le « hobbit », Alicia, Alex et Marie-Françoise (ma fidèle relectrice), pour leur soutien et pour n'avoir jamais douté.

Merci à tous mes chers amis de pimenter la vie de nombreux moments d'amitié. Une pensée particulière à Emeline ma Partenaire Docteur !

Parce que ce travail est aussi en partie le tien, merci à toi Arnaud, d'être mon exhausteur de réussite et de bonheur.

SOMMAIRE

Liste des tables et figures	13
Liste des abreviations	15
INTRODUCTION	17
CHAPITRE I - L'ARSENIC DANS TOUS SES ETATS	19
I Géochimie	21
1 Propriétés physico-chimiques	21
2 Mobilisation de l'arsenic	23
II Toxicité	24
1 Au niveau cellulaire	24
2 Toxicité chez l'homme	26
<i>a. Symptômes et maladies</i>	26
<i>b. L'arsenic dans le corps humain</i>	27
<i>c. La toxicité en quelques chiffres</i>	28
<i>d. Traitements</i>	28
3 Utilisation de l'arsenic en thérapie	29
III L'arsenic dans l'environnement	30
1 Répartition dans l'environnement	30
2 La répartition géographique des zones contaminées	31
3 Sources de la contamination à l'arsenic	32
<i>a. Origine naturelle</i>	32
<i>b. Origine anthropogénique</i>	32
4 Législation	33
5 La remédiation de l'arsenic	34
CHAPITRE II - LES ROLES DES BACTERIES DANS LE CYCLE BIOGEOCHIMIQUE DE L'ARSENIC	37
I Les transformations bactériennes et leurs rôles dans le cycle biogéochimique de l'arsenic	39
II L'oxydation de l'arsénite en arséniate	40
1 Phylogénie et diversité	41
2 L'arsénite-oxydase	42
<i>a. L'enzyme</i>	42
<i>b. Les sous-unités</i>	43

c.	Le <i>cytochrome</i>	44
3	Le transport de l'arsénite-oxydase	44
4	Le fonctionnement de l'arsénite-oxydase	44
5	Les gènes <i>aox</i>	46
6	La régulation transcriptionnelle et post-transcriptionnelle de l'arsénite-oxydase	47
7	Arsénite-oxydase et milieu originel	49
8	Conclusions	49
III	Reduction de l'arséniate en arsénite	50
1	La réduction cytoplasmique de l'AsV et le système d'efflux	50
a.	<i>ArsC : La réductase</i>	50
b.	<i>ArsB/Acr3 : La pompe d'efflux</i>	51
c.	<i>ArsA : l'ATPase</i>	53
d.	<i>ArsRD : les protéines régulatrices</i>	54
e.	<i>Quel rôle pour ArsH ?</i>	55
f.	<i>L'opéron ars</i>	55
g.	<i>Evolution des gènes ars</i>	56
2	La réduction dissimilatrice	56
a.	<i>Les souches</i>	57
b.	<i>La protéine</i>	57
c.	<i>Reversibilité des fonctions de la protéine ArrB</i>	60
IV	Methylation	60
V	Sequestration	62
VI	Un exemple particulier : le rôle des bactéries dans les drainages miniers acides (AMD)	64
1	Origine et chimie	64
2	Les micro-organismes dans les drainages miniers acides	66
a.	<i>Les souches bactériennes présentes</i>	66
b.	<i>Le rôle de ces procaryotes dans la physico-chimie des drainages miniers acides</i>	68
3	Un exemple : la rivière le Reigous à Carnoulès, Gard (France)	71
4	Conclusions	73
CHAPITRE III - L'UTILISATION DES OUTILS « META-OMICS » DANS L'EXPLORATION D'UNE COMMUNAUTE MICROBIENNE COMPLEXE		75

I Introduction	77
II Métagénomique	79
1 Principe	80
2 L'analyse des pseudo-génomés	82
III Métatranscriptomique	84
IV Métaproteomique	85
V Conclusions	88
CHAPITRE IV - OBJECTIFS DES TRAVAUX	89
TRAVAUX PERSONNELS	93
CHAPITRE I - CARACTERISATION DE BACTERIES ARSENITE-OXYDANTES	95
Première partie - étude d'une souche arsénite-oxydante <i>Pseudomonas</i> S11	97
I Introduction	97
II Manuscrit	99
Deuxième partie - étude d'une souche arsénite-oxydante <i>Leptothrix</i> <i>arsenicoxydans</i>	123
I Introduction	123
II Manuscrit	125
Troisième partie - génomique descriptive et comparative d'une souche arsénite-oxydante <i>Thiomonas</i> 3As	147
I Introduction	147
II Manuscrit	149
CHAPITRE II - ETUDE DE LA DIVERSITE ET L'EVOLUTION DE GENES <i>AOXB</i> CHEZ DES BACTERIES	189
I Introduction	191
II Manuscrit	192
CHAPITRE III - ETUDE DE LA DIVERSITE DES GENES <i>AOXB</i> DANS LA COMMUNAUTE BACTERIENNE DE L'ANCIENNE MINE DE SAINTE-MARIE-AUX- MINES	203
I Introduction	205
II Manuscrit	207

CHAPITRE IV - ETUDE D'UNE COMMUNAUTE COMPLEXE ISSUE D'UN AMD PAR DES APPROCHES PROTEO-GENOMIQUES	237
I Introduction	239
II Manuscrit	241
III Résultats complémentaires	271
1 L'opéron <i>aox</i> dans le métagénome	271
2 Analyses des régions <i>aox</i> de Carn2	274
<i>a. Les régions aox dans le pseudo-génome Carn2</i>	274
<i>b. L'opéron aox et ses gènes associés</i>	275
<i>c. Les gènes impliqués dans la synthèse du cofacteur à molybdène</i>	277
<i>d. Les gènes ars</i>	278
3 Conclusions	279
CONCLUSIONS ET PERSPECTIVES	281
MONITORAT	287
ANNEXES	291
ANNEXE 1 - COMMUNICATIONS SCIENTIFIQUES	293
ANNEXE 2 - HOW CLAYS STIMULATE MARINE OIL BIOREMEDIATION	299
BIBLIOGRAPHIE	327

LISTE DES TABLES ET FIGURES

Figure I1.1 :	Schéma des formes +III et +V de l'arsenic et leur pKa.	22
Figure I1.2 :	Diagramme de l' E_h -pH pour l'arsenic à 25°C et 1013 hPa (1 atm).	22
Figure I1.3 :	Détoxification de l'arsenic dans les cellules procaryotes et eucaryotes.	26
Figure I1.4 :	Diagramme illustrant la relative toxicité des différentes formes d'arsenic.	28
Figure I1.5 :	Cristal d'orpiment, de réalgar, d'énergite, d'arsénopyrite.	31
Figure I1.6 :	Carte du monde montrant les zones connues pour leurs fortes concentrations en arsenic.	32
Table I1.T1 :	Effets potentiels d'un traitement à base d'arsenic sur différents cancers.	29
Figure I2.1 :	Cycle de l'arsenic dans l'environnement.	39
Figure I2.2 :	Diagramme des différents processus microbiens impliqués dans la biochimie de l'arsenic.	40
Figure I2.3 :	Représentation biochimique de l'arsénite-oxydase d' <i>Alcaligenes faecalis</i> .	43
Figure I2.4 :	Représentation schématique du mécanisme d'oxydation de l'arsénite en arséniate par l'arsénite-oxydase.	45
Figure I2.5 :	Modèle de transfert d'électrons de l'arsénite-oxydase lors de l'oxydation de l'As III en As V.	45
Figure I2.6 :	Représentation putative de la chaîne de transport d'électrons pour une oxydation hétérotrophique de l'arsénite.	46
Figure I2.7 :	Organisation des clusters de gènes <i>aox</i> chez <i>H. arsenicoxydans</i> et chez différentes souches métabolisant l'arsenic.	47
Figure I2.8 :	Représentation conceptuelle de l'oxydation de l'arsénite chez <i>H. arsenicoxydans</i> .	48
Figure I2.9 :	Le mécanisme réactionnel d'ArsC.	51
Figure I2.10 :	La pompe ArsAB.	52
Figure I2.11 :	Modèle de l'activation allostérique de ArsA par l'As III.	54
Figure I2.12 :	Organisation des opérons <i>ars</i> chez <i>H. arsenicoxydans</i> .	56
Figure I2.13 :	Modèle de l'arséniate-réductase respiratoire.	58
Figure I2.14 :	Gènes codant pour l'arséniate réductase respiratoire (<i>arrAB</i>) et cytoplasmique (<i>arsABCD</i>) de <i>Shewanella</i> souche ANA-3.	58
Figure I2.15 :	Arbre phylogénétique (non-enraciné neighbor joining) des enzymes à molybdène de la famille DMSO réductase.	59
Figure I2.16 :	Etapes hypothétiques de la méthylation de l'arsenic inorganique.	61
Figure I2.17 :	Synthèse d'EPS par <i>H. arsenicoxydans</i> en présence de 300 ppm d'arsenic.	62
Figure I2.18 :	Photographie de microscopie électronique à transmission d' <i>H. arsenicoxydans</i> .	63
Figure I2.19 :	Photographie de différentes AMDs.	65
Figure I2.20 :	Analyse phylogénétique des gènes 16S rRNA issus d'AMD.	67
Figure I2.21 :	Proposition schématique de développement d'un AMD.	70

Figure I2.22 :	Présentation du site de Carnoulès dans le Gard.	71
Figure I2.23 :	Schéma de la section transversale de la naissance du Reigous sous les stériles miniers jusqu'au point de prélèvement classiques COWG.	72
Table I2.T1 :	Noms et abréviations des différentes formes d'arsenic méthylées.	61
Table I2.T2 :	Caractéristiques physico-chimiques d'AMD de divers endroits sur la planète.	64
Figure I3.1 :	Vue d'ensemble des principaux outils de génomique, de protéomique, de métagénomique et de métaprotéomique.	78
Figure I3.2 :	Représentation schématique de la construction et du criblage de la banque métagénomique issue d'un échantillon environnemental.	81
Figure I3.3 :	Représentation de la plateforme d'annotation de séquence MaGe développée par le Génoscope.	82
Figure I3.4 :	Représentation schématique de la métatranscriptomique.	85
Figure I3.5 :	Schéma synthétique de l'analyse métaprotéomique.	87
Figure 4.1 :	Représentation schématique de cinq pseudo-génomés reconstruits à partir des données de métagénomique et leurs interactions.	272
Figure 4.2 :	Représentation synthétique des opérons <i>aox</i> du pseudo-génome 2 et de leurs environnements génétiques.	273
Figure 4.3 :	Carte génomique circulaire du pseudo-génome Carn 2.	275
Figure 4.4 :	Organisation synténique des opérons <i>aox</i> chez <i>Herminiimonas arsenicoxidans</i> , <i>Thiomonas</i> 3As et dans les trois régions du pseudo-génome Carn 2.	276
Figure 4.5 :	Proposition de représentation schématique de la synthèse du cœur molybdoptérine et représentation schématique et du transfert d'électrons de l'arsénite-oxydase.	278

LISTE DES ABREVIATIONS

ABC	<i>ATP-Binding Cassette</i>
ADN	Acide DésoxyriboNucléique
ADP	Adénosine DiPhosphate
AMD	<i>Acid Mine Drainage</i>
ANR	Agence Nationale de la Recherche
ARN	Acides RiboNucléique
ATP	Adénosine TriPhosphate
BRGM	Bureau des Recherche Géologiques et Minières
bp	paire(s) de base
CDS	<i>CoDing Sequence</i>
CGH	<i>Comparative Genomic Hybridization</i>
DMA	<i>DiMethylarsine Acid</i>
DMSO	DiMéthylSulfOxyde
EPS	ExoPolySaccharide
GDR	Groupement De Recherche
GSH	Glutathion
Gb	Giga bases
HPLC	<i>high performance liquid chromatography</i>
kpb	kilo paire(s) de base
LUCA	<i>Last Universal Common Ancestor</i>
MMA	<i>MonoMethylarsine Acid</i>
MBS	<i>Metal Binding Site</i>
Mb	Mega base
MS	<i>Mass Spectrometry</i>
NBS	<i>Nucleotide Binding Site</i>
OMS	Organisation Mondiale de la Santé
ORF	<i>Open Reading Frame</i>
PCR	<i>Polymerase Chain Reaction</i>
RARE	<i>Reactivity of an Arsenic-Rich Ecosystem</i>
TAT	<i>Twin ArginineTransporter</i>

Introduction

Chapitre 1

L'arsenic dans tous ses états

Une réputation bien acquise

L'arsenic est connu depuis l'antiquité comme un puissant poison, ce qui lui vaut la dénomination de « poudre d'héritage ». Il est selon Martin Villemein (Académicien) « (une) arme adroite pour faire passer l'arme à gauche ». Il reste encore aujourd'hui tristement célèbre comme le responsable d'empoisonnements massifs liés à l'ingestion d'eau contaminée, en particulier au Bangladesh. L'organisation mondiale de la santé décrit cette contamination comme le plus grand empoisonnement de l'histoire de l'Humanité, puisqu'on estime que plus de 80 millions de personnes au Bangladesh seraient exposées et 40 millions au Bengal (Chowdhury *et al.*, 2000) (Vaughan, 2006). La découverte de l'élément lui-même est généralement attribuée à Albertus Magnus en 1250. Cependant, il était déjà bien connu au IV^e siècle avant JC puisque Aristote décrit le réalgar (AsS) et Pline appela l'orpiment *auri pigmentum* en raison de sa belle couleur jaune. Il tient son nom du latin *arsenicum*, issu du grec *arsenikon* : « qui dompte le mâle ».

Toutefois, l'arsenic sut s'attirer d'autres convoitises. En effet, les alchimistes savaient que l'arsenic ajouté au cuivre donnait un métal blanc qu'on assimilait à l'argent. Un pas était alors franchi « d'un métal vil vers un métal noble », il ne restait plus alors qu'à « transmuter » l'argent en or.

I GEOCHIMIE

1. Propriétés physico-chimiques

L'élément arsenic est un membre du groupe V dans la classification périodique de Mendeleïev. Il se situe entre le phosphore et l'antimoine. De symbole chimique As, il possède un numéro atomique de 33, son poids moléculaire est de 74,9 et sa densité de 5,7. Il peut exister sous différents états de valence : l'état métallique As(0), trivalent As(+III) [noté As III] et As(-III) et pentavalent As(+V) [noté As V] et As(-V). C'est un métalloïde, c'est-à-dire que ses propriétés physiques et chimiques sont intermédiaires entre les métaux et les non-métaux, son comportement en solution le rapproche des non-métaux, alors que ses capacités de conduction le rapprochent des métaux.

Dans les milieux aqueux, les formes dominantes sont l'arsénite H_3AsO_3 , As(+III) et l'arséniate $H_2AsO_4^-$, As(+V).

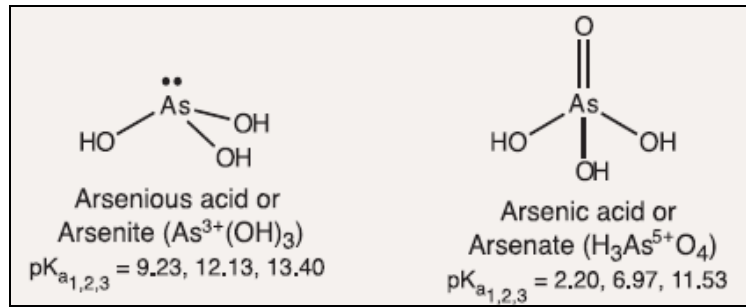


Figure I1.1 : Schéma des formes +III et +V de l'arsenic et leur pKa (extrait de (O'Day, 2006)).

La configuration structurale de l'As III présente trois liaisons pyramidales et une paire d'électrons libres sur le quatrième bras du tétraèdre. L'As V quant à lui possède la même base, mais les deux électrons supérieurs participent à une double liaison avec un atome d'oxygène comme illustré dans la Figure I1.1 (O'Day, 2006).

Le potentiel redox (E_h) et le pH contrôlent la spéciation de l'arsenic dans les environnements naturels (Garelick *et al.*, 2008), ainsi les courbes de distribution des différentes formes d'arsenic en fonction du pH et de l' E_h sont représentées Figure I1.2. Brièvement, dans les milieux aqueux à pH neutre l'As III est l'espèce dominante (Ramírez-Solís *et al.*, 2004) (Sharma & Sohn, 2009). Dans des conditions d'oxydation (valeurs d' E_h élevées) l'arsenic inorganique se trouve sous forme As V, alors que dans les conditions basses d' E_h , c'est-à-dire des conditions réductrices, la forme As III sera principalement retrouvée. Enfin dans des conditions réductrices extrêmes, il peut se former de l'arsine ou de l'arsenic élémentaire, mais ceci ne se produit que rarement dans l'environnement.

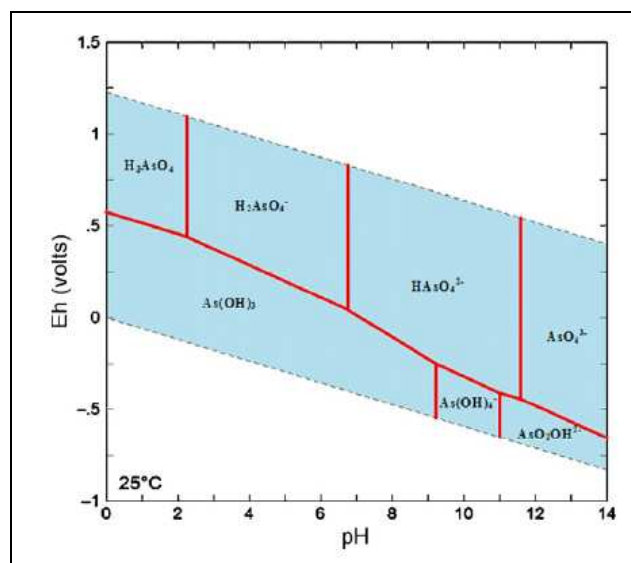


Figure I1.2 : Diagramme de l' E_h -pH pour l'arsenic à 25°C et 1013hPa (1 atm) (Extrait de (Cheng *et al.*, 2009))

Cependant les données du ratio As III/As V retrouvé dans l'environnement ne correspondent que rarement aux valeurs théoriques. En effet, les processus biologiques jouent un rôle prépondérant dans la mobilisation ou l'immobilisation de l'arsenic et affectent donc sa spéciation (Sharma & Sohn, 2009) (Oremland *et al.*, 2005a).

2. Mobilisation de l'arsenic

L'étude des mécanismes de mobilisation/immobilisation de l'arsenic est indispensable en raison de leur implication dans la remédiation de l'arsenic dans les eaux de consommation. Dans les sols, ces mécanismes dépendent du type et de la quantité de composés adsorbants du sol, du pH et du potentiel redox (Bissen & Frimmel, 2003b). Trois processus différents sont en cause dans cette mobilisation/immobilisation dans les aquifères (Lievremont *et al.*, 2009) :

- L'adsorption/désorption due au pH, aux conditions d'oxydoréduction, à la présence d'ions en compétition et aux changements structuraux des phases solides. L'arsénite et l'arséniate s'adsorbent sur des surfaces d'une grande variété de matériaux présents dans les aquifères. Par exemple l'adsorption maximale d'As III apparaît autour du pH 7 alors que l'adsorption de l'As V est indépendante du pH à des concentrations faibles d'arsenic. L'adsorption de l'As V augmente avec la diminution du pH (pH 9-4) à des concentrations plus importantes d'arsenic (Sharma & Sohn, 2009). Cette adsorption peut se faire sur les oxydes et hydroxydes de Fe III, d'Al III et de Mn III/IV (Choi *et al.*, 2009).

En outre l'effet matière organique constitue un autre système de mobilisation, d'adsorption et de spéciation de l'arsenic. La matière organique correspond à des substances humiques de haut poids moléculaire ou à des molécules présentant des groupes fonctionnels réactifs tels que des groupements phénol, ou acide carboxylique. Ces groupes fonctionnels liés à la matière organique peuvent avoir un rôle dans la spéciation/mobilisation de l'arsenic par des réactions redox, des adsorbants inorganiques contenus dans les particules de matières organiques ou par des complexations des espèces d'arsenic. Les acides humiques peuvent adsorber jusqu'à 100-120 mmol d'arsenic par kilogramme (Sharma & Sohn, 2009) (Bissen & Frimmel, 2003b).

- Précipitation/dissolution due au pH, au potentiel redox et à la composition chimique. Les phases solides qui composent les aquifères peuvent subir des dissolutions ou des précipitations, l'arsenic contenu dans ces phases solides (comme constituant ou impureté) sera alors libéré. Dans certaines conditions du milieu, il peut y avoir par exemple co-précipitation avec le Fe(OH)₃. On peut, entre autres, trouver ce phénomène de co-précipitation As-Fe dans les drainages miniers acides (AMD), comme par exemple le site de Carnoulès (Travaux personnels -Chapitre IV).

▪ Les phénomènes bactériens dus à l'oxydoréduction enzymatique. Ces phénomènes bactériens ont été montrés dans diverses études comme étant parmi les principaux acteurs de la mobilisation/immobilisation de l'arsenic (Charlet & Polya, 2006; Lloyd & Oremland, 2006; Oremland *et al.*, 2005b; Rhine *et al.*, 2005; Sharma & Sohn, 2009). Par exemple les bactéries Fe-réductrices engendrent la dissolution de l'arsenic dans les sédiments, les bactéries sulfato-réductrices interviennent dans la mobilisation de l'arsenic en produisant du sulfure d'hydrogène qui va dissoudre les minerais et libérer l'arsenic (Ahmann *et al.*, 1997). Enfin les bactéries qui respirent l'arséniate peuvent elles aussi libérer l'As III des sédiments (Islam *et al.*, 2004) (Jones *et al.*, 2000) (Oremland & Stolz, 2005). Les mécanismes d'oxydo-réduction seront traités en détail dans le chapitre II de l'introduction.

L'arsenic est donc un composé aisément biodisponible dans l'environnement et en particulier il pourra se trouver en forte concentration dans les eaux de consommation. De par sa toxicité la consommation d'arsenic engendre alors un grave problème de santé publique.

II TOXICITE

L'arsenic est probablement le contaminant responsable du plus grand risque de mortalité et de morbidité au monde, risque dû à sa toxicité et au nombre de personnes exposées. Sa toxicité s'étage du niveau cellulaire, tissulaire jusqu'à l'organisme.

1. Au niveau cellulaire

Les cellules n'ont pas développé de système d'import, l'arsénite et l'arséniate empruntent les systèmes de transport existants.

L'arséniate est un analogue moléculaire du phosphate qui va inhiber la phosphorylation oxydative. L'As V va interrompre ce système en formant un ester d'arséniate d'ADP instable qui va court-circuiter la production d'énergie. D'autre part l'As V perturbe le métabolisme cellulaire puisqu'il peut y avoir production de glucose-6-arséniate à la place du glucose-6-phosphate. Enfin l'As V peut remplacer le phosphore dans la molécule d'ADN et ainsi empêcher les mécanismes de réparation (Mandal & Suzuki, 2002).

L'entrée de l'arséniate dans les cellules procaryotes se fait par les transporteurs du phosphate : les systèmes Pit ou Pst (Figure I1.5) (Rosen, 2002a) (Rosen & Liu, 2009). En excès de

phosphate, l'As V emprunte le système Pit exprimé constitutivement et peu spécifique. Néanmoins en présence plus faible de phosphate la cellule peut discriminer l'entrée de l'As V de celle du phosphate en induisant la synthèse d'un système d'entrée Pst cent fois plus spécifique au phosphate qu'à l'arséniate (Nies & Silver, 1995).

L'As III est non chargé à pH inférieur à 9,2 (Ramírez-Solís *et al.*, 2004) et se retrouve alors sous forme $\text{As}(\text{OH})_3$. Chez les procaryotes, l'arsenic va pouvoir entrer dans la cellule *via* la protéine GlpF, correspondant à un transporteur de polyols similaire aux aquaglycéroporines (transporteur de petites molécules non chargées type urée ou glycérol) (Figure I1.5) (Meng *et al.*, 2004). Néanmoins il doit exister d'autres systèmes d'entrée dans la cellule puisque la disruption du gène codant pour GlpF ne confère pas une résistance complète à l'As III (Meng *et al.*, 2004), *H. arsenicoxydans* par exemple ne possède pas de gènes codant le GlpF dans son génome. Chez les levures, le système d'entrée de l'As III dans les cellules se fait par un analogue d'aquaglycéroporine Fps1p (Wysocki *et al.*, 2001) (Rosen, 2002b) ou par un système de transport du glucose Hxt (Liu *et al.*, 2004a). Enfin dans les cellules mammifères, l'As III emprunte soit le transporteur du glucose GLUT (Rosen & Liu, 2009) (Figure I1.5), soit un système analogue aux aquaglycéroporines (Liu *et al.*, 2004b).

Une fois dans les cellules, l'arsénite se lie au groupement sulfhydrile. La formation de la liaison As III-S inhibe l'activité de nombreuses enzymes comme la glutathione réductase, glutathione peroxydase, thiorédoxine réductase et thiorédoxine peroxydase. Ce processus empêche aussi la respiration en se fixant sur les groupements thiols de la pyruvate deshydrogénase et de la 2 oxo-glutarate deshydrogenase (Oremland & Stolz, 2003). Ainsi l'As III inhibe le fonctionnement de plus d'une centaine d'enzymes (Gailer, 2007) (Duker *et al.*, 2005). Il a d'autre part été montré que l'arsénite a un effet mutagène sur l'ADN *via* les espèces réactives de l'oxygène (Hei *et al.*, 1998) (Pi *et al.*, 2002).

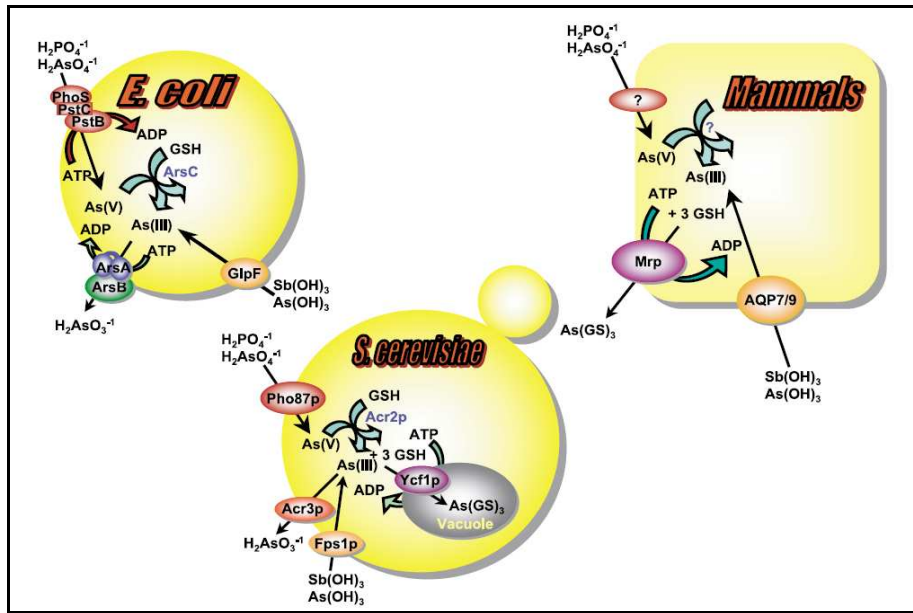


Figure 11.3 : Détoxification de l'arsenic dans les cellules procaryotes et eucaryotes. L'As V est transporté dans la cellule *via* les transporteurs du phosphate et As III entre *via* les aquaglycéroporines (Extrait de (Rosen, 2002a)).

2. Toxicité chez l'homme

Il est à noter que l'arsenic a un effet combiné avec d'autres paramètres tels que la nutrition, l'état général de santé, le tabagisme, l'exposition à d'autres substances chimiques, ... Ces différents facteurs peuvent avoir des actions synergiques et ont une responsabilité importante dans la modulation de la toxicité de l'arsenic (Schoen *et al.*, 2004) (Hopenhayn, 2006).

a. Symptômes et maladies

▪ L'exposition aiguë

L'ingestion d'une quantité toxique d'arsenic génère un effet après 30 à 60 minutes. Une toxicité sévère a été rapportée pour 1mg d'As III (Abernathy *et al.*, 2003). L'exposition aiguë à l'arsenic cause des vomissements, une sécheresse de la bouche et de la gorge, des crampes musculaires, des tremblements des mains et des jambes, des problèmes circulatoires,... Elle peut entraîner la mort par faiblesse rénale, hépatique ou cardiaque (Duker *et al.*, 2005).

▪ **Exposition chronique**

L'intoxication chronique à court terme engendre des symptômes non spécifiques comme des faiblesses chroniques, une perte des réflexes, une gastrite, des colites, la perte de poids et de cheveux,... (Bissen & Frimmel, 2003a).

La contamination à long terme due à l'absorption d'eau contaminée entraîne chez l'homme des affections de la peau (hyperkératose ou « Blackfoot disease »), des cancers de la peau, des poumons, de la vessie et des reins, mais peut aussi entraîner des dysfonctionnements du système vasculaire périphérique, des désordres neurologiques, de l'hypertension et du diabète. Il semble avoir un effet reprotoxique et tératogène (Schoen *et al.*, 2004) (Abernathy *et al.*, 2003).

Il est cependant difficile d'appréhender les effets de l'arsenic sur l'homme autrement que par des études épidémiologiques. La différence de métabolisme de ce métalloïde entre les espèces ne permet pas d'extrapoler les effets de l'arsenic sur le modèle animal aux effets de l'arsenic sur l'homme (Abernathy *et al.*, 2003).

De nombreuses études épidémiologiques ont montré qu'il existait une relation dose-réponse à l'arsenic ; ainsi la peau est très sensible et l'un des premiers signes d'intoxication chronique est alors l'hyperpigmentation (Yoshida *et al.*, 2004).

b. L'arsenic dans le corps humain

De façon générale, il a été montré que l'arsenic génère un stress oxydatif (Pi *et al.*, 2002), le blocage de la méthylation de l'ADN, l'inhibition des enzymes de réparation de l'ADN, des dommages au niveau chromosomique et l'altération de la transcription (Schoen *et al.*, 2004).

La toxicité de l'arsenic dépend naturellement de sa forme : les formes inorganiques As III et As V sont plus toxiques que les formes organiques. Par exemple, l'arsénobétaine ou l'arsénocholine retrouvées dans les produits de la mer sont « aisément » éliminées par les urines. Les formes inorganiques ingérées sont rapidement absorbées par le tractus gastro-intestinal (>80%). L'As V pourra aussi être réduit en As III par une réductase ou chimiquement par le glutathion (Radabaugh & Aposhian, 2000) (Gailer, 2007). Les différentes formes d'arsenic peuvent aussi être méthylées en MMA (acide monométhylarsine) ou DMA (acide diméthylarsine), ce qui facilite l'élimination par voie urinaire où environ 70 % de l'arsenic ingéré est éliminé (Cohen *et al.*, 2007). Il a récemment été montré que la toxicité du MMA et du DMA serait plus grande que celle de l'As III, ces composés méthylés ne sont pas seulement des mécanismes de détoxication, mais ils génèreraient des dommages directs sur l'ADN (Hughes, 2002). De plus la méthylation de l'arsenic génère une hypométhylation de l'ADN, due à l'épuisement de groupes méthyls. Cet ADN hypométhylé va alors avoir pour conséquence l'expression aberrante de certains gènes (Roy & Saha, 2002) (Zhao *et al.*,

1997). Toutefois une grande quantité d'arsenic se retrouvera stockée dans les os, la peau, les ongles et les cheveux.

c. La toxicité en quelques chiffres

De façon globale, la consommation d'eau contaminée à l'arsenic à hauteur de 0,05 mg/L induit l'augmentation du risque de cancer des poumons et de la vessie (Donnée OMS). Pour l'homme, la dose létale correspond à 250 mg d'arsenic/m³ durant 30 min d'exposition. Les doses létales (LD50) chez le rat pour l'As III, l'arsine, l'arséniate, le MMA et le DMA sont illustrées en Figure I1.4. Ainsi l'arsine est de loin la forme d'arsenic la plus toxique, suivie de l'arsénite et de l'arséniate, le même rang de toxicité pouvant être observé chez l'homme (Bissen & Frimmel, 2003a) (Yamauchi & Fowler, 1994) (Vaughan, 2006).

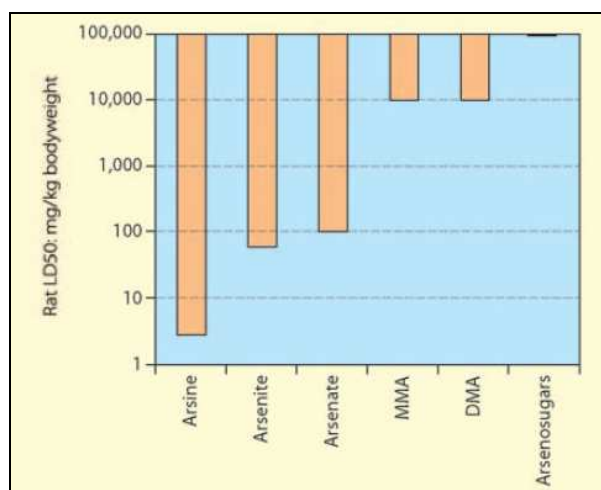


Figure I1.4 : Diagramme illustrant la relative toxicité des différentes formes d'arsenic. « Rat LD50 » correspond à la quantité d'arsenic en mg/masse corporelle des rats exprimée en kg, nécessaire pour tuer 50% de la population. MMA est l'acide monométhylarsonique et DMA est l'acide diméthylarsonique (Extrait de (Vaughan, 2006)).

d. Traitements

La glutathion peroxydase et la catalase modulent la génotoxicité de l'arsenic, donc l'augmentation du niveau d'antioxydants dans la cellule, ce qui présente un effet préventif et thérapeutique de l'intoxication à l'arsenic (Roy & Saha, 2002).

D'autre part, il est possible de mettre en place une thérapie par chélation. Les agents chélateurs sont des composés organiques qui vont fixer les cations métalliques en formant des complexes stables, les chélates, solubles dans l'eau et non toxiques qui seront facilement éliminés par les reins. Par exemple, il est possible d'administrer du dimercaprol (BAL : British Anti-Lewiste),

dimava (DMPS) ou méso-2,3-dimercaptosuccinate (DMSA) (Roy & Saha, 2002). Le dimercaprol, très utilisé, est un dérivé thiol du propanol, les groupes thiols forment un chélate relativement labile. Cependant la chélation présente de nombreux effets secondaires comme l'hypertension, la tachycardie, la fièvre, les vomissements... Aujourd'hui une nouvelle tendance émerge en thérapie qui consiste à combiner plusieurs modes de thérapie avec la chélation. Ainsi l'apport en vitamines (thiamine, vitamine E...), en acides aminés ou en métaux essentiels (Cu, Fe, Zn...) couplé à la chélation augmente l'effet du traitement et la rapidité de guérison de l'intoxication (Kalia & Flora, 2005).

3. Utilisation de l'arsenic en thérapie

L'arsenic est une substance naturelle qui a été utilisée comme médicament depuis plus de 2000 ans. Il a longtemps été employé, sous le nom de solution de Fowler, dans les traitements contre la malaria, la syphilis ou le psoriasis. Paul Ehrlich reçut le Prix Nobel de médecine en 1908 pour ses travaux sur l'utilisation de composés arséniés comme agents chimiothérapeutiques et combinés avec le composé « Salvarsan », comme agent antimicrobien.

Aujourd'hui il est utilisé pour ses effets contre certaines formes de cancers et en particulier certaines leucémies (Table I1.1). L'arsenic a un effet dose dépendant (Hu *et al.*, 2005) : il cause la différenciation cellulaire à faible dose et l'apoptose à dose relativement élevée. Il est utilisé en traitement sous deux formes principalement : l'arsenic trioxyde As_2O_3 (donné à une concentration entre 0,08 et 0,16 mg/kg/jour) et sous forme As_4S_4 . Comme tout produit, l'utilisation de l'arsenic peut entraîner l'apparition d'effets secondaires comme des réactions dermiques, des perturbations gastro-intestinales, et pouvant aller jusqu'à des neuropathies périphériques. Néanmoins ces effets sont moindres comparés à ceux d'une chimio- ou radio-thérapie.

Table I1.T1 : Effet potentiel d'un traitement à base d'arsenic sur différents cancers (Adapté de (Hu *et al.*, 2005)).

Efficacité potentielle en cancérologie	Effets biologiques
Leucémie	Induction de la différenciation
Leucémie, myélome, lymphome, cancer du sein, cancer de la prostate, cancer du pancréas, cancer du foie	Induction de l'apoptose
Fibrosarcome, myélome	Inhibition de l'angiogénèse
Cancer du sein, myélome	Immunomodulation

L'arsenic est extrêmement toxique pour l'homme en particulier, et pour tous les organismes de façon générale. Outre sa toxicité, l'arsenic cause de nombreux empoisonnements puisqu'il est très largement répandu dans les aquifères.

III L'ARSENIC DANS L'ENVIRONNEMENT

1. Répartition dans l'environnement

L'arsenic est le 20^{ème} élément le plus abondant dans la croûte terrestre, avec des concentrations d'environ 1-5 mg/kg en moyenne (0,1-1000 mg/kg) (Mukhopadhyay *et al.*, 2002) (Mandal & Suzuki, 2002) correspondant à un total d'environ 4.01×10^{16} kg d'arsenic. L'arsenic libéré de la lithosphère correspond à $2,2 \times 10^7$ kg par an suite aux éruptions volcaniques terrestres et marines (Matschullat, 2000). Selon le résumé du même auteur, $1,74 \times 10^6$ kg d'arsenic sont retrouvés dans l'atmosphère, ceci étant dû en partie aux émissions polluantes des pays industrialisés et en particulier à la combustion de carburant fossile. Enfin, la teneur en arsenic dans la pédosphère et l'hydrosphère est respectivement de $0.6-1.7 \times 10^9$ kg et 3,7 µg/L (en concentration totale). En moyenne, 0,6 à 0,9 mg/L d'arsenic dissous sont retrouvés dans les zones rurales et 3,2 à 5,6 mg/L dans les rivières contaminées par une activité industrielle, cette valeur pouvant atteindre plus de 25 mg/L. Ainsi l'homme est exposé à l'arsenic *via* l'inhalation d'air pollué, l'ingestion d'aliments et d'eau contaminés. La dose maximale non toxique s'élève à 0,3 µg d'As par kg et par jour, alors que la législation tolère 10 µg/l dans l'eau (WHO, 2001).

L'arsenic natif est rare et est souvent associé à d'autres métaux comme le cuivre, l'or ou le plomb, il est retrouvé dans plus de 200 minéraux différents. A l'état naturel, il existe principalement sous forme de minéraux sulfurés : le réalgar AsS, l'orpiment As₂S₃, l'énargite Cu₃AsS₄ et l'arsénopyrite (FeAsS) (Figure I1.5). A ces sulfures s'ajoutent deux oxydes, l'arsénolite (As₂O₃) et la claudétite (As₄O₆), et de nombreux arséniures métalliques de fer, de nickel et de cobalt.



Figure I1.5 : Cristal d'orpiment (Extrait de (Vaughan, 2006), de réalgar (<http://www.thunderhealing.org/rock/realgar.jpg>), d'énargite (http://www.futura-sciences.com/uploads/RTEmagicC_cuivre_302b.jpg.jpg), d'arsenopyrite (<http://www.geowiki.fr/images/thumb/d/d6/Arsenopyrite-4-hbt.JPG/640px-Arsenopyrite-4-hbt.JPG>)

2. La répartition géographique des zones contaminées

Depuis le début du siècle dernier, de nombreux cas de contamination collective ont été répertoriés dans plusieurs régions à travers le monde : en Argentine, 1917, Taiwan, 1920 ,où est décrite pour la première fois la maladie «Blackfoot disease » , puis au Chili en 1958, à Taiwan à nouveau dans les années 1980... Et encore aujourd'hui on pense en premier lieu à la situation alarmante du Bangladesh et de l'Inde où plus de 80 millions de personnes sont exposées et plus de 300 000 personnes présentent des lésions cutanées ou des cancers. Dans ces régions, un adulte sur dix pourrait développer un cancer dû à l'exposition à l'arsenic (Chowdhury *et al.*, 2000).

Bhattacharya et collaborateurs citent de nombreuses études qui tendent à montrer que de plus en plus de pays sont concernés. A titre d'exemple on peut citer le sud du Népal, le Pakistan, le Vietnam et le Cambodge, le centre du Gana, le sud de la Thaïlande, la Slovénie, la Turquie, la Province de Guizhou en Chine,... (Figure I1.6) (Bhattacharya *et al.*, 2007) (Garelick *et al.*, 2008) (Smedley & Kinniburgh, 2002). Il est d'autre part aisé d'imaginer la sous-évaluation de l'ampleur de la catastrophe, due aux difficultés d'investigation et de diagnostic de cancers dans les zones rurales de ces pays (Chowdhury *et al.*, 2000). L'utilisation d'eau polluée est également un problème majeur, car cette eau est utilisée dans l'irrigation des cultures et en particulier dans celle des plants de riz qui sera ensuite retrouvé dans la chaîne alimentaire (Bhattacharya *et al.*, 2007).

La France n'est pas en reste puisque de nombreux sites sont contaminés par l'arsenic (<http://infoterre.brgm.fr/viewer/MainTileForward.do>). Le cas de Carnoulès dans le Gard est à citer (Casiot *et al.*, 2003a) (Casiot *et al.*, 2003b) (Bruneel *et al.*, 2003) (Morin *et al.*, 2003). De nombreuses contaminations sont aussi à noter dans l'ouest de la France : l'ancienne mine de Loperec par exemple

(Battaglia-Brunet *et al.*, 2006a) ou le site de Salsigne dans l'Aude (Chatain *et al.*, 2005) ou encore à Ferrette en Alsace (Meyer *et al.*, 1999).

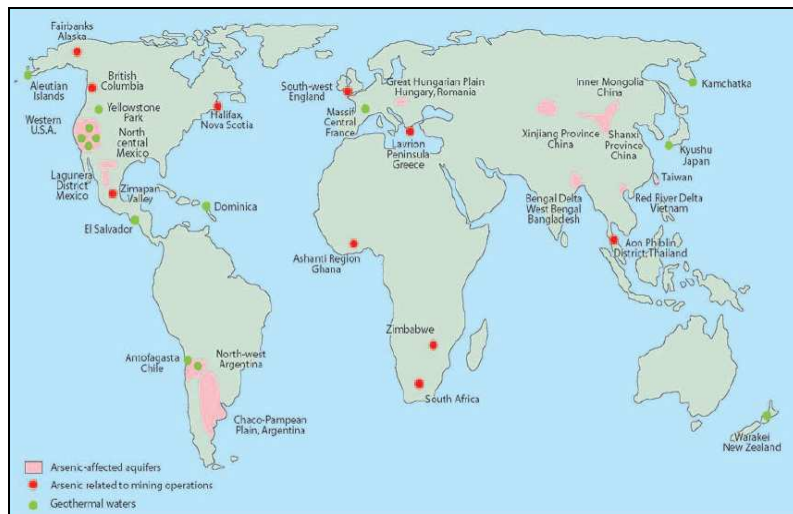


Figure 11.6 : Carte du monde montrant les zones connues pour leurs fortes concentrations en arsenic. Elles sont généralement associées aux sources hydrothermales, à des activités minières ou des aquifères contaminés (Extrait de (Smedley & Kinniburgh, 2002)).

3. Sources de la contamination à l'arsenic

a. Origine naturelle

L'arsenic se trouve à l'état naturel dans les roches arsénifères, ainsi 1 à 2 ppm d'arsenic sont retrouvés dans les quartz, feldspath, jusqu'à 10 ppm dans les péridoties, le basalte,... Et des concentrations très élevées apparaissent quand l'arsenic est un élément constitutif de minerais comme le réalgar, l'orpiment,... Celui-ci se retrouve dans les eaux de la nappe phréatique par la dissolution des minéraux ou minerais riches en arsenic.

Les pollutions humaines sont très localisées alors qu'au contraire, les contaminations d'origine naturelle des eaux de consommation dues aux fonds géochimiques riches en roches arsénifères sont très largement répandues au travers de la planète. Ces causes de contamination naturelle sont paradoxalement celles qui génèrent les plus grands problèmes de santé humaine. C'est le cas au Bangladesh où la contamination, d'origine naturelle, peut s'élever à 1000 µg d'arsenic par litre d'eau.

b. Origine anthropogénique

La principale utilisation d'arsenic est consacré à des fins industrielles et agricoles. Il est utilisé sous forme d'arséniate de sodium à des fins phytosanitaires : herbicides, insecticides, fongicides dans

l'industrie du bois. Son utilisation industrielle se concentre, d'autre part, dans la fabrication de verre (utilisé en décoloration, en se complexant à différents minerais colorés comme le fer ou le manganèse), dans la fabrication de peinture (acéto-arsénite de cuivre : vert de Paris), et dans les industries électroniques (production de semi-conducteurs),... Il en résulte environ 250 sites industriels contaminés par l'arsenic en France [BASOL, 2005].

L'exploitation minière constitue une autre source importante de contamination issue d'une activité humaine. En effet, les sources d'arsenic dans ces environnements sont la conséquence de la mobilisation des métaux et métalloïdes contenus dans les minerais, le lavage des galeries minières par les eaux souterraines ou le lavage des terrils par les eaux de pluie et peuvent conduire à des concentrations importantes dans les cours d'eau et les aquifères.

En particulier, les drainages miniers acides (*l'abréviation AMD sera utilisée dans la suite de l'exposé pour ne pas interférer avec l'abréviation DMA Diméthylarsine*) sont un exemple extrême de contamination par des métaux de sites miniers. Ils se forment par percolation d'eau sur les stériles (déchets des minerais excavés) contenant des sulfures ou dans les galeries minières par des infiltrations d'eau ou lors de mouvements de la nappe phréatique. Cette production d'acide conduit alors à une augmentation de la solubilité des éléments métalliques et de leur mobilité. Les effluents contiennent des concentrations très élevées en métaux, souvent toxiques, et constituent l'un des problèmes environnementaux majeurs de l'industrie extractive mondiale.

Par exemple, le site minier de Carnoulès dans le Gard présente de 100 à 350 mg/l d'arsenic soluble (majoritairement retrouvé sous forme As III) dans le ruisseau Reigous, alimenté par les eaux de percolation des déchets miniers (Bruneel *et al.*, 2006) (Koffi *et al.*, 2002). Néanmoins dans ces environnements extrêmes (acidité forte et haute concentration en métaux), des formes de vie peuvent être retrouvées, constituées essentiellement de bactéries et d'archés qui jouent un rôle important dans la mobilisation de ces métaux (Tyson *et al.*, 2004).

Enfin il est intéressant de noter que la production annuelle d'arsenic s'élevait à plus de 30 000 t en 1993 (Matschullat, 2000) (Bissen & Frimmel, 2003a) et 37 500 t en 2004 (Vaughan, 2006). Le prix de ce métal est de 3,50 dollars pour 0,454 kg (une once).

4. Législation

L'OMS a déterminé en 2001 la valeur maximale d'arsenic à 10µg/L dans l'eau de consommation, cette valeur ayant été reprise dans le droit européen (directive 98/83/CE, conseil du

03/11/98) et dans le droit français (décret 2001-1220 du 20/12/2001, entré en vigueur le 25/12/2003).

Des réglementations particulières sont en vigueur pour les rejets industriels ou les effluents miniers. Ainsi il existe une réglementation nationale et des arrêtés préfectoraux qui sont établis au cas par cas en fonction du type de production et du flux de rejet. Par exemple, le cas de l'ancienne mine de Salsigne (France), pour laquelle l'arrêté préfectoral fixe une concentration d'arsenic à 0,05 mg/L dans les eaux de rejet ou un flux d'arsenic inférieur à 30 g/j.

5. La remédiation de l'arsenic

De nombreuses études sont publiées concernant les différentes méthodes de remédiation de l'arsenic, parmi elles Sharma et Sohn (2009), Battaglia-Brunet et collaborateurs (2008) et Lièvremon et collaborateurs (2009) résument quelques-unes des techniques utilisées (Sharma & Sohn, 2009) (Battaglia-Brunet *et al.*, 2008) (Lievremont *et al.*, 2009). Ainsi les sites contaminés par l'arsenic peuvent être remédiés par, entre autres, l'adsorption, la phytoremédiation, l'oxydation photochimique, l'oxydation photocatalytique, la filtration, la coagulation et l'osmose inverse, ... (Sharma & Sohn, 2009).

L'adsorption de l'arsenic se fait par différents composés, par exemple le chlorure de fer, le sulfate de fer, des résines échangeuses d'ions, des polymères organiques, ... ou encore sur des charbons actifs. Cette technique d'adsorption très utilisée s'avère très efficace puisque plus de 95 % d'As V est adsorbé pour un taux de 10-50 mg/L à des pH entre 5,0 et 7,5.

L'oxydation photochimique est basée sur l'irradiation UV de l'As III. Cette oxydation se produit par la photolyse de peroxyde d'hydrogène en condition acide ou par l'intermédiaire de FeIII – hydroxyde et –chlorure. Ceux-ci absorbent les photons et libèrent des radicaux (hydroxyl et dichloro) qui oxydent l'As III en As V. La lumière du soleil est elle aussi suffisante au processus et il pourrait donc avoir lieu dans l'environnement et en particulier dans les AMDs.

L'oxydation photocatalytique par le TiO₂ est suivie de l'adsorption de l'As V sur le TiO₂. Lors de la photocatalyse par le TiO₂ des ions superoxydes sont libérés et vont oxyder l'As III en As V qui peut alors se fixer sur le TiO₂ (Sharma & Sohn, 2009).

Enfin la phytoremédiation est une technique très étudiée et présente l'avantage d'être bon marché. Certaines plantes ont été décrites comme résistantes à l'arsenic, mais surtout comme étant hyperaccumulatrices. C'est le cas par exemple de *Pteris vittata* qui peut accumuler jusqu'à 23 g d'arsenic par kilogramme de plante (Tu *et al.*, 2002). Cette technique présente aussi l'avantage d'avoir un coût faible et de pouvoir ainsi être mise en place dans des pays en voie de développement.

Néanmoins ces techniques ne semblent pas toujours réalisables ou adaptées au traitement *in situ* des aquifères souterraines contaminées, en particulier dans les pays en voie de développement (Roy & Saha, 2002). Différentes techniques de remédiation de l'arsenic sont tout de même disponibles : des méthodes de remédiation bon marché ont été développées pour les pays concernés, consistant en des adsorptions de l'arsenic sur des matériaux naturels comme la latérite ou l'oxisol Fe-riche (Bhattacharya *et al.*, 2007).

Certaines de ces techniques nécessitent l'oxydation préalable d'As III en As V. Cette oxydation peut être réalisée par l'ajout d'oxydants forts comme le permanganate de potassium, le peroxyde d'hydrogène ou l'ozone. Cette étape d'oxydation se faisant indépendamment de l'étape d'élimination, les étapes d'oxydation classiques (ajout de chlore, permanganate de potassium) peuvent donc être facilement remplacées par de l'oxydation biologique.

Aujourd'hui des tests montrent l'efficacité de cette oxydation biologique de l'arsenic ou de l'oxydation enzymatique directe. Des tests en réacteurs à lit-fixé utilisant des bactéries oxydantes présentent une oxydation efficace entre 10 µg/L et 100 mg/L (Battaglia-Brunet *et al.*, 2002).

Chapitre II

Les rôles des bactéries dans le cycle
biogéochimique de l'arsenic

Avec 10^6 - 10^9 cellules par cm^3 de sol (Haferburg & Kothe, 2007), les bactéries sont les organismes les plus nombreux dans l'environnement. De par leur petite taille et leur multiplicité, elles offrent une surface globale importante donc des possibilités d'interactions fortes avec leur environnement.

I LES TRANSFORMATIONS BACTERIENNES ET LEURS ROLES DANS LE CYCLE BIOGEOCHIMIQUE DE L'ARSENIC

La proportion relative des différentes espèces d'arsenic et des différents états d'oxydation dans un environnement dépend des conditions redox et du pH du milieu, mais également des biotransformations liées au métabolisme bactérien. En effet, en absence de micro-organismes, les réactions d'oxydation et de réduction entre les différentes espèces d'arsenic sont extrêmement lentes, même dans les conditions thermodynamiquement favorables. Les bactéries jouent donc un rôle majeur dans la mobilisation/immobilisation de l'arsenic et donc dans le cycle biogéochimique de l'élément (Figure I2.1). Pour les procaryotes, ces réactions sont soit utilisées comme moyen de détoxification soit pour leur apport énergétique.

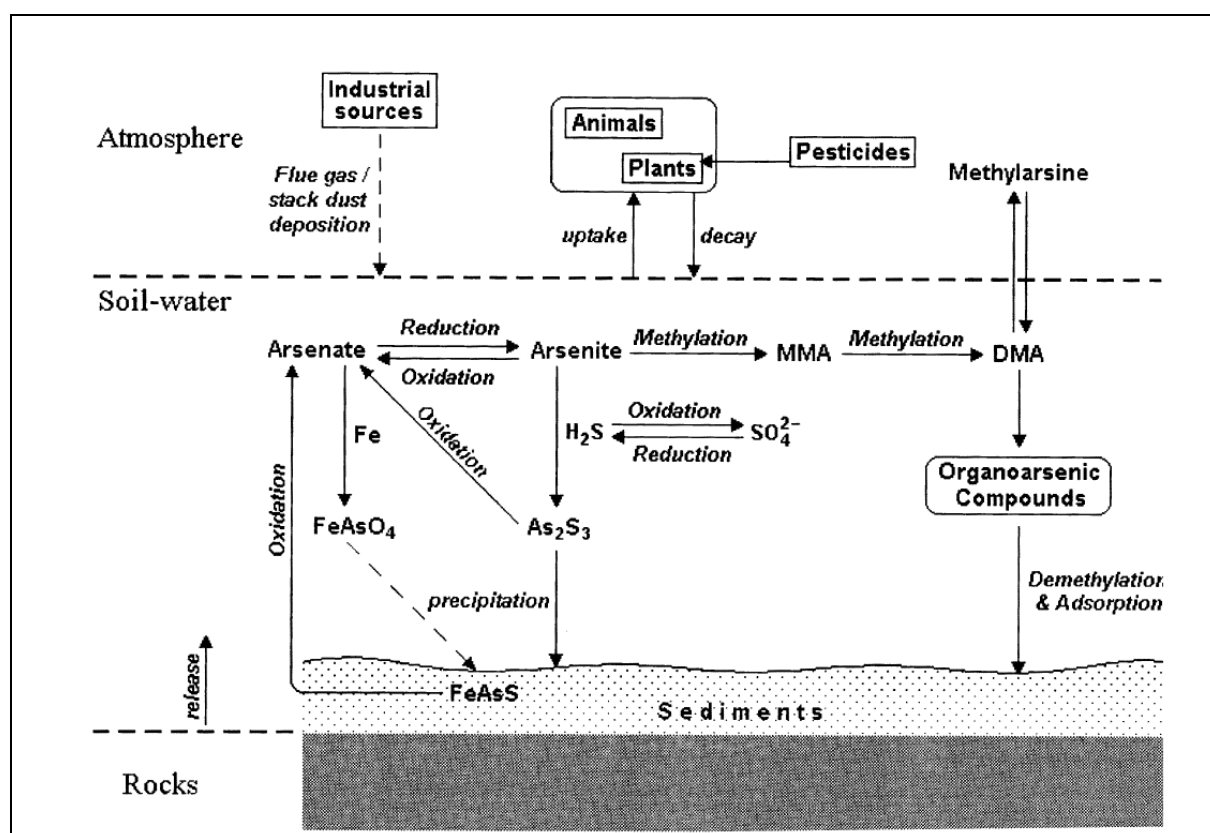


Figure I2.1 : Cycle de l'arsenic dans l'environnement. Représentation des réactions majeures qui influencent la mobilisation/immobilisation, le transport et la disponibilité de l'arsenic (Extrait de (Roy & Saha, 2002)).

Les biotransformations bactériennes de l'arsenic sont : la réduction de l'arséniate en arsénite, l'oxydation de l'arsénite en arséniate, les différentes réactions de méthylation qui permettent aux formes inorganiques de l'arsenic d'être transformées en formes organiques, la séquestration de l'arsenic et la respiration de l'arséniate (Páez-Espino *et al.*, 2009). Ces différents mécanismes sont présentés dans la Figure I2.2 et seront détaillés dans les paragraphes suivants.

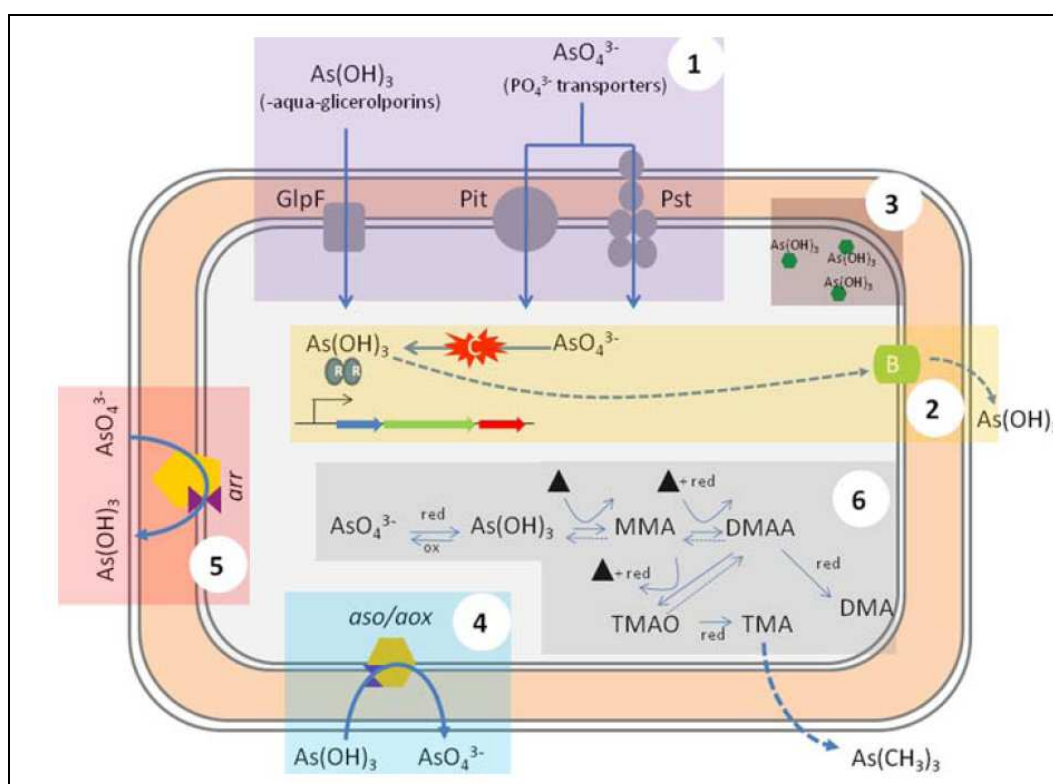


Figure I2.2 : Diagramme des différents processus microbiens impliqués dans la biochimie de l'arsenic. (1) L'arsenic entre dans la cellule via les transporteurs du phosphate (arséniate) ou via les aquaglycéroporines (arsénite). (2) Une fois dans la cellule, l'arséniate est réduit en arsénite par ArsC puis l'arsénite est exporté de la cellule par la pompe d'efflux ArsB. (3) L'arsénite peut être complexé par des peptides Cys-riches. (4) L'arsénite peut être oxydé en arséniate par l'arsénite-oxydase AoxAB. Il peut dans certains cas être utilisé comme donneur d'électrons. (5) L'As V peut être utilisé comme accepteur d'électrons dans la respiration via le système Arr. (6) Les formes inorganiques de l'arsenic peuvent être transformées en formes organiques par une cascade de méthylation (Extrait de (Páez-Espino *et al.*, 2009)).

II L'OXYDATION DE L'ARSENITE EN ARSENIATE

En raison de la plus faible toxicité et de la mobilité moindre de l'arséniate par rapport à l'arsénite, l'oxydation de l'As III en As V conduit donc à la détoxification du milieu environnant la

bactérie. Ce mécanisme conduit à la diminution de la biodisponibilité de l'arsenic dans l'environnement.

L'oxydation de l'arsénite en arséniate suit la réaction suivante :



1. Phylogénie et diversité

Depuis la première description d'une bactérie arsénite-oxydante, *Bacillus arsenoxydans*, par Green en 1918 (Green, 1918) une centaine de procaryotes arsénite-oxydants ont été isolées d'environnements variés tels que : des sédiments miniers (Santini *et al.*, 2000) (Santini *et al.*, 2002a) (Duquesne *et al.*, 2008), des eaux d'égout (Phillips & Taylor, 1976), des lacs hyper-salins (Oremland *et al.*, 2005b), des sols (Osborne & Ehrlich, 1976), des sources géothermiques (Gihring & Banfield, 2001) (Salmassi *et al.*, 2002), des eaux industrielles arséniées (Weeger *et al.*, 1999), etc.

Ces bactéries phylogénétiquement diverses appartiennent principalement aux alpha-, beta- et gamma-protéobactéries, au genre *Chloroflexus*, mais aussi au phylum des Firmicutes et des actinobactéries et au genre *Thermus*. Enfin des gènes *aox* ont été retrouvés chez des archées ; c'est le cas, par exemple de *Sulfolobus acidocaldarius*, *Sulfolobus tokodaii* ou *Aeropyrum pernix* (Santini & Vanden Hoven, 2004).

Les bactéries arsénite-oxydantes peuvent aujourd'hui être classées en deux groupes : les hétérotrophes (qui poussent en présence de matière organique) et les chimiolithoautotrophes (aérobies ou anaérobies, elles utilisent l'arsénite comme donneur d'électrons et le $\text{CO}_2/\text{HCO}_3^-$ comme source de carbone).

L'oxydation hétérotrophique¹ de l'arsénite par les bactéries sert dans un premier temps à la détoxification du milieu environnant la bactérie. Les gènes codant l'oxydase sont retrouvés dans des souches oxydantes isolées, mais aussi directement sur l'ADN extrait d'échantillons de sols contaminés (Inskeep *et al.*, 2007). Les souches décrites sont aujourd'hui très nombreuses, on peut citer par exemple *Alcaligenes faecalis* (Phillips & Taylor, 1976), *Herminiimonas arsenicoxydans* (Muller *et al.*, 2003), *Hydrogenophaga* sp. str. NT-14 (nommée NT-14 dans la suite de l'exposé) (Hoven *et al.*, 2004), *Agrobacterium tumefaciens* (Kashyap *et al.*, 2006), *Thiomonas arsenivorans* (Michel *et al.*, 2007) (Duquesne *et al.*, 2008), *Ochrobactrum tritici* SCII24 (Branco *et al.*, 2009), *Achromobacter* sp. SY8 et *Pseudomonas* sp. TS44 (Cai *et al.*, 2009b). De façon particulière, en condition aérobie, les souches *Thermus aquaticus* et *T. thermophilus* oxydent l'As III dans le but de

¹ Les bactéries hétérotrophiques arsénite-oxydantes oxydent l'arsénite mais ont nécessairement besoin d'une source d'énergie alternative et de matière organique pour se développer

détoxifier le milieu alors qu'en condition anaérobie ces souches peuvent utiliser l'As V comme donneur d'électrons (Gihring & Banfield, 2001). Enfin certaines souches hétérotrophes peuvent tirer de l'énergie de l'oxydation de l'arsénite, comme la souche NT-14 (Hoven *et al.*, 2004), néanmoins celle-ci ne peut pas croître comme une chimioautotrophe.

Les bactéries arsénite-oxydantes chimiolithoautotrophes² utilisent l'As III comme source d'énergie et le CO₂ comme source de carbone. C'est-à-dire qu'elles couplent l'oxydation de l'arsénite (comme donneur d'électrons) à la réduction de l'oxygène ou du nitrate. Quelques exemples (Oremland & Stolz, 2003) : *Rhizobium* NT26 (nommée NT-26 dans la suite de l'exposé) (Santini *et al.*, 2000), *Thiomonas arsenivorans* (Battaglia-Brunet *et al.*, 2006b), *Thiomonas* 3As (Duquesne *et al.*, 2007), *Alkalilimnicola ehrlichei* MLHE-1 qui utilise NO₃⁻ comme accepteur terminal d'électrons (Oremland *et al.*, 2002) (Hoeft *et al.*, 2007), les souches *Sinorhizobium* sp. DAO10 et *Azoarcus* sp. DAO1 couplent l'arsénite oxydation à la dénitrification (Rhine *et al.*, 2005) (Rhine *et al.*, 2006).

2. L'arsénite-oxydase

L'enzyme la plus étudiée est celle issue d'*Alcaligenes faecalis* pour laquelle la structure cristalline a été déterminée par Ellis et collaborateurs, 2001.

a. L'enzyme

L'arsénite-oxydase d'*A. faecalis* est un hétérodimère $\alpha\beta$ constitué de deux sous-unités : une petite sous-unité Rieske à centre [2Fe-2S] appelé AoxA (14kD-134 résidus) et une grande sous-unité à centre [3Fe-4S] et à cœur molybdoptérique appelée AoxB (85kD-825 résidus) (Figure I2.3), ce qui la classe dans la famille des enzymes diméthylsulfoxides (DMSO) réductases (Ellis *et al.*, 2001). Elle est assez éloignée en terme d'identité de séquence des autres DMSO réductases (par exemple 23% d'homologie de séquence avec la nitrate-réductase dissimilatrice de *Desulfovibrio desulfuricans*) et forme un groupe à part, néanmoins sa structure est conservée (le centre molybdoptérine par exemple) (Inskeep *et al.*, 2007). Les arsénite-oxydases de *A. faecalis* ou de *H. arsenicoxydans* sont des hétérodimères $\alpha\beta$ (100kDa) (Muller *et al.*, 2003), alors que celles de NT26 et de NT14 sont hétérotétramériques $\alpha_2\beta_2$ (219 kDa) (Santini & Vanden Hoven, 2004) et hétérohexamériques $\alpha_3\beta_3$ (309 kDa) (Hoven *et al.*, 2004), respectivement.

² Les bactéries chimiolithoautotrophiques arsénite-oxydantes génèrent de l'énergie via l'oxydation de l'arsénite et utilisent le carbone atmosphérique comme source de carbone.

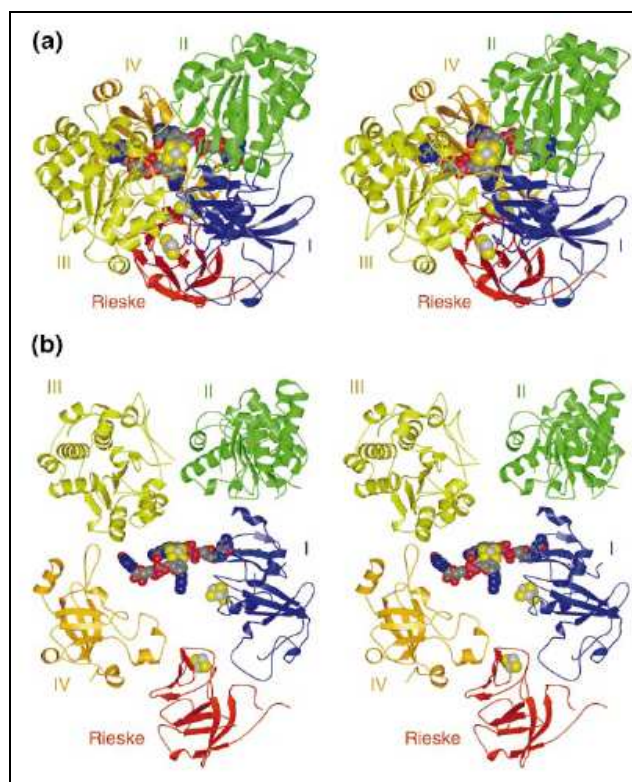


Figure 12.3 : Représentation biochimique de l'arsénite-oxydase d'*Alcaligenes faecalis*. (a) Les domaines I-IV de la grande sous-unité sont de couleur bleue, verte, jaune et orange, respectivement. La sous-unité Rieske est de couleur rouge. Le cofacteur à molybdène, les clusters [3Fe-4S] et [2Fe-2S] sont représentés. (b) Vue éclatée de l'arsénite-oxydase (Extrait de (Ellis *et al.*, 2001)).

b. Les sous-unités

La grande sous-unité AoxB est divisée en quatre domaines. Le premier est composé des résidus 1-119, 532-558 et 622-657 (trois feuilletts β antiparallèles et six hélices). Il fixe le cluster [3Fe-4S] au niveau du motif Cys21-X₂-Cys24-X₃-Cys28-X₇₀-Ser99 ainsi que la sous-unité Rieske. Les domaines II –IV fixent le centre molybdène, les domaines II et III ont des topologies $\alpha\beta$ similaires et sont constitués par les résidus 120-195, 424-531 et 559-621 pour le domaine II et des résidus 196-423 et 658-682 pour le domaine III. Enfin le Domaine IV est composé des résidus 683-825.

Le cœur à molybdène se situe au centre de la protéine et est lié par les domaines I, II et III. Il est constitué par un atome de Mo lié symétriquement par deux cofacteurs antiparallèles molybdoptérines (Ellis *et al.*, 2001).

La petite sous-unité Rieske AoxA possède un centre [2Fe-2S] lié par les résidus Cys60, His62, Cys78 et His81 (Ellis *et al.*, 2001) et un peptide signal TAT comme décrit ci-après (Muller, 2004) (Duquesne *et al.*, 2007).

c. Le cytochrome

Chez NT-14, un cytochrome de la famille c-551 a été mis en évidence comme intervenant dans le transfert des électrons suite à l'oxydation de l'arsénite. Il est constitué de 20 acides aminés et contient en N-terminal, le motif Cys-X₂-Cys-His caractéristique d'une liaison à un hème (Hoven *et al.*, 2004).

Le gène *aoxD* chez *H. arsenicoxydans* possède 35 % d'identité de séquence avec le gène codant pour le cytochrome de NT-14.

3. Le transport de l'arsénite-oxydase

Chez *A. faecalis*, l'arsénite-oxydase est une protéine périplasmique localisée sur la face externe de la membrane interne (Ellis *et al.*, 2001). Dans les cas de NT-26 et NT-14 l'enzyme serait soluble et localisé dans le périplasme (Santini *et al.*, 2000) (Hoven *et al.*, 2004). Il contient en partie N terminale de la sous-unité Rieske une séquence signal TAT (*Twin Arginine Transporter*) avec un motif RRGFLK très fortement conservé. La séquence TAT permet le guidage de la protéine repliée du cytoplasme au périplasme. Les cofacteurs molybdoptérines et le cluster [Fe-S] sont incorporés au polypeptide dans le cytoplasme, la protéine est repliée puis transportée dans le périplasme ou vers la membrane grâce au système TAT (Silver & Phung le, 2005) (Silver & Phung, 2005). Contrairement aux systèmes TAT classiques, Duquesne et collaborateurs proposent que le peptide signal d'AoxA ne serait pas clivé une fois localisé dans la membrane, mais servirait d'ancre à la protéine (Duquesne *et al.*, 2007).

4. le fonctionnement de l'arsénite-oxydase

Les centres molybdoptérines sont orientés tête-bêche dans la structure de l'AoxB. Le site de liaison du substrat se situe sur les résidus His195, Glu203, Arg419 et His423 (Ellis *et al.*, 2001). Le site actif montre alors quatre interactions Mo-S et une double liaison à l'oxygène quand l'arsénite-oxydase est réduite et une liaison additionnelle Mo-O lorsque l'enzyme est oxydé. Il se produit une attaque nucléophile de la paire d'électrons de l'arsénite sur la double liaison Mo VI=O (Figure I2.4). Les électrons sont transférés de l'arsénite à l'atome de molybdène, en réduisant le Mo VI en Mo IV, la réoxydation du molybdène se fait par la déprotonation de l'eau au niveau du site actif.

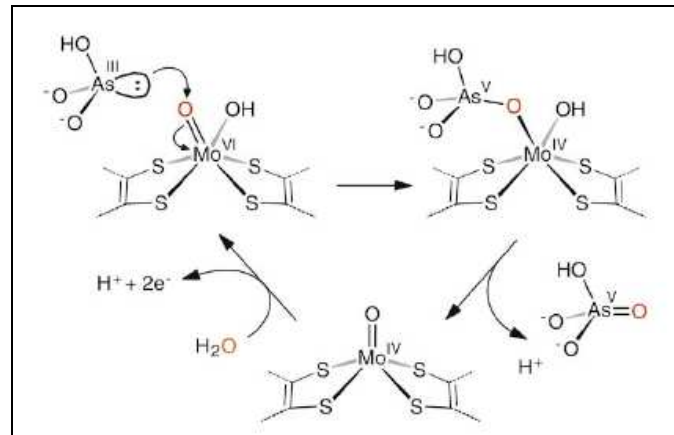


Figure 12.4 : Représentation schématique du mécanisme d'oxydation de l'arsénite en arséniate par l'arsénite-oxydase. L'atome d'oxygène transféré de l' H_2O à l'arsénite au cours du cycle catalytique est représenté en rouge (Extrait de (Ellis et al., 2001)).

Après la réduction du Mo VI en Mo IV, deux électrons sont transférés au centre [3Fe-4S] de la même sous-unité, puis au centre [2Fe-2S] de la petite sous-unité Rieske (Figure 12.5) (Silver & Phung le, 2005). Ces électrons sont alors transférés à une protéine de la chaîne respiratoire, un cytochrome ou une azurine par exemple.

Dans le cas de NT-14, les électrons seraient transmis au cytochrome c-551 puis à une cytochrome oxydase aa_3 , quatre électrons sont requis pour réduire l' O_2 en H_2O , un gradient de proton se forme alors (*via* le complexe aa_3) qui pourra être utilisé pour la synthèse d'ATP (Figure 12.6) (Hoven et al., 2004).

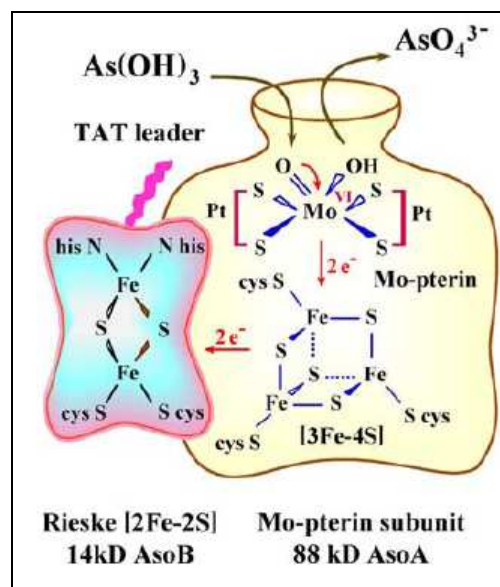


Figure 12.5 : Modèle de transfert d'électron de l'arsénite-oxydase lors de l'oxydation de l'As III en As V (Extrait de (Silver & Phung le, 2005)).

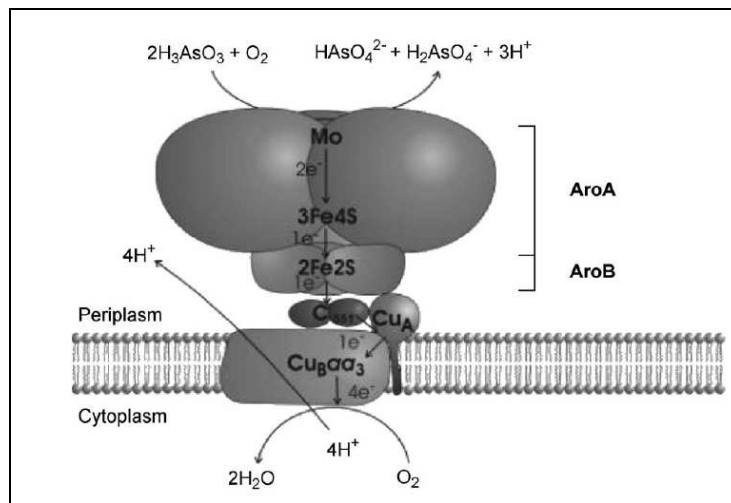


Figure 12.6 : Représentation de la chaîne de transport d'électrons pour une oxydation hétérotrophe de l'arsénite. L'arsénite est oxydé en arséniate au niveau du site actif Mo de la sous-unité AoxB. Les deux électrons sont alors transférés au centre [3Fe-4S], puis un par un au centre [2Fe -2S] de la sous-unité Rieske. Les électrons sont transférés au cytochrome c-551 (dimériques) lié à l'AoxA. Enfin les électrons sont transférés à l'hème du cytochrome c oxydase aa₃. Quatre électrons sont nécessaires pour réduire O₂ en H₂O dans le cytoplasme de la cellule et quatre protons sont pompés du cytoplasme générant un gradient de force proton-motrice au travers de la membrane, gradient qui peut par conséquent être utilisé pour générer de l'ATP (Extrait de (Hoven *et al.*, 2004)).

5. Les gènes *aox*

L'arsénite-oxydase est codée par les gènes *aoxA* et *aoxB* qui sont responsables respectivement de la synthèse de la petite et la grande sous-unité catalytique de l'enzyme. Ces gènes sont organisés en opéron avec les gènes *aoxC* et *aoxD* codant respectivement pour une nitroréductase et un cytochrome. L'opéron *aox* a été pour la première fois mis en évidence chez *H. arsenicoxydans* (Muller *et al.*, 2003).

Trois différents noms ont été donnés aux gènes codant pour l'arsénite-oxydase, ainsi *aoxAB* (arsénite-oxydase) pour *H. arsenicoxydans* (Muller *et al.*, 2003). Les noms *asoBA* (**As** oxydante) sont donnés pour *A. faecalis* (Ellis *et al.*, 2001) et l'homologie en acides aminés correspond à 62 % et 73 % avec *aoxAB* de *H. arsenicoxydans* (522 pb et 2478 pb, respectivement). Enfin la troisième terminologie est donnée pour NT-26 (Santini *et al.*, 2000) et correspond à *aroBA* (arsénite-oxydase) (528 pb et 2538 pb, respectivement). Les identités de séquences protéiques sont de 43 % et 49 % avec *asoBA* de *A. faecalis*.

De façon globale, on peut considérer que les termes *Aox/aox* sont utilisés pour les enzymes/gènes des souches hétérotrophes et la terminologie de *Aro/aro* pour ceux des souches chimioautotrophes (Hoeft *et al.*, 2007).

Les gènes *aoxAB* sont fréquemment suivis d'un gène codant pour un cytochrome (*aoxD*) et d'un gène pour une nitroréductase (*aoxC*). Le cytochrome est impliqué dans le transfert des électrons suite à l'oxydation de l'arsénite par l'arsénite-oxydase, en particulier chez NT-14 il correspond à un cytochrome de la famille c-551 (Hoven *et al.*, 2004). Aucune fonction n'a encore été donnée à AoxD.

Une variabilité de synténie des opérons *aox* existe, même si l'ordre *aoxAB* est toujours conservé (Figure I2.7). Bien souvent l'opéron *aox* se trouve à proximité de l'opéron *ars* formant alors un « îlot de gènes arsenic » comme déterminé chez *A. faecalis* ou chez *H. arsenicoxydans*. De nombreuses souches présentent aussi les gènes *aoxRS* en amont de l'opéron. Il y a une plus grande diversité en ce qui concerne la présence ou non des gènes *aoxCD* composant l'opéron *aox* chez *H. arsenicoxydans*.

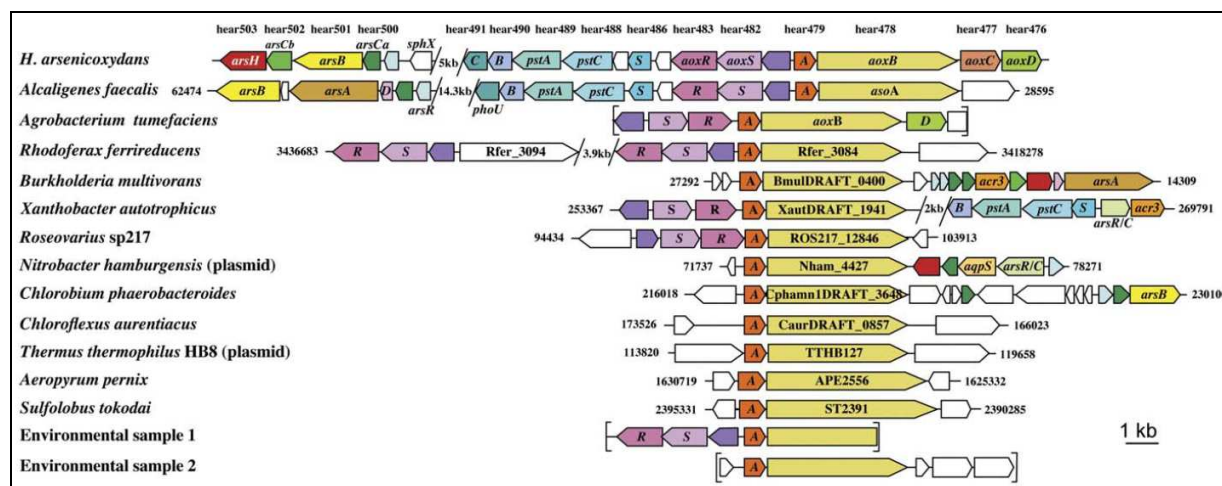


Figure I2.7 : Organisation des clusters de gènes *aox* chez *H. arsenicoxydans* et chez différentes souches métabolisant l'arsenic. Les opérons *aox* des souches *H. arsenicoxydans*, *A. faecalis*, *X. autotrophicus*, *N. hamburgensis*, et *C. phaeobacteroides*. Dans les trois premières bactéries ces gènes sont associés avec un système de régulation à deux composants *aoxRS* (Extrait de (Muller *et al.*, 2007)).

6. La régulation transcriptionnelle et post-transcriptionnelle de l'arsénite-oxydase

L'opéron *aox* est régulé par le système de transduction du signal à deux composants AoxRS codé par les gènes *aoxRS* (Kashyap *et al.*, 2006), localisé chez *H. arsenicoxydans* en amont de

l'opéron *aox*. Une étude réalisée chez *A. tumefaciens* a montré que l'AoxS possède une séquence comportant les motifs d'un senseur kinase, c'est-à-dire une séquence de 14 acides aminés qui jouerait le rôle de récepteur signal périplasmique. AoxS va s'autophosphoryler après détection d'As III dans le système et il va ensuite phosphoryler AoxR (Figure 12.8) (Kashyap *et al.*, 2006). AoxR phosphorylé va se fixer au facteur sigma RpoN ; en effet les séquences de fixation au facteur sigma ont été détectées chez AoxR d'*H. arsenicoxydans*. Le facteur sigma, son coactivateur (AoxR) et l'ARN polymérase vont alors se fixer sur la séquence consensus Shine-Dalgarno (-12/-24) du promoteur et activer la transcription de l'opéron (Figure 12.8) (Kashyap *et al.*, 2006) (Koechler *et al.*, 2009).

Ainsi chez les souches possédant les gènes *aoxRS*, l'expression de l'opéron *aox* est induite par la présence d'As III (Kashyap *et al.*, 2006) (Koechler *et al.*, 2009). Les études sur *A. tumefaciens* montrent d'autre part que la régulation de l'opéron serait aussi contrôlée par le quorum sensing, c'est-à-dire qu'elle serait dépendante de la densité cellulaire.

Pour finir, une régulation post-transcriptionnelle pourrait moduler l'expression de l'enzyme. En effet, DnaJ (protéine chaperon) serait impliquée dans la conformation de l'arsénite-oxydase, avant son export dans le périplasma par le système de sécrétion TAT (Koechler *et al.*, 2009) (Figure 12.8).

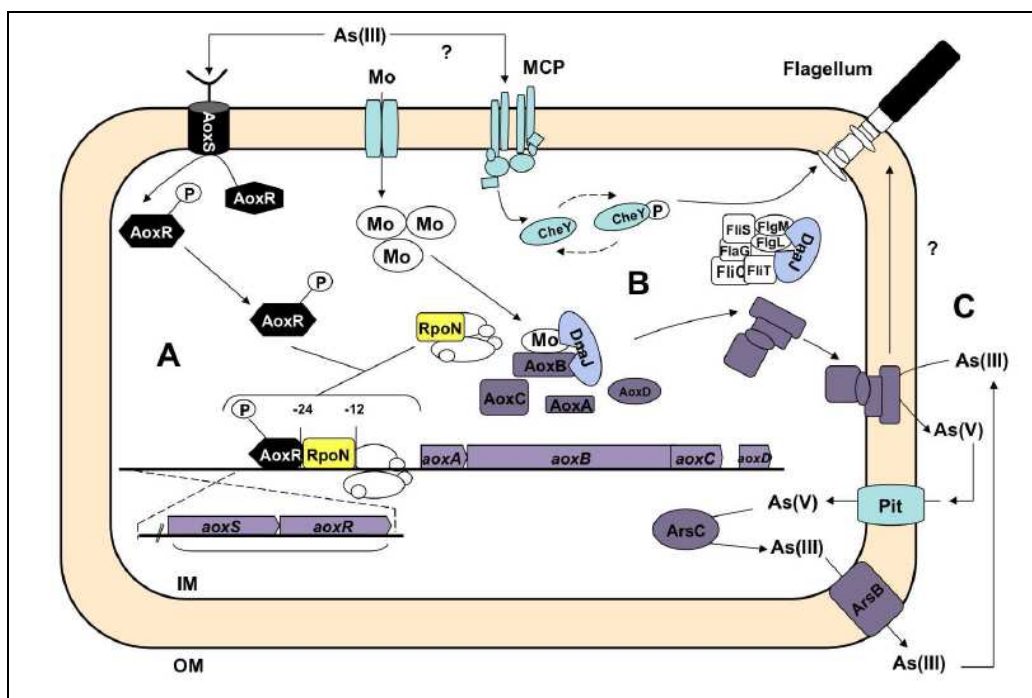


Figure 12.8 : Représentation conceptuelle de l'oxydation de l'arsénite chez *H. arsenicoxydans*. (A) La régulation transcriptionnelle : AoxS joue le rôle de senseur de l'As III et phosphoryle AoxR. L'AoxR phosphorylé se lie à RpoN, qui va interagir avec l'ARN polymérase et l'ensemble va initier la transcription de l'opéron *aox*. (B) La régulation post-transcriptionnelle : DnaJ est impliquée dans la conformation de l'arsénite-oxydase avant son export dans le périplasma par le système de sécrétion

TAT. (C) L'arsénite-oxydase oxyde la forme toxique As III en une forme moins toxique l'As V. Ce processus est connu pour affecter la motilité, qui doit faire intervenir une protéine MCP de chimiotactisme et nécessite DnaJ comme co-chaperon. IM = membrane interne, OM = membrane externe (Extrait de (Koechler *et al.*, 2009)).

Un gène supplémentaire *aoxX*, trouvé entre *aoxRS* et *aoxABCD* serait, cotranscrit avec *aoxRS* et induit par l'As III. La fonction de ce gène n'a pas encore été trouvée ; il code pour une protéine qui participerait à la régulation de l'opéron puisqu'il est cotranscrit avec *aoxR*. Cependant, une séquence signal TAT a été retrouvée dans la séquence du gène, ce qui localise la protéine dans la membrane cellulaire (Cai *et al.*, 2009a). Ces deux éléments contradictoires ne permettent donc pas de conclure quant au rôle de cette protéine.

7. Arsénite-oxydase et milieu originel

Les prémices de la vie se seraient construits dans un habitat volcanique aquatique. Dans ces environnements ayant abrités les premières formes de vie, l'arsenic se trouvait en quantité importante. L'As III devait alors être un substrat énergétique pour les premiers organismes chimolithotrophes. La colonisation de cet environnement primitif et l'oxydation de l'As III en As V, moins disponible aurait détoxiqué ces milieux hostiles, ces bactéries chimolithotrophes ont donc « repoussé les limites de la vie » pour permettre à d'autres types de vie d'émerger et d'évoluer (Oremland & Stolz, 2003) (Hoven *et al.*, 2004) (Muller *et al.*, 2007). Les travaux de Lebrun et collaborateurs tendent à montrer que le système d'oxydation aurait existé avant la différenciation entre les procaryotes et les eucaryotes (Lebrun *et al.*, 2003). Le système d'oxydation de l'arsenic serait donc un système très ancien qui aurait existé chez le dernier ancêtre commun LUCA (*Last universal common ancestor*) (Baymann *et al.*, 2003).

8. Conclusions

Tous les intervenants de l'oxydation n'ont pas encore été mis en évidence. En effet, des travaux récents tendent à montrer que compléter une souche d'*Escherichia coli* non oxydante par les gènes *aoxRSAB-cytc₂* de *A. tumefaciens* (Kashyap *et al.*, 2006) ou par les gènes *aoxABCD* d'*H. arsenicoxydans* (Koechler, Cleiss et Delavat, données personnelles) n'est pas suffisant pour procurer à cette souche un phénotype arsénite-oxydant. D'autre part les gènes *aox* n'ont pas été retrouvés sur le génome d'*Alkalilimnicola erlichii*, malgré sa capacité à oxyder l'As III en As V en condition anaérobie (Hoeft *et al.*, 2007). Ceci suggère que d'autres intervenants du système *aox* et qu'au moins un autre système d'oxydation de l'arsénite existent et restent à découvrir.

III REDUCTION DE L'ARSENATE EN ARSENITE

De façon générale, l'arsénate est plus aisément adsorbé que l'arsénite et est donc considéré comme moins mobile. Ainsi la réduction de l'As V en As III impacte ce cycle biogéochimique de l'arsenic et son transport dans la nature. Deux types de réduction de l'arsénate ont été décrits : la réduction cytoplasmique qui permet de produire de l'As III qui sera alors transporté hors de la cellule, et la réduction dissimilatrice qui permet d'utiliser l'As V comme accepteur final d'électron.

1. La réduction cytoplasmique de l'As V et le système d'efflux

Les systèmes de résistance à l'arsénate ont été décrits pour la première fois chez *E. coli* et *S. aureus* comme des gènes inductibles situés sur des plasmides et présentant une organisation opéronique (Hedges & Baumberg, 1973) (Silver *et al.*, 1981). De façon générale, les protéines du système Ars ont la fonction de « sentir » et de « détoxifier » l'arsénate. Ces gènes sont largement répandus chez les souches à Gram positif et à Gram négatif (Páez-Espino *et al.*, 2009). Outre leur localisation plasmidique, ils sont aussi retrouvés associés à des transposons (Neyt *et al.*, 1997) et/ou sur le chromosome bactérien (Carlin *et al.*, 1995).

Suite à son entrée dans la cellule, l'As V va être réduit dans le cytoplasme en As III grâce à l'enzyme cytoplasmique l'arsénate-réductase ArsC. L'As III sera ensuite excrété hors de la cellule *via* la/les protéine(s) transmembranaire(s) ArsB/ArsA (ou Acr3).

a. ArsC : La réductase

L'arsénate-réductase ArsC est une petite protéine monomérique de 16 kDa et d'environ 135 acides aminés appartenant à la classe des tyrosine-phosphatases ; son rôle est la réduction de l'arsénate intracellulaire en arsénite (Ji & Silver, 1992a) (Rosen, 1999).

Il existe deux familles d'ArsC. La première est issue de l'enzyme codé par *arsC* d'*E. coli* (plasmide R773). Elle contient des résidus cystéines et nécessite l'interaction avec le glutathion (GSH) et la glutarédoxine (Grx1, Grx2 ou Grx3) (Shi *et al.*, 1999).

L'ion arsénate va se lier au site de fixation, composé de trois résidus basiques : Arg60, Arg94 et Arg107 (les ions sulfate et phosphate sont des compétiteurs de l'arsénate pour ce site de liaison). L'arsénate va alors former un intermédiaire covalent arsénate thioester avec le site actif et sera ensuite réduit en deux étapes par la glutarédoxine et le glutathion (Figure I2.9) (Rosen, 2002b) (Rosen, 2002a).

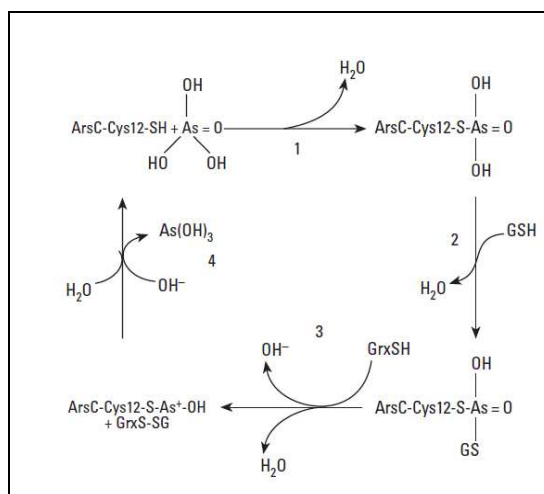


Figure 12.9 : Le mécanisme réactionnel d'ArsC. Dans la première étape l'As V se lie de façon non covalente à l'enzyme. Il y a alors formation d'une liaison thiol entre Cys12 de ArsC et l'As V. Puis le GSH va se lier pour former un intermédiaire ArsC-Cys12-S-As-GS. Grx va alors réduire l'intermédiaire et fixer le GS. L'enzyme se retrouve lié à l'As III, le complexe va s'hydrolyser pour régénérer l'ArsC et libérer l'As III (Extrait de (Mukhopadhyay & Rosen, 2002)).

La seconde famille est composée des enzymes proches de celle décrite chez *S. aureus* (plasmide pI258). Dans ce cas l'enzyme utilise la thiorédoxine comme donneur d'électron et ne présente pas les mêmes sites catalytiques (l'arsenic est lié à une cystéine en N-terminal de ArsC) (Rosen, 1999) (Stolz *et al.*, 2006).

La réductase a d'autre part été caractérisée chez certains eucaryotes. Chez *S. cerevisiae* le gène impliqué dans la réduction a été nommé *acr2p* (Mukhopadhyay & Rosen, 2002). Un gène équivalent a été décrit chez l'homme dans les cellules hépatiques (Radabaugh & Aposhian, 2000).

b. ArsB/Acr3 : la pompe d'efflux

Suite à la réduction de l'As V en As III par l'ArsC, l'arsénite sera pris en charge par le système d'efflux.

Deux familles de gènes codants pour ce système d'efflux ont été identifiées :

1) La famille ACR3 dont les gènes ont été identifiés pour la première fois chez *Bacillus subtilis* (Beloin *et al.*, 1997). Les gènes *acr3* sont principalement présents chez les actinobactéries et les alpha-protéobactéries (Páez-Espino *et al.*, 2009). Chez les eucaryotes, un système similaire intervient dans la résistance à l'arsénite : Acr3p correspond à la pompe d'efflux chez *S. cerevisiae* (Wysocki *et al.*, 1997). Acr3 est un transporteur spécifique de l'arsénite, c'est une pompe à efflux qui fonctionne en utilisant le potentiel de membrane. Il est composé, selon les données de séquence, de 10 hélices

transmembranaires. D'autre part les gènes *acr3* sont très largement répandus dans l'environnement (Achour *et al.*, 2007).

2) La famille des ArsB appartient au groupe des MRP (*multidrug resistance-associated protein*) correspondant aux protéines de résistance multidrogues et au sous-groupe des ABC-transporteurs (*ATP-binding cassette*). Ces gènes sont principalement retrouvés chez les Firmicutes et les gamma-protéobactéries (Páez-Espino *et al.*, 2009).

ArsB est une protéine de transport de 45 kDa, composée de 429 résidus, qui catalyse l'extrusion de l'arsénite et de l'antimonite (Rosen, 1999). La protéine est enchâssée dans la membrane interne d'*E. coli* avec 12 segments transmembranaires (hélices α), cinq boucles cytoplasmiques et six boucles périplasmiques (Figure I2.10).

ArsB est généralement couplée à ArsA pour former le complexe ArsAB qui permet de pomper l'arsénite hors de la cellule grâce à l'action ATPasique d'ArsA (Figure I2.10). Néanmoins ArsB est capable d'exporter l'As III en absence d'ArsA, mais de façon moins efficace. Il fonctionne alors comme un antiport métalloïde/proton utilisant le potentiel de membrane (Dey & Rosen, 1995) (Meng *et al.*, 2004), comme par exemple chez *S. aureus* (Ji & Silver, 1992b) (Ji & Silver, 1992a) ou *Thiobacillus ferrooxidans* (Butcher *et al.*, 2000).

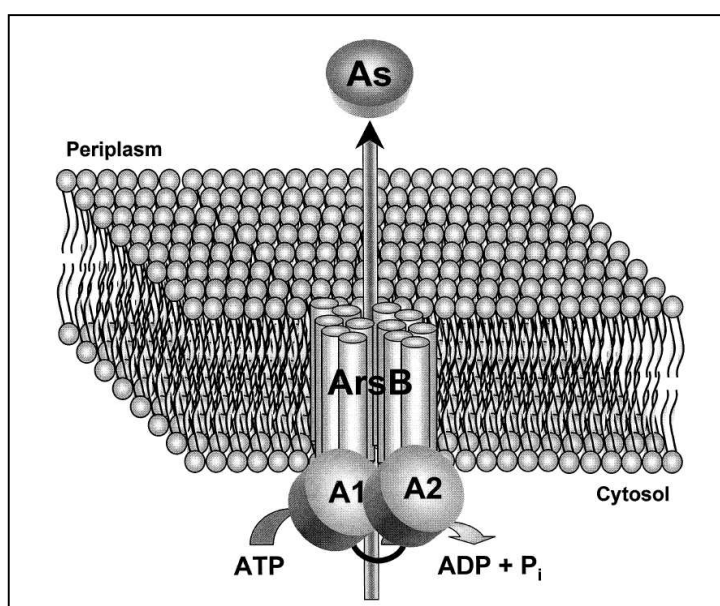


Figure I2.10 : La pompe ArsAB. Le complexe ArsA et ArsB forme une ATPase qui catalyse l'extrusion de l'arsénite et de l'antimonite. ArsB est une protéine de la membrane interne chez *E. coli* et sert d'ancrage à la membrane pour ArsA et au transport d'anion. ArsA est composé de deux sous-unités A1 (N-terminal) et A2 (C-terminal). ArsA est la sous-unité catalytique ATPasique stimulé par l'As III/Sb III (Extrait de (Rosen *et al.*, 1999)).

Récemment un système différent de transport de l'arsénite a été mis en évidence chez *Campylobacter jejuni*, (en plus du gène *acr3*). Il s'agit du gène *arsP* qui présente une faible homologie de séquence avec les *arsB* déjà connus. ArsP correspondrait à une perméase de la famille MFS (*major facilitator superfamily*). Il possède huit hélices transmembranaire prédites et une grande boucle hydrophile extracellulaire. ArsP est codé par le gène *arsP* situé au sein de l'opéron *ars* et est contrôlé par ArsR. Cependant aucune preuve expérimentale n'a permis de mettre en évidence l'implication de ce gène dans la résistance de la souche (Wang *et al.*, 2009).

c. ArsA : l'ATPase

La protéine ArsA (63 kDa – 583 résidus) est une ATPase dont la synthèse est activée par l'arsénite et l'antimonite et présentant une haute affinité pour l'As III (Rosen *et al.*, 1999) (Zhou *et al.*, 2000). Elle est composée de deux domaines identiques A1 (N-terminal) et A2 (C-terminal) de 30 kDa chacun connecté par un peptide *linker* de 25 acides aminés.

L'arsenic se fixe sur trois sites catalytiques composés des résidus suivants : site 1 : Cys-113 (A1) et Cys-422 (A2), sites 2 : Cys-172 (A1) et His-453 (A2) et site 3 : His-148 (A1) et Ser-420 (A2) (Figure 12.11) (Rosen *et al.*, 1999). Cette fixation va alors générer une modification allostérique et rapprocher les domaines A1 et A2 de la protéine. Une fois le substrat fixé, les domaines sont liés et, à l'autre extrémité de la protéine, deux sites de liaison nucléotidiques (NBS : *nucleotide-binding site*) se retrouvent en contact. Le rapprochement de ces deux sites NBS est nécessaire à l'activité ATPasique de la protéine (Rosen *et al.*, 1999) (Zhou *et al.*, 2000) (Bhattacharjee *et al.*, 2008). L'activité ATPasique est activée par la fixation du métalloïde (Rosen, 2002b). En absence de l'activateur (As III ou SbIII), ArsA présente une activité basale faible ATPasique joué par le site NBD du domaine A1 (Rosen, 2002b).

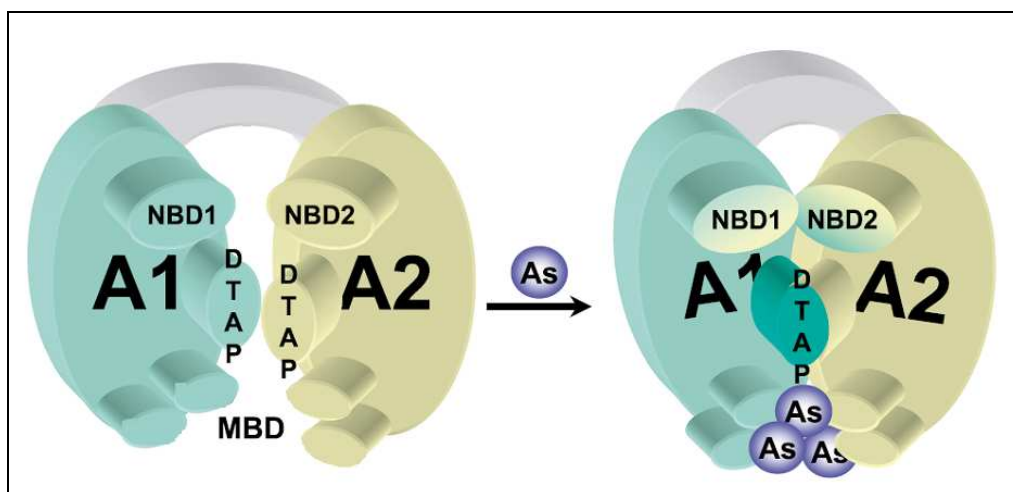


Figure I2.11 : Modèle de l'activation allostérique de ArsA par l'As III. En l'absence d'As III, les domaines A1 et A2 de ArsA ne sont pas en contact. La liaison d'As III ou de Sb III au site de liaison (MBD : *metal binding site*) colle les deux domaines A1 et A2 de ArsA ensemble, ce qui joint les sites NBD (*nucleotide binding site*) et accélère la catalyse (Extrait de (Rosen, 2002b)).

La protéine présente deux séquences signatures qui servent de domaines de transduction du signal $D_{142}TAPTGH_{148}TIRLL$ dans (A1) et $D_{447}TAPTGH_{453}TLLLL$ en A2. Ces domaines seraient impliqués dans la transmission de l'énergie, issue de l'hydrolyse de l'ATP, au transporteur du métalloïde (Ruan *et al.*, 2006).

Il est intéressant de noter que l'activateur de ArsA est en fait la même molécule qui est transportée par la pompe ArsAB. Néanmoins il n'existe pas de preuve que l'ion qui active le transporteur est également celui qui est directement transporté par la pompe (Zhou *et al.*, 2000).

d. ArsRD : les protéines régulatrices

ArsR (13k Da) est un régulateur transcriptionnel qui contrôle le niveau basal de l'expression de l'opéron *ars* (Ji & Silver, 1992b). C'est un dimère qui se fixe en trans sur le promoteur de l'opéron *ars*. Il se fixe sur la région -64 à -31 du promoteur chez *E. coli* (Xu *et al.*, 1996a), l'ARN polymérase se fixe sur les régions -50 à +20, la fixation du dimère ArsR interfère donc avec l'initiation de la transcription (Wu & Rosen, 1993).

Les protéines de régulation ont une haute affinité pour leur substrat, en particulier ArsR doit répondre à de faibles concentrations de substrat (Qin *et al.*, 2006). La liaison de l'As III sur ArsR induit un changement conformationnel du domaine hélice-tour-hélice de fixation à l'ADN et conduit à une dissociation du répresseur sur le promoteur de l'opéron et permet sa transcription (Rosen *et al.*, 1999). Le site consensus du répresseur présente, chez *E. coli*, la structure suivante: TCAT-N₇-TTTG (Xu *et al.*, 1996a) (Xu *et al.*, 1996b).

Il existe une certaine variabilité de séquences puisque les souches *Pseudomonas* sp. (Patel *et al.*, 2007) ou *Acidithiobacillus ferrooxidans* (Butcher & Rawlings, 2002) ne présentent pas la séquence consensus de site de liaison à l'arsenic, même si les protéines ArsR sont relativement proches les unes des autres au sein de la famille des ArsR, composée aussi des répresseurs de systèmes de résistance à d'autres métaux comme Zn II, Cd II et Pb II.

ArsD (13 kDa et 120 résidus) correspond à un deuxième régulateur transcriptionnel indépendant de ArsR. Contrairement à AsrR qui contrôle le niveau basal de l'expression de l'opéron *ars*, ArsD est un homodimère qui contrôle le niveau haut de transcription ; ceci probablement pour éviter une surexpression d'ArsB qui est toxique à de fortes concentrations (Rosen, 1999). ArsD est un chaperon qui se lie à l'As III cytoplasmique et le livrerait directement au site de liaison de ArsA (Silver & Phung, 2005) (Lin *et al.*, 2006).

e. Quel rôle pour ArsH ?

Décrit pour la première fois chez *Yersinia enterocolitica* (Neyt *et al.*, 1997) où il est essentiel à la résistance de la souche, le gène *arsH* est retrouvé chez toutes les souches à Gram négatif contenant un opéron *ars* (Páez-Espino *et al.*, 2009), entre autres les souches *T. ferrooxidans* (Butcher *et al.*, 2000), *Pseudomonas putida* (Canovas *et al.*, 2003), *Pseudomonas* sp. TS44 (Cai *et al.*, 2009b), ... Pour certaines d'entre elles, il a été montré que ce gène n'est pas nécessaire à leur résistance à l'arsenic. Neyt et collaborateurs suggèrent que ArsH serait un régulateur transcriptionnel (Neyt *et al.*, 1997). Cependant elle a été cristallisée par Ye et collaborateurs (Ye *et al.*, 2007) qui ont montré qu'il s'agit d'une NADPH-flavin mononucléotide oxidoreductase de 24,6 kDa. Son rôle dans le système Ars reste donc encore inconnu (Ye *et al.*, 2007).

f. L'opéron ars

Les gènes *ars* sont fréquemment retrouvés dans les génomes de nombreuses souches, les opérons y présentent des structures différentes et il n'est pas rare de retrouver plusieurs opérons *ars* dans un même génome avec ou non les mêmes compositions. C'est par exemple le cas de la souche *H. arsenicoxydans* qui possède quatre opérons *ars* avec des structures opéroniques différentes présentées dans la Figure I2.12.

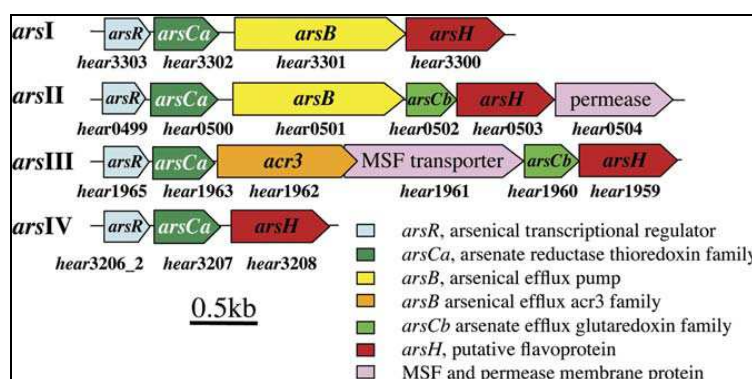


Figure I2.12 : Organisation des opérons *ars* chez *H. arsenicoxydans* (Extrait de (Muller *et al.*, 2007))

g. Evolution des gènes *ars*

Dans le milieu originel, l'arsenic devait être l'une des espèces chimiques prépondérantes, ainsi les systèmes de résistance à l'arsenic devaient déjà être présents dans les premiers organismes. Quand l'atmosphère devint oxydante, l'accumulation d'arséniate a conduit à l'émergence d'enzymes permettant de réduire l'As V en As III qui était alors évacué de la cellule *via* les systèmes d'extrusion déjà présents.

Ainsi il est aisé d'imaginer l'histoire évolutive de l'opéron *ars* : *ArsB* est suffisant pour conférer la résistance à l'As III, la structure ancêtre était donc composée uniquement du gène *arsB*. Il se serait ensuite couplé au gène *arsR* codant le régulateur, l'opéron devient alors *arsRB*. Cette structure opéronique de deux gènes existe toujours aujourd'hui chez *Bacillus subtilis* (Beloin *et al.*, 1997). Suite à l'augmentation de la concentration en arséniate (dus aux conditions oxydantes) un gène de réduction de l'As V en As III est donc requis, *arsC*. Lors de l'acquisition de ce gène *arsC*, des divergences ont dû apparaître puisqu'il existe aujourd'hui deux familles de gènes *arsC*. Enfin la pompe *ArsAB* présente un avantage évolutif, puisque les opérons *arsRADBC* confèrent un niveau de résistance plus élevé que les opérons à trois gènes. En effet, le couple *ArsAB* est plus efficace que la seule pompe *ArsB*. Il y aurait donc une pression évolutive pour acquérir *ArsA*. Cette acquisition de *arsA* ainsi que celle de *arsD* devait être tardive dans l'histoire évolutive de l'opéron (Rosen, 1999).

2. La réduction dissimilatrice

La réduction dissimilatrice correspond à l'utilisation de l'As V comme accepteur final d'électrons. Cette réaction est couplée à un donneur d'électron qui peut être organique comme l'acétate, le lactate ou des composés aromatiques comme le toluène, le phénol... ou inorganique comme l'hydrogène ou le sulfure. La réduction de l'As V en As III a un rendement de -172 kJ/mol sur lactate

a. Les souches

La respiration de l'As V est décrite chez quelques procaryotes, des archées, des bactéries à Gram positif, des γ -, δ -, ϵ -protéobactéries ; comme *S. barnesii* SES-3 (Stolz *et al.*, 1999), *B. arsenicoselenatis*, *Desulfotomaculum auripigmentum* (Newman *et al.*, 1997b), *Chrysiogenes arsenatis* (Stolz & Oremland, 1999), *Desulfovibrio* sp., *Desulfomicrobium* sp. (Macy *et al.*, 2000), *Bacillus* sp. JMM-4 (Santini *et al.*, 2002b), *Shewanella* sp. (Malasarn *et al.*, 2004), *Halarsenatibacter silvermanii* SLAS-1 (Blum *et al.*, 2009),... La souche arsénite-oxydante chimiolithoautotrophe *Thermus* HR13 est capable d'alterner entre l'oxydation de l'As III et la réduction de l'As V dans un processus de respiration durant les périodes d'anoxie, elle utilise l'As V comme accepteur d'électrons et le lactate comme donneur d'électrons (Gihring & Banfield, 2001).

b. La protéine

Seuls deux arséniate-réductases respiratoires, Arr, ont été purifiées et caractérisées chez *C. arsenatis* (Krafft & Macy, 1998) et *Bacillus selenitireducens* (Afkar *et al.*, 2003). Chez ces deux bactéries, la protéine Arr est constituée de deux sous-unités ArrA (87 kDa, 234 résidus) et ArrB (29kDa, 854 résidus). L'enzyme présente une conformation hétérodimérique $\alpha\beta$, alors que la réductase de *S. barnesii* est membranaire hétérotrimérique, $\alpha\beta\gamma$ (65 kDa, 31 kDa et 22 kDa respectivement). ArrAB est une molybdoenzyme retrouvée soit liée à la membrane (*B. selenitireducens* ou *S. barnesii*), soit périplasmique (*C. arsenatis*) (Afkar *et al.*, 2003) (Krafft & Macy, 1998). Appartenant à la famille DMSO réductase, elle intervient dans le métabolisme énergétique. Elle est constituée d'une grande sous-unité à molybdoptérine, ArrA, qui contient un centre fer-souffre, probablement [4Fe-4S]. ArrA présente dans sa séquence le motif (C-X₂-C-X₃-C-X₂₇-C) qui permet la liaison au centre fer-souffre et à sa partie N-terminal une séquence signale TAT. Elle est également constituée d'une petite sous-unité à centre [Fe-S], ArrB qui est homologue à la petite sous-unité Rieske de AoxB (Figure I2.13) (Silver & Phung, 2005). Les cofacteurs impliqués sont le molybdène, le fer et le soufre. Enfin, le complexe ArrAB est probablement lié à un cytochrome b (Stolz & Oremland, 1999) (Santini *et al.*, 2002b).

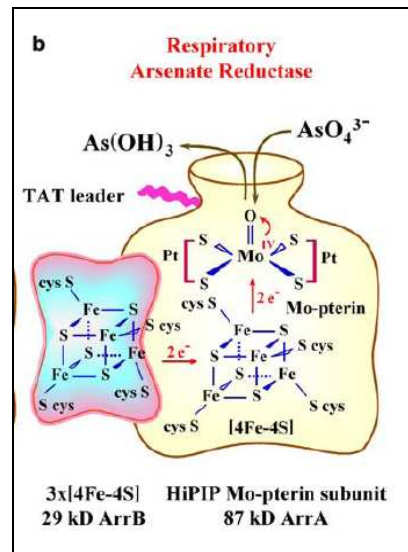


Figure I2.13 : Modèle de l'arséniate-réductase respiratoire (Extrait de (Silver & Phung, 2005)).

L'organisation générale est la même que celle de l'arsénite-oxydase. Deux différences notoires existent pourtant : le transfert d'électrons se fait dans le sens opposé puisqu'il s'agit là d'une réductase et la présence d'une séquence TAT sur la grande sous-unité ArrA et chez la petite sous-unité AoxA (Silver & Phung, 2005).

Comme pour *Shewanella*, les gènes *arr* sont fréquemment retrouvés directement en aval de l'opéron *ars* (Figure I2.14). De façon générale, ces gènes sont régulés et induits par l'As III (Krafft & Macy, 1998).

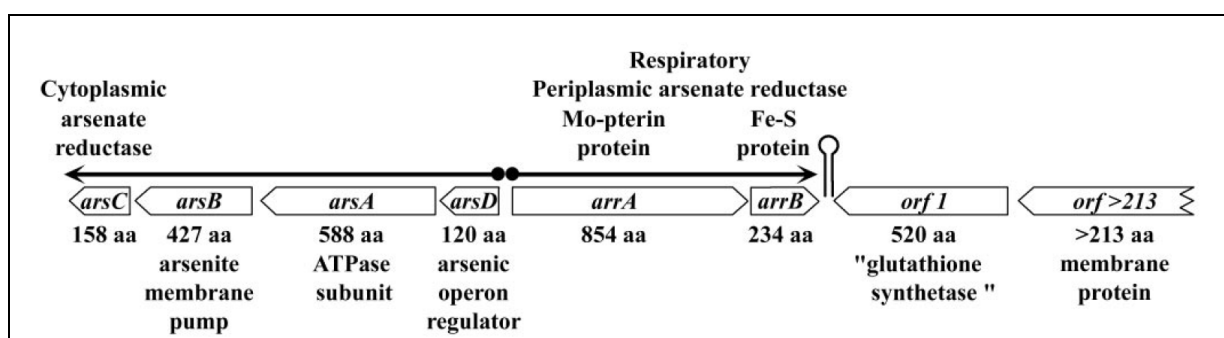


Figure I2.14 : Gènes codant pour l'arséniate réductase respiratoire (*arrAB*) et cytoplasmique (*arsABCD*) de *Shewanella* souche ANA-3. aa : acides aminés (extrait de (Silver & Phung, 2005)).

Récemment les données de génomes ont révélé une hétérogénéité dans le nombre de gènes codant la protéine ainsi que son système de régulation. La souche MLMS-1 par exemple présente un troisième gène qui a été nommé *arrC* et code une troisième sous-unité de l'enzyme. Les auteurs ont

d'autre part mis en évidence chez cette souche la présence de gènes indiquant un système de régulation à deux composants (Stolz *et al.*, 2006).

Enfin les protéines ArrA forment un groupe à part dans la famille des DMSO réductases. On remarque un haut degré de conservation entre les ArrA connus jusqu'alors. Le gène correspondant peut alors être utilisé comme marqueur de la respiration de l'As V (Malasarn *et al.*, 2004) (Figure 12.15).

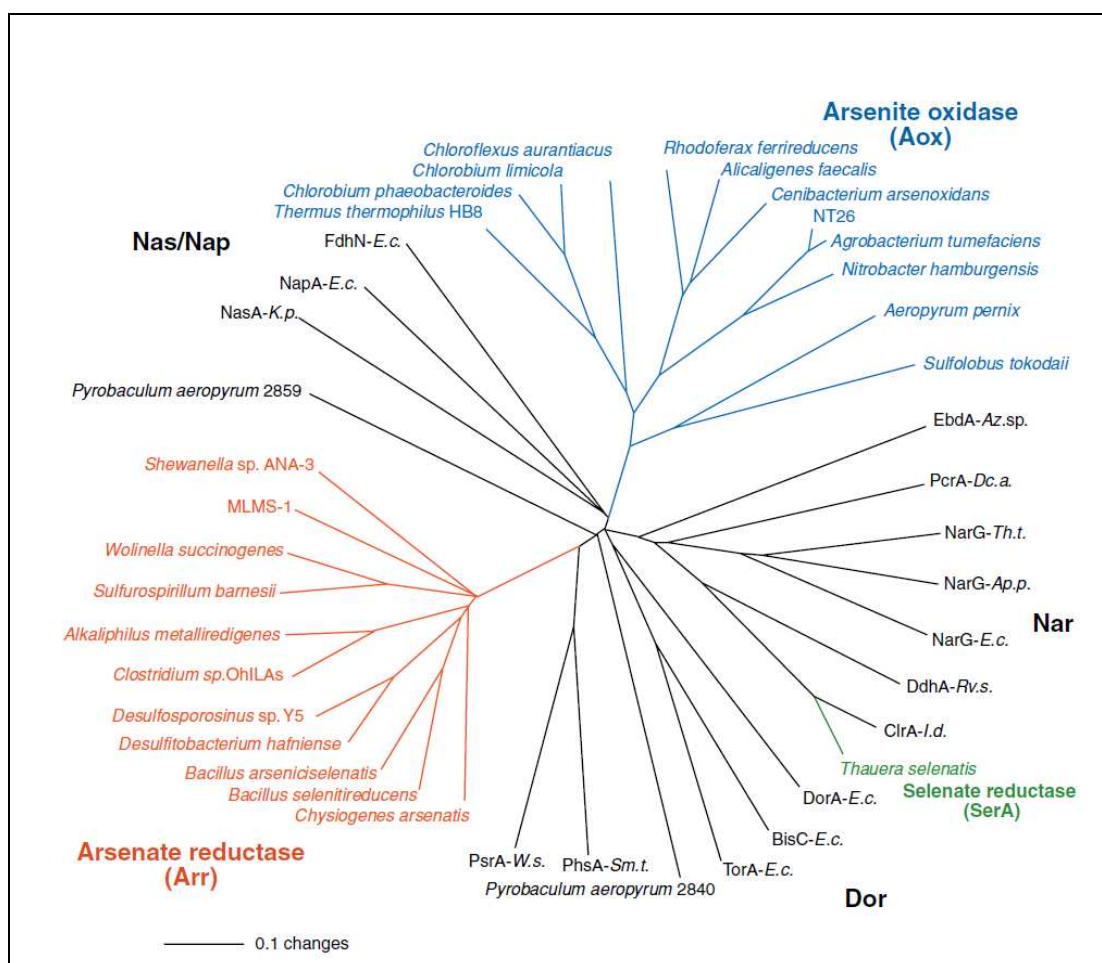


Figure 12.15 : Arbre phylogénétique (non-enraciné neighbor joining) des enzymes à molybdène de la famille DMSO réductase. L'arséniate réductase (Arr, orange) et l'arsénite-oxydase (Aox, bleu) sont séparées et sur des branches distinctes. Les autres principaux clades dans la famille sont les DMSO réductases (Dor), nitrate réductase périplasmique et nitrate réductase assimilatrice (Nas/Nap) et nitrate respiratoire réductase (Nar) dont la séniate réductase respiratoire (SerA, vert). Abréviations : Ap. p. : *Aeropyrum pernix*, K. p. : *Klebsiella pneumoniae* ; Sm. t. : *Salmonella typhimurium* ; Th. t. : *Thermus thermophilus* ; W. s. : *Wolinella succinogenes*. *Pyrobaculum aeropyrum* 2859 et 2840 seraient des homologues Arr. A. sp. EB1 : *Azoarcus sp. EB1* ; A. f. : *Alcaligenes faecalis* ; D. d. : *Desulfovibrio desulfuricans* ; D. a. : *Dechloromonas agitata* ; E. c. : *Escherichia coli* ; I. d. : *Ideonella dechloratans* ; R. sp. : *Rhodobacter sphaeroides* ; R. su. : *Rhodovulum sulfidophilum* ; S. m. : *Shewanella massilia* ; T. s. : *Thauera selenatis* (Extrait de (Stolz *et al.*, 2006)).

c. Réversibilité des fonctions de la protéine ArrB

Récemment Richey et collaborateurs ont mis en évidence qu'*Alkalilimnicola ehrlichii* ne possède pas de gène *aox* dans son génome alors qu'elle est physiologiquement capable d'oxyder l'As III en As V. Les auteurs ont alors montré que cette oxydation est réalisée par le système Arr. Ils ont donc découvert la réversibilité de l'enzyme chez *Shewanella* (oxydation d'As III et réduction d'As V) (Richey *et al.*, 2009).

Ceci qui implique que ArrB peut transférer les électrons dans les deux sens : de ArrA vers As V ou d'As III vers ArrA.

Deux autres gènes *arr* ont également été mis en évidence, *arrC* qui coderait une unité d'ancrage à la membrane homologue avec NrfD et *arrD* qui pourrait coder un chaperon homologue à TorD (Richey *et al.*, 2009).

IV METHYLATION

La biométhylation correspond à la formation de composés méthylés, volatiles ou non, de métaux et de métalloïdes. La méthylation de l'arsenic est très largement répandue dans la nature, elle a longtemps été considérée comme un processus de détoxication, mais de récentes études montrent que certaines formes méthylées de l'arsenic seraient plus toxiques que les formes inorganiques (comme exposé dans le chapitre I). De nombreuses études ont été menées sur ce processus chez les champignons ou chez d'autres eucaryotes, mais peu d'entre elles ont été menées chez les procaryotes (Páez-Espino *et al.*, 2009). Le rôle de la méthylation de l'arsenic chez les procaryotes reste encore non élucidé.

Ces mécanismes de méthylation ne seront détaillés ici que partiellement, puisque ces concepts ne sont pas nécessaires à la compréhension de l'exposé.

De façon générale, la méthylation inclut une série d'étapes où la réduction de la forme As V est suivi d'une addition oxydative d'un groupe méthyl. La réaction se produit entre le S-adénosyl méthionine (donneur du groupement méthyl) et l'arsénite. L'arsénite est alors réduit en méthylarsoniate. La production de méthylarsine est quant à elle réalisée en présence de thiorédoxine réduit (Figure I2.16) (Silver & Phung Ie, 2005). Les composés méthyls formés sont listés dans la Table I2.T1.

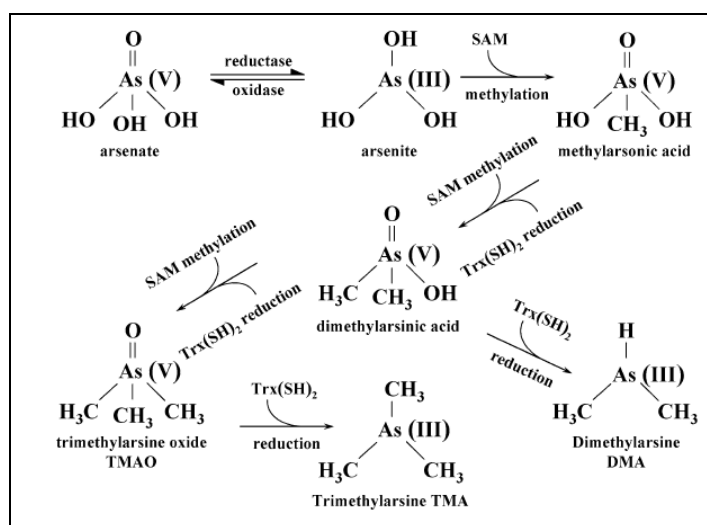


Figure I2.16 : Etapes hypothétiques de la méthylation de l'arsenic inorganique (Extrait de (Silver & Phung le, 2005)).

Table I2.T1 : Nom et abréviations des différentes formes d'arsenic méthylées (Extrait de (Bentley & Chasteen, 2002))

Groupe et composés	Abréviations	Formules
Pentavalent		
Arséniate	As V	$(\text{HO})_2\text{AsO}(\text{OH})$
Méthylarsoniate (acide méthylarsonic ou acide méthylarsinic ou acide monométhylarsonic)	MMA V et MMAA	$(\text{CH}_3)(\text{OH})\text{AsO}(\text{OH})$
Diméthylarsiniate (acide diméthylarsinic ou acide diméthylarsonic)	DMA V et DDMMA	$(\text{CH}_3)_2\text{AsO}(\text{OH})$
Trivalent		
Arsénite	As III	$(\text{HO})_2\text{As}(\text{OH})$
Méthylarsonite (acide monométhylarsonous)	MMA III	$(\text{CH}_3)(\text{OH})\text{As}(\text{OH})$
Diméthylarsinite (acide diméthylarsinous)	DMA III	$(\text{CH}_3)_2\text{As}(\text{OH})$

D'autres formes méthylées d'arsenic sont volatiles, c'est le cas par exemple du mono-, di- et tri-méthylarsine ou du diméthyléthylarsine. Ce processus de volatilisation de l'arsenic, découvert chez *Methanobacterium bryantii*, permet l'élimination de l'arsenic et peut donc illustrer le rôle prépondérant des micro-organismes dans le cycle biogéochimique de cet élément. Les bactéries impliquées dans la méthylation de l'arsenic sont entre autres *Methanobacterium formicium*, *Clostridium collagenovorans*, *Desulfovibrio gigas* ou *D. vulgaris* (Páez-Espino *et al.*, 2009).

ArsM (29 kDa et 283 acides aminés) correspond à une As(III) S-adenosylmethyltransférase et a récemment été montré comme étant impliqué dans la synthèse du gaz triméthylarsine chez *Rhodopseudomonas palustris* (Qin *et al.*, 2006). Le gène correspondant *arsM* est retrouvé dans de nombreuses souches bactériennes (Qin *et al.*, 2006) (Yuan *et al.*, 2008).

V SEQUESTRATION

Il est largement reconnu que les bactéries qui colonisent une surface au sein d'un biofilm sont plus résistantes aux molécules toxiques (entre autre l'arsenic) que celles ayant adopté un mode planctonique (Harrison *et al.*, 2005) (Harrison *et al.*, 2007). La production de biofilm ou la sécrétion de polysaccharides suite à un stress lié à l'oxygène, la prédation, la dessiccation, la présence de substance bactéricide ou encore la présence de métaux lourds constitue un mécanisme courant de défense développé par les bactéries (Stewart, 1994) (Teitzel & Parsek, 2003) (Harrison *et al.*, 2007). Les biofilms sont un assemblage de cellules se développant sur une surface souvent solide, en général recouverte d'une phase aqueuse. Ainsi dans certaines régions de la structure, les bactéries sont protégées des fortes concentrations en contaminant (Nikolaev & Plakunov, 2007).

L'arsenic ne déroge pas à la règle et certains micro-organismes présentent une augmentation de la production d'exopolysaccharides (EPS) après exposition à ce métalloïde toxique. C'est le cas notamment d'*H. arsenicoxydans* (Muller *et al.*, 2007) (Figure I2.17) et de certaines souches de *Thiomonas* (Michel *et al.*, 2007) (Bryan *et al.*, 2009).

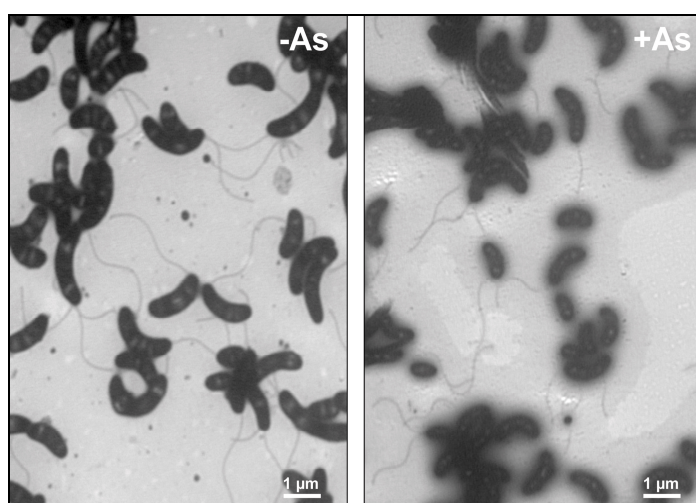


Figure I2.17 : Synthèse d'EPS en présence de 300 ppm d'arsenic par la souche *H. arsenicoxydans*. Les cellules ont été colorées au tétr oxyde d'osmium et observées au microscope électronique à transmission (Extrait de (Muller *et al.*, 2007)).

D'autre part la production d'EPS est liée à la séquestration de certaines molécules (Teitzel & Parsek, 2003). Les résidus chargés négativement à l'extrémité des polymères composant le biofilm affectent la pénétration dans le biofilm des ions métalliques toxiques (Kumar *et al.*, 2007). Les métaux lourds, sous forme de cations, réagissent avec les groupes phosphodiester et carboxylates du biofilm ce qui retarde donc la diffusion des cations métalliques dans le biofilm. Ces cations

métalliques peuvent, d'autre part, se lier aux thiolates, sulfates, phosphates... et sont ainsi réellement séquestrés (Harrison *et al.*, 2007). La séquestration de métaux dans les biofilms a été mise en évidence chez de nombreuses souches et pour de nombreux métaux, citons par exemple le cas de la séquestration de cuivre, fer ou d'or par *P. aeruginosa* PAO1 (Langley & Beveridge, 1999).

La séquestration d'arsenic extracellulaire ou intracellulaire a été mise en évidence pour la première fois chez la souche *Desulfatamaculum auripigmentum*. En présence d'As V et de soufre la souche produit des précipités d'As₂S₃ visible en microscopie électronique comme étant intra- et extra-cellulaire (Figure I2.18A) (Newman *et al.*, 1997a).

De même, Muller et collaborateurs ont mis en évidence la séquestration d'arsenic par les EPS chez la souche *H. arsenicoxydans* (Muller *et al.*, 2007). L'observation des cellules en microscopie électronique à transmission, couplée à la spectrométrie à rayon X à dispersion d'énergie (TEM /EDS), montre la présence de granules riches en arsenic piégés dans la gangue d'EPS (Figure I2.18B).

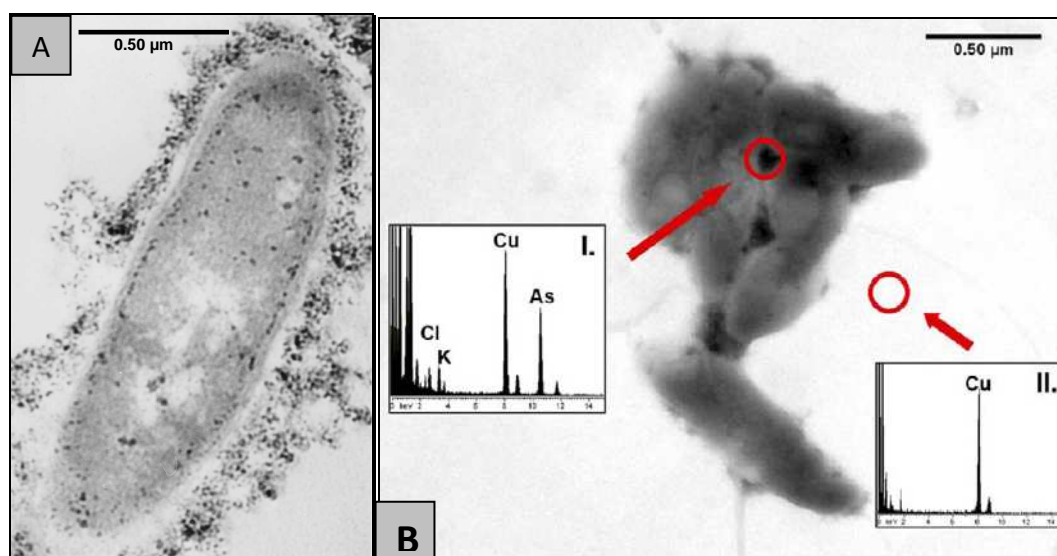


Figure I2.18 : Photographie de microscopie électronique à transmission. A : Précipitation intra et extracellulaire d'As₂S₃ chez *D. auripigmentum*. B : Synthèse d'EPS en réponse à l'arsenic chez *H. arsenicoxydans*. Les ronds représentent les points d'analyse des rayons X. Les spectres Cl et K représentent les constituants organiques et Cu est un constituant de la grille. La teneur en arsenic est de 16,5 % en I. et de 0,0 % en II (Muller *et al.*, 2007).

Il faut noter que la capacité de synthèse des biofilms par les bactéries oxydantes peut être exploitée dans les procédés de biotraitement. C'est ainsi que cette propriété d'adhésion a été utilisée pour la mise au point de réacteurs destinés à traiter de l'eau de surface contaminées par de l'arsenic (Michel *et al.*, 2007).

VI UN EXEMPLE PARTICULIER : LE ROLE DES BACTERIES DANS LES DRAINAGES MINIERS ACIDES (AMD)

1. Origine et chimie

Dans divers endroits de la planète, des rivières ou des lacs sont pollués par des effluents miniers. En 1989, il a été estimé que 19 300 km de rivières et cours d'eau et 72 000 ha de lac et étendues d'eau étaient contaminés par ces effluents. Les exploitations minières augmentent la surface d'exposition à l'air et à l'eau, de divers minerais riches en métaux, ce qui conduit à de fortes contaminations des sols et des eaux de surface. La nature de ces contaminations est très variable et la plus documentée est le cas des drainages miniers acides (AMD). Les AMDs correspondent à des pollutions d'eau de surface liées à une exploitation minière. Celles-ci résultent de la dissolution oxydative des minerais, principalement des sulfures, ce qui conduit à l'acidification du milieu et à l'enrichissement en métaux dissous, principalement en fer et en arsenic. Généralement le pH est acide, situé entre 1 et 4 (Johnson & Hallberg, 2003).

Les AMDs sont retrouvés dans de très nombreuses zones de la planète (Table I2.2) et arborent des caractéristiques physico-chimiques diverses. Néanmoins, la pyrite FeS_2 , très présente dans les zones minières, contient des métaux sous forme d'impureté comme l'or, l'argent, le cuivre, le zinc, le plomb ou l'arsenic. Ces métaux seront alors libérés et dissous avec la pyrite lors de son oxydation (Baker & Banfield, 2003).

Table I2.T2 : Caractéristiques physico-chimiques d'AMD de divers endroits sur la planète. Les concentrations en métaux et sulfate sont en mg/l. – non déterminé (Extrait de (Johnson & Hallberg, 2003)).

	Storwartz mine, Norway ^a	Ynysarwed, Wales ^b	Bullhouse, England ^b	Wheal Jane, England ^b	Killingdal mine dump, Norway ^a	King's mine stream, Norway ^c	Parys mine, Wales ^b	Rio Tinto, Spain ^d	Iron Mountain, California ^e
pH	6.5	6.2	5.9	3.4	2.77	2.75	2.5	2.2	0.5 – 1
Eh (mV)	–	+214	+257	+462	–	–	+685	+450	–
Fe _{total}	1.60	160	61	290	265	172	650	2.3×10^3	$13-19 \times 10^3$
Fe ²⁺	–	140	45	250	–	–	650	1.5×10^3	$13-19 \times 10^3$
Al	0.03	–	1.2	27	38.3	22.5	70	–	$1.4-6.7 \times 10^3$
Mn	1.35	–	15	8	4.24	0.78	10	–	17–120
Cu	0.06	–	–	1.2	5.65	15.8	40	109	120–650
Zn	2.13	–	–	132	61.1	25.4	60	225	700–2600
Sulfate	151	464	–	400	1219	668	1550	10×10^3	$20-108 \times 10^3$

La pyrite (FeS_2) s'oxyde spontanément en présence d'air et d'humidité, ce qui correspond à la première étape de génération d'un AMD. C'est une oxydation abiotique lente qui libère du Fe^{2+} et du 2SO_4^{2-} et qui acidifie le site. Il est ensuite nécessaire de recycler le Fe^{2+} ; ce processus lent est catalysé par certaines bactéries Fe-oxydantes acidophiles ou acidotolérantes. Le Fe^{3+} va ensuite

pouvoir soit permettre l'oxydation de la pyrite dans un second processus abiotique d'autocatalyse qui dégrade la pyrite, régénère le Fe^{2+} et produit de l'acidification, soit s'hydrolyser et ainsi générer une grande quantité d' H^+ . Ces différents processus sont résumés dans la Figure I2.21.

Ainsi la régénération du Fe^{3+} est un mécanisme clef du système. Le fer est présent dans beaucoup d'AMD, pour ne pas dire dans toutes. Il est le métal dominant et son oxydation en Fe-hydroxide lui confère cette coloration rouge caractéristique dans les eaux très acides ou jaune/orange dans les eaux à pH plus élevé (Figure I2.19) (Johnson & Hallberg, 2003).

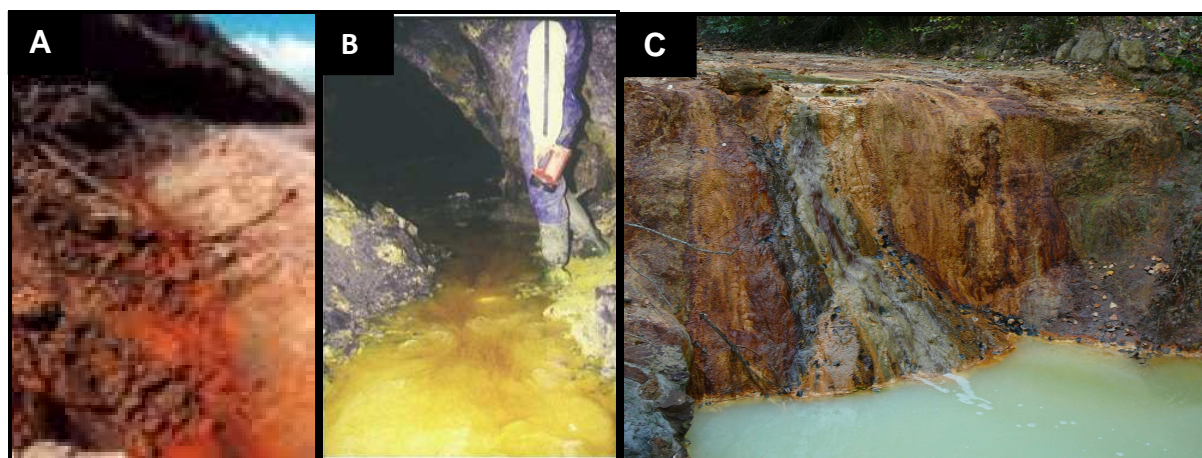


Figure I2.19 : Photographies de différentes AMDs A : Mine d'uranium à Thuringia (Allemagne). B : Ancienne mine de pyrite, Wales (Grande-Bretagne), C : Ancienne mine de plomb, zinc et argent, Carnoulès (France) (Johnson & Hallberg, 2003) (Haferburg & Kothe, 2007).

L'acidité va alors dissoudre d'autres minéraux contenant du cuivre, du zinc, du manganèse ou de l'arsenic, ce dernier est un des contaminants majeurs des AMDs. Il provient essentiellement de l'oxydation de l'arséno-pyrite FeAsS . L'As III et l'As V sont présents, mais la forme plus oxydée (As III) prévaut. Les concentrations en arsenic peuvent être extrêmement élevées dans les AMDs jusqu'à $340\,000\ \mu\text{g/L}$ dans les mines Richmond, USA (Smedley & Kinniburgh, 2002).

L'impact des AMDs sur les rivières ou les cours d'eau peut être vraiment dramatique. Le pH très bas, la présence de métaux et métalloïdes toxiques peuvent conduire à un stress qui induit la disparition des populations indigènes, en particulier des organismes supérieurs. Dans ces milieux extrêmement hostiles à la vie, certains micro-organismes ont su s'adapter et se développer, même si la quantité de matière organique est très faible, $<20\ \text{mg/l}$ et favorise ainsi le développement de bactéries oligotrophes (Johnson & Hallberg, 2003). Les AMDs sont des milieux très divers en termes d'acidité ou de température, ainsi la flore bactérienne sera différente en fonction des conditions.

Dans les paragraphes suivants ne seront considérés que les AMDs ayant des températures de l'ordre de 25°C.

2. Les micro-organismes dans les drainages miniers acides

De nombreuses études ont mis en évidence le rôle prépondérant des bactéries dans les AMDs. Etant donné le caractère extrême des AMDs, ils présentent une faible diversité bactérienne intrinsèque (Tyson *et al.*, 2004).

a. Les souches bactériennes présentes

La flore présente est essentiellement constituée de souches acidophiles. Elles ont été classées en différentes catégories selon leur température optimale de croissance (mésophiles, modérément thermophiles ou thermophiles), leur type de trophisme (autotrophes assimilant le CO₂, hétérotrophes assimilant la matière organique ou mixotrophes assimilant les deux).

Certaines de ces bactéries peuvent accélérer la solubilisation des minéraux par oxydation ou réduction. C'est les cas des fer-oxydantes et soufre-oxydantes qui participent à la dissolution des minéraux soufrés. Elles sont par ailleurs utilisées en biolixiviation pour extraire l'or, le cobalt ou le cuivre.

Les micro-organismes impliqués sont représentés dans la Figure I2.20, ils appartiennent entre autres aux gamma-Protéobactéries : *Acidithiobacillus ferrooxidans*, *A. thiooxidans*, aux beta-Protéobactéries : *Thiomonas sp.*, et aux alpha-Protéobactéries : *Acidiphilum cryptum*, aux Nitrospira : *Leptospirillum ferrooxidans*, aux Firmicutes : *Acidimicrobium ferrooxidans*, *Sulfolobus metallicus* (Fe-oxydante), aux Acidobactéries : *Acidobacterium capsulatum* et (Johnson & Hallberg, 2003). Des archées, comme *Ferroplasma sp.*, ont également été mises en évidence.

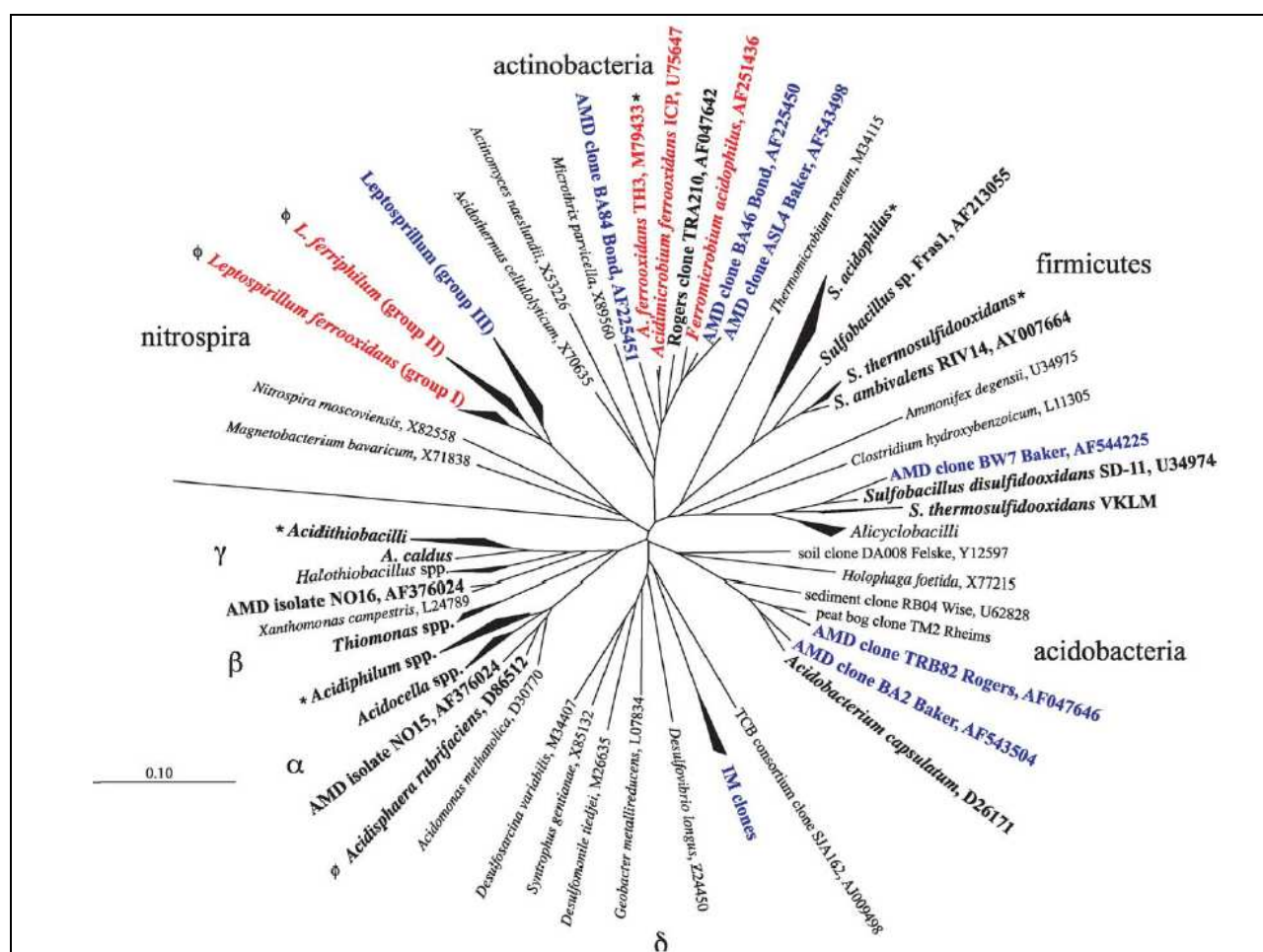


Figure 12.20 : Analyse phylogénétique des gènes 16S rRNA issus d'AMD. *contenant des souches anaérobies facultatives et Φ correspondant à des anaérobies obligatoires. En bleu sont représentés les phyla n'ayant pas de représentants cultivés connus. En rouge les souches sont Fe-oxydantes. L'analyse a été réalisée selon la méthode de neighbor joining. L'échelle représente 0.1 changement par site ou 10 % de différence en séquence. IM = Iron Mountain (Extrait de (Baker & Banfield, 2003)).

Certains eucaryotes peuvent aussi avoir un rôle dans ce type d'écosystème, c'est le cas d'*Euglena mutabilis* (Casiot *et al.*, 2004), de *Cyanidioschyzon sp* (Qin *et al.*, 2009) ou de certains champignons (Baker *et al.*, 2009). Ces organismes génèrent du carbone organique grâce à la fixation de carbone atmosphérique (système photosynthétique) et augmentent, par leur activité, la quantité d'oxygène dissous (Johnson & Hallberg, 2003).

Les communautés indigènes actives

Dans les environnements modérément acide (pH 5-7), on peut retrouver les souches Fe-oxydantes ou les souches qui permettent la précipitation du fer comme *Gallionella ferruginosa* ou *Leptothrix discophora*. Elles se développent préférentiellement en milieu anoxique où la concentration en O₂ est inférieure à 10 %. Le développement de ces bactéries génère également une

diminution du pH dans les AMDs qu'elles colonisent et permet ainsi aux souches plus acidophiles de s'y développer (Baker & Banfield, 2003).

En revanche dans les zones très acides (pH<3) et particulièrement riche en fer, les souches comme *Acidithiobacillus ferrooxidans* et *L. ferrooxidans* sont fréquemment retrouvées. Des bactéries sulfato-réductrices (SRB) ont aussi été mises en évidence ainsi que certaines archées comme *Ferroplasma acidiphilum* et *Thermoplasma acidiphilum* (Baker & Banfield, 2003).

b. Le rôle de ces procaryotes dans la physico-chimie des drainages miniers acides

Dans ces environnements, certaines espèces ou combinaisons d'espèces effectuent l'oxydation du fer, du soufre, de carbone organique, la fixation du carbone, de l'azote, la production d'exopolysaccharides, la réduction du fer et du soufre.

Certaines d'entre elles interviennent également dans les cycles biogéochimiques des métaux et métalloïdes présents. Mais seules quelques souches réalisent ces mécanismes clefs de l'écosystème. Ainsi des interactions trophiques fortes s'établissent entre les souches.

La première étape de génération d'une AMD comprend l'oxydation abiotique de FeS_2 , ce qui libère du Fe^{2+} , du SO_4^{2-} , divers métaux et métalloïdes et génère l'acidification du milieu. Les micro-organismes vont alors se développer dans ces milieux riches en fer et en divers métaux et métalloïdes. Prenons pour exemple le cas d'une contamination à l'arsenic. Les bactéries vont alors oxyder le Fe^{2+} en Fe^{3+} , ce qui régénère la quantité de Fe^{3+} qui pourra ensuite oxyder la pyrite. Dans ces AMDs il y a un cycle du fer qui fait intervenir des réactions biotiques et abiotiques. L'oxydation du Fe^{2+} et Fe^{3+} peut se produire de façon abiotique, mais c'est une réaction lente qui est accélérée d'un facteur 10^6 en présence de certaines bactéries (Banfield *et al.*, 2005a) (Chang *et al.*, 2009). Les bactéries pouvant jouer ce rôle sont par exemple *Thiobacillus*, *Leptospirillum*, *Ferroplasma* (Tyson *et al.*, 2004) ou *Acidithiobacillus ferrooxydans* (Schrenk *et al.*, 1998). Le Fe^{3+} produit lors de ces réactions va aussi réagir avec l'eau ce qui libère des protons et ainsi diminue encore le pH du milieu. Ces processus sont résumés Figure I2.21.

Certaines souches présentes sont autotrophes, elles vont fixer le CO_2 atmosphérique et le N_2 , comme par exemple *A. ferrooxydans* ou *Leptospirillum* sp. (1 dans la Figure I2.21). Ces bactéries produisent de la matière organique utilisée par les organismes hétérotrophes présents.

Lors de l'oxydation de la pyrite il y a aussi libération de SO_4^{2-} (sulfate) qui sera réduit par certaines bactéries sulfato-réductrices. D'autres encore vont oxyder le soufre ; un cycle complet biogéochimique du soufre est en place dans ces AMDs (SR et SO sur la Figure I2.21).

Enfin ces milieux sont extrêmes aussi par leur teneur en métaux et métalloïdes pour lesquels là aussi, des cycles biogéochimiques se mettent en place. Dans le cas de l'arsenic, la forme en solution est principalement l'As III qui sera oxydé en As V par des souches arsénite-oxydantes comme par exemple *Thiomonas* sp. (4 sur la Figure I2.21). L'As V formé va alors se complexer au Fe III (et au soufre) et précipiter sous forme d'arsénopyrite ou de tooeleite par exemple. Ce mécanisme permet de diminuer la teneur en arsenic du milieu et peut constituer un mécanisme de décontamination naturelle du site (Morin *et al.*, 2003).

D'autres bactéries encore peuvent aussi intervenir dans la bioremédiation. Plus particulièrement, les sulfato-réductrices permettent la réduction des métaux en anaérobiose (c'est-à-dire en conditions redox basses) par deux voies différentes : d'une part par la réduction dissimilatrice indirecte *via* la production de S^{2-} et du H_2S . Le S^{2-} qui pourra se complexer à certains métaux et précipiter, (par exemple : $S^{2-} + Zn^{2+} \longrightarrow ZnS$) et d'autre part, par la réduction directe via l'activité de certaines enzymes à bas potentiel redox comme certains cytochromes (Lovley, 2000).

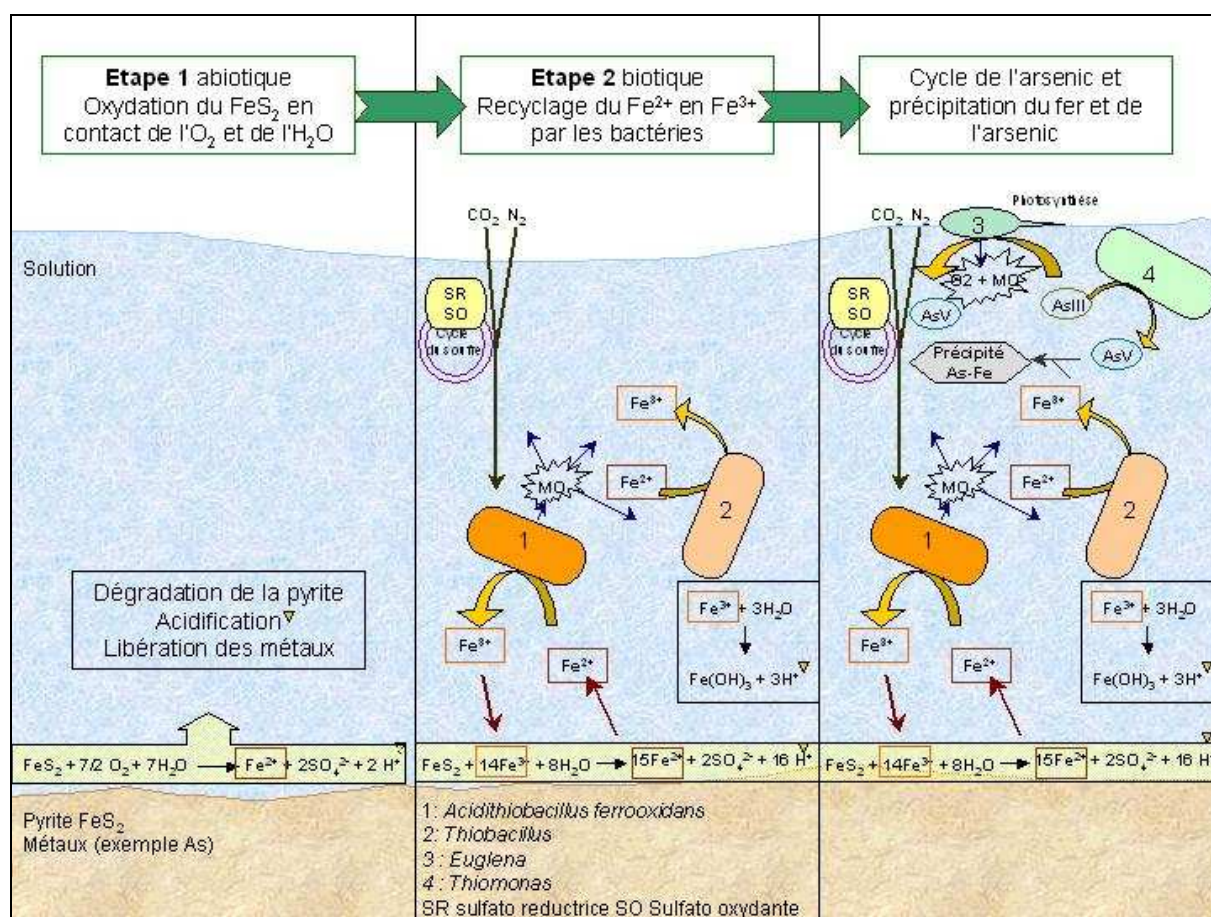


Figure 12.21 : Proposition schématique de développement d'un AMD. Etape abiotique – Etape biotique et induction de la génération de l'AMD – rôle des micro-organismes dans le cycle des métaux contaminants.

Acidithiobacillus ferrooxidans et *Gallionella* sp. sont sans conteste les micro-organismes clefs des AMDs (Bruneel *et al.*, 2006) (He *et al.*, 2007) puisqu'ils interviennent dans le cycle du fer. C'est *Thiomonas* sp. qui joue ce rôle clef en ce qui concerne l'arsenic puisqu'elle est capable dans les AMDs d'oxyder l'As III en As V (Bruneel *et al.*, 2003). La découverte de l'oxydation anoxique bactérienne de l'As III représente une étape importante du cycle biogéochimique dans les environnements aquatiques.

En conclusion les AMDs sont des milieux extrêmes et très diversifiés. La richesse de leur flore bactérienne est corrélée à la diversité de leur composition. De nombreux cycles sont en jeu et les interactions trophiques des différentes souches y sont très fortes et encore peu élucidées. Il est ainsi très difficile de résumer en quelques lignes le développement et le fonctionnement de ces AMDs.

3. Un exemple : la rivière le Reigous à Carnoulès, Gard (France)

Le site de Carnoulès dépend de la commune de St Sébastien d'Aigrefeuille au pied du Massif central dans le Gard (Figure I2.22). Cette mine de plomb, de zinc et d'argent a été exploitée jusqu'en 1963. Les 1,5 Mt de stériles (déchets des minerais extraits) sont aujourd'hui stockés sur le site et contiennent de fortes quantités de sulfure de fer, de plomb et de l'arsenic (Morin *et al.*, 2003). Les stériles arborent des concentrations d'arsenic parmi les plus importantes au monde, jusque 12 000 mg/l (Koffi *et al.*, 2002). La quantité totale d'arsenic dans le stock de stériles est estimée à 3 000 t, ce qui correspond, compte tenu de la masse annuelle d'arsenic rejetée par la source (6 t), à 500 ans de pollution (LeBlanc *et al.*, 2002). Sous l'action combinée de l'air et des eaux de pluie, les affleurements sulfurés s'oxydent ; ce qui conduit à la production d'acide et à la dissolution des métaux présents.

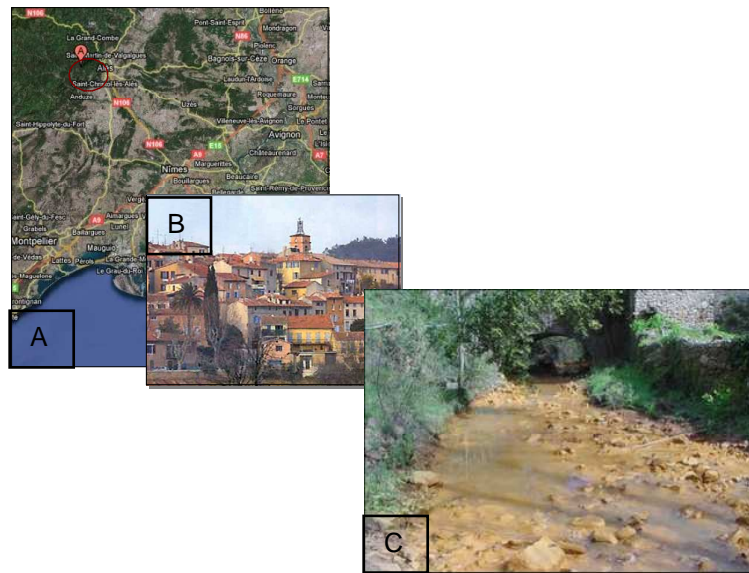


Figure I2.22 : Présentation du site de Carnoulès dans le Gard. A : Carte de localisation du site. B : Photographie de la commune de Carnoulès. C : Photo du cours d'eau contaminé le Reigous (AMD).

Les eaux acides de percolation se jettent dans le Reigous, une source d'eau qui prend naissance sous les stériles. Ce ruisseau est chargé de divers métaux (fer (1 000-2 000 mg/L), plomb (1-2 mg/L), zinc (15-48 mg/L), cadmium (100-200 µg/L), cuivre (50-1 100 µg/L), antimoine (45-120 µg/L), ...) (Bruneel, 2004 thèse) et en particulier d'une très forte quantité d'arsenic dissous de 100 à 350 mg/L, en fonction de la saison (Figure I2.23). Le Reigous se jette dans la rivière Amous ce qui conduit à sa pollution : en aval du confluent, aucune trace d'organismes supérieurs et de vie piscicole n'est retrouvée (Bruneel, 2004 thèse). Les eaux sont acides et présentent un pH de l'ordre de 3 en moyenne.

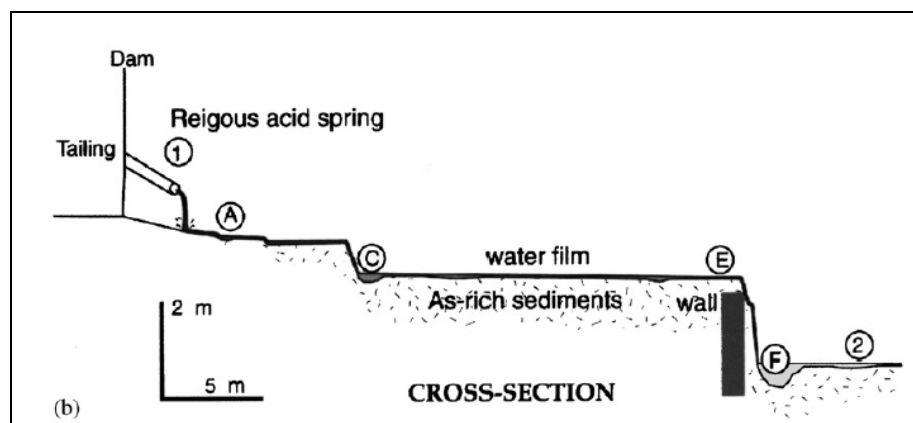


Figure 12.23 : Schéma de la section transversale de la naissance du Reigous sous les stériles miniers jusqu'au point de prélèvement classique COWG (F) 50 mètres en aval (Extrait de (Casiot *et al.*, 2003b)).

Ce site arbore une atténuation biologique naturelle de 95 % entre la source du Reigous et sa confluence avec l'Amous (1.5 km en aval). Dans les 10 premiers mètres une grande quantité de Fe et d'As précipite dans les sédiments qui contiennent alors 20 % d'arsenic (Morin *et al.*, 2003). Cette bioremédiation est le fait des micro-organismes présents qui participent à l'oxydation de l'arsénite en arséniate (*Thiomonas sp.*, *Gallionella sp.*) et à la co-précipitation de l'arséniate avec les ions ferriques FeIII (*Acidithiobacillus ferrooxidans* qui oxyde le Fe II en Fe III) (Leblanc *et al.*, 1996) (Casiot *et al.*, 2003b) (Bruneel *et al.*, 2003) (Duquesne *et al.*, 2008). D'autre part l'As V, à pH acide, s'adsorbe plus facilement aux ions Fe III que l'As III, l'oxydation bactérienne de l'arsenic influence donc fortement l'immobilisation de l'élément.

Néanmoins le site présente une faible diversité bactérienne (ce qui est dû au très faible pH, à la toxicité des métaux dissous mais aussi au nombre limité d'accepteurs et de donneurs d'électrons). La population dominante est *Gallionella ferruginea* capable d'oxyder le fer, d'*Actinobacteria*, Fe-oxydante et *Acidithiobacillus ferrooxidans* Fe-oxydante, sulfate-réductrice. Enfin les divers travaux ont conduit à la mise en évidence de la présence de *Desulfobacterium* (bactérie sulfato-réductrice), de *Thiomonas* (As-oxydante) (Duquesne *et al.*, 2003) (Bruneel *et al.*, 2006), et d'archées (Bruneel *et al.*, 2008).

Des eucaryotes ont également été trouvés sur le site, c'est notamment le cas d'*Euglena mutabilis* qui pourrait oxyder l'As III en As V. Cet organisme participe de plus à l'oxygénation du milieu, grâce à la photosynthèse et à l'apport en matière organique (Casiot *et al.*, 2004).

4. Conclusions

Les conditions requises pour générer un AMD sont donc : le contact avec l'oxygène atmosphérique, une aquifère et la présence de bactéries Fe-oxydantes acidophiles ou acidotolérantes. Outre leur rôle dans la génération des AMDs, des bactéries ont aussi un rôle clef dans les différents cycles biogéochimiques des métaux contaminant ces AMDs.

La bonne connaissance de la biogéochimie des AMDs est importante en raison de leurs effets sur l'environnement. Une meilleure connaissance des processus microbiens qui participent à la génération des AMDs permettrait le développement de technologies empêchant la formation de ce type de pollution. De plus une meilleure connaissance de la flore microbienne présente et de son interaction redox avec l'AMD pourrait aider au développement des techniques de bioremédiation permettant de réhabiliter ces environnements. Cependant encore peu de choses sont connues sur les bactéries non-cultivables de ces environnements et sur leur rôle dans ces écosystèmes. Le développement récent des technologies de microbiologie moléculaire, en particulier les avancées en génomique et en protéomique, donneront accès à des micro-organismes non-cultivables et aideront à élucider leur rôle dans ces environnements complexes.

Chapitre III

L'utilisation des outils « Méta-OMICS »
dans l'exploration d'une communauté
microbienne complexe

I INTRODUCTION

La diversité microbienne ainsi que les interactions trophiques des bactéries dans de nombreux environnements, en particulier les milieux extrêmes, restent encore très peu connues. Le manque de méthodes de culture adéquate ou de méthodes d'enrichissement pour isoler ces micro-organismes est limitant pour explorer ces niches écologiques (Bertin *et al.*, 2008). Les approches de métagénomique, de métatranscriptomique et de métaprotéomique permettent ainsi de s'affranchir des étapes de culture, donc du biais des conditions de laboratoire et ainsi d'explorer le potentiel génétique, métabolique et fonctionnel de la communauté. De façon générale, la génomique, la protéomique, la transcriptomique, etc. sont appelées 'OMICS' par certains auteurs (Jerez, 2008) ; dans la suite de ce chapitre les OMICS appliqués aux communautés complexes seront nommés 'MétaOMICS'.

De nombreux développements technologiques existent et évoluent sans cesse dans ce domaine. Dans ce chapitre les principes MétaOMICS seront développés sans entrer dans les détails techniques, mais dans le souci d'une vision globale et intégrée de l'apport des MétaOMICS à la compréhension du fonctionnement des communautés microbiennes d'environnements complexes.

Les pré-requis :

Pour comprendre les relations entre les souches et leur environnement, et plus précisément le rôle de chaque bactérie dans cet environnement et son interaction vis-à-vis du milieu, il est nécessaire d'avoir préalablement caractérisé le site par des analyses physico-chimiques, comme le pH, la température, la minéralogie, les sources potentielles d'énergie et de nutriments (Allen & Banfield, 2005).

De plus, il est indispensable de connaître la composition biologique de cet écosystème. Pour cela diverses techniques d'analyses culture-indépendantes d'une communauté microbienne complexe sont aujourd'hui utilisées (Figure I3.1) : le séquençage d'ADNr 16S d'une banque de clones, (menée récemment par Pope et collaborateurs (Pope & Patel, 2008)), l'analyse FISH (*Fluorescence in situ hybridization*), etc.

La description phylogénétique de la diversité de la communauté microbienne est une première étape importante, mais comprendre le rôle de ces organismes au sein de leur écosystème est un challenge bien plus important (Allen & Banfield, 2005).

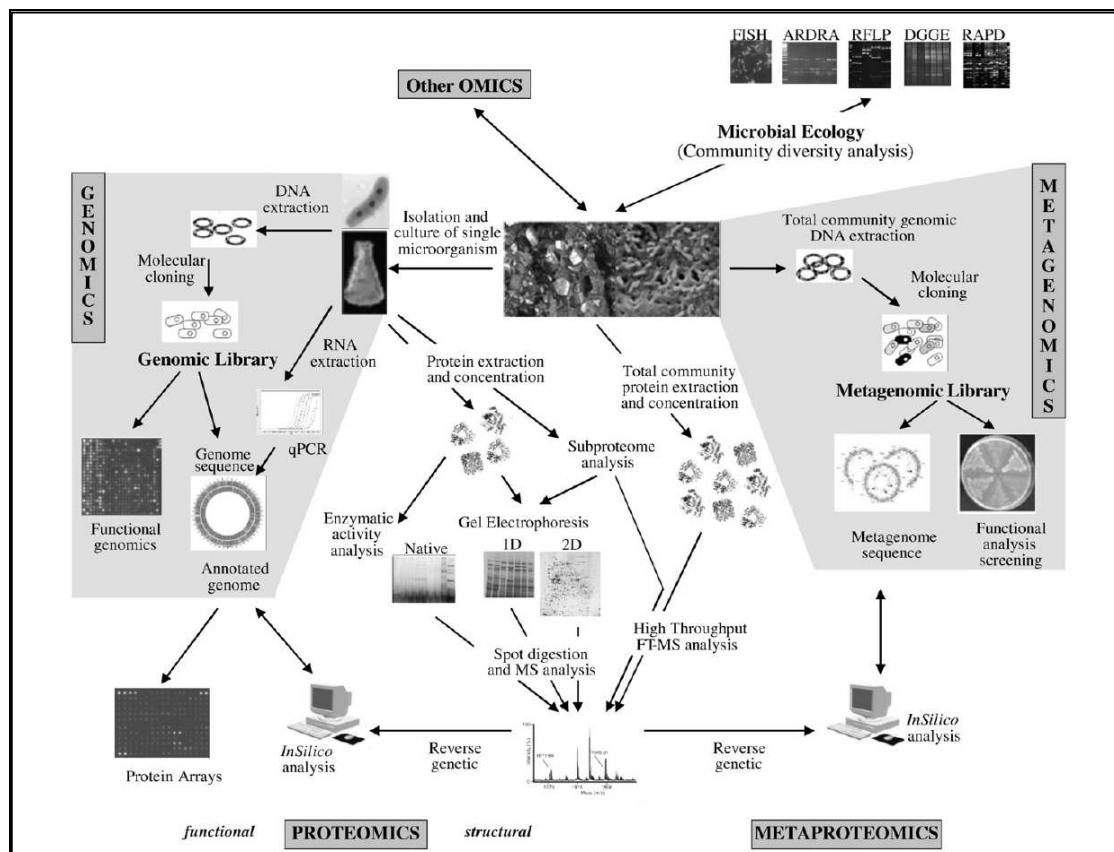


Figure 13.1 : Vue d'ensemble des principaux outils de génomique, de protéomique, de métagénomique et de métaprotéomique utilisables pour l'étude d'une communauté bactérienne d'un AMD (Extrait de (Jerez, 2008)).

Dans la nature, des organismes indépendants sont très rares, les micro-organismes étant le plus souvent organisés en communauté. La compréhension du fonctionnement d'une communauté ne peut pas être réalisée uniquement en se focalisant sur certaines souches isolées qui ne sont pas représentatives des potentialités métaboliques et fonctionnelles de la communauté dans son ensemble. Néanmoins l'étude de souches isolées de ces environnements, en particulier les données de séquences des génomes ou les résultats de protéomique, permet d'apporter des clefs au système et d'enrichir la suite de l'analyse (Allen & Banfield, 2005). Les études métaOMICS sont souvent plus aisées si elles peuvent se référer à des données issues des études OMICS préalables des souches qui composent la communauté.

Applications des techniques OMICS:

Ces techniques peuvent aussi être développées dans un but industriel (Streit & Schmitz, 2004). En effet, les analyses métagénomiques permettent de mettre en évidence des dizaines d'ORFs dont la plupart codent des enzymes encore inconnues et pourraient avoir un intérêt industriel.

Les enzymes sont utilisés dans de très nombreux domaines et représentent un important marché (2,3 billions de dollars en 2003) réparti dans plusieurs secteurs comme par exemple ceux des détergents, de l'agriculture, de l'industrie du papier (Lorenz & Eck, 2005). Il est donc aisé d'imaginer que la découverte de nouveaux enzymes représente un intérêt certain pour les industries.

Pour mettre ce type d'enzymes en évidence, il est possible de cribler une banque d'ADN métagénomique sur une activité métabolique particulière (comme par exemple la dégradation d'ampicilline ou de l'amidon) (Schloss & Handelsman, 2003). Les enzymes correspondants sont ensuite analysés et développés dans l'objectif de leur mise sur le marché (Handelsman, 2004) (Lorenz & Eck, 2005).

Parmi les enzymes produits et commercialisés à partir des données de métagénomique, on peut citer une nitrilase (Diversa), une alcool dehydrogenase (Schering Plough) et une glycosylhydrolase (BRAIN AG).

II METAGENOMIQUE

La métagénomique est une analyse génomique culture indépendante d'une communauté microbienne (Riesenfeld *et al.*, 2004). Elle est aujourd'hui en plein essor puisque plus de 130 projets ont été référencés en janvier 2009 sur GOLD (*Genomes OnLine Database*). L'approche métagénomique consiste à extraire l'ADN d'un échantillon environnemental, de le cloner dans un vecteur, puis de transformer un hôte bactérien, ce qui constitue la banque métagénomique de cet environnement (Figure 13.2). Il est alors possible soit de cribler cette banque sur de différentes fonctions d'intérêt : économiques, environnementales (comme la résistance au nickel (Mirete *et al.*, 2007) ou à l'arsenic (Chauhan *et al.*, 2009)), soit de séquencer complètement la banque dans le but de réaliser une analyse phylogénétique (pour revue (Riesenfeld *et al.*, 2004)) ou dans le but d'une analyse des potentialités génétiques des populations microbiennes d'un écosystème donné (Handelsman, 2004). Le séquençage à grande échelle mené sur les microorganismes de différents océans (Rusch *et al.*, 2007) (Biers *et al.*, 2009) ou sur la mer des Sargasses (Tyson *et al.*, 2004) sont des exemples d'analyses des potentialités de populations microbiennes d'écosystème.

Dans ce chapitre seront détaillés la construction et le séquençage d'une banque métagénomique dans le but d'assembler les séquences en pseudo-génomés et d'analyser leurs potentialités.

1. Principe

Il est possible d'estimer la quantité de séquences nécessaires pour couvrir la communauté bactérienne, préalablement au séquençage. Les analyses de communautés bactériennes présentes sur le site, la composition et l'abondance des différentes souches permettent de définir le nombre de séquences nécessaires (Allen & Banfield, 2005). Par exemple, si l'abondance d'une souche donnée est de 1 %, il faudra une couverture de la banque métagénomique de 100 à 1 000 pour avoir une représentation de cet organisme. Ceci aura pour conséquence de générer une quantité énorme de séquences pour les organismes dominants (Riesenfeld *et al.*, 2004).

Pour tenter de reconstruire les génomes d'une communauté microbienne, il est idéal d'avoir une faible diversité microbienne (Chivian *et al.*, 2008). Ces conditions sont retrouvées dans le cas de symbiose (Woyke *et al.*, 2006) ou dans certains environnements comme les AMDs (Tyson *et al.*, 2004), dans lesquels seuls quelques types d'organismes sont présents.

Après prélèvement de l'échantillon à analyser, il est dans certains cas nécessaire de procéder à une étape préliminaire de séparation des procaryotes. Certains échantillons peuvent en effet contenir de nombreux eucaryotes dont la présence compromet l'analyse. Un gradient de Nycodenz peut par exemple être réalisé et permettre la séparation des procaryotes par gradient de densité (Goulhen-chollet *et al.*, 2009).

L'ADN extrait est ensuite soumis à fragmentation (par nébulisation). Les fragments obtenus (environ 6-8 Kb) sont insérés dans des vecteurs plasmidiques pour être clonés dans une souche d'*E. coli* ; ce qui constitue la banque d'ADN métagénomique. Il est d'autre part possible de réaliser des banques de BAC (*Bacterial artificial chromosome*) avec de grands inserts qui auront pour but de faciliter l'assemblage (Streit & Schmitz, 2004) (Riesenfeld *et al.*, 2004).

Chaque clone est alors séquencé par la technique de séquençage Shotgun³ Sanger (Fleischmann *et al.*, 1995) (MacLean *et al.*, 2009). Cependant les étapes de clonage peuvent être problématiques. En effet les séquences ne sont pas toujours stables et certains gènes peuvent être toxiques pour l'hôte, ce qui biaisera les données de séquence. Afin de s'affranchir du clonage, le séquençage pourra être réalisé par pyroséquençage 454⁴ (MacLean *et al.*, 2009). Cette technique

³ Séquençage de fragments aléatoires.

⁴ Le pyroséquençage est une méthode basée sur l'émission de lumière. L'insertion d'un nucléotide lors des réactions de réplication s'accompagne d'un relargage de pyrophosphate qui sera converti en ATP. Celui-ci sera utilisé par la luciférase pour produire de la lumière (MacLean *et al.*, 2009).

moins onéreuse et plus rapide, génère un très grand nombre de séquences mais qui restent encore trop petites pour permettre un assemblage robuste. Dinsdale et collaborateurs ont réalisé par cette approche une comparaison métagénomique de 15 millions de séquences afin de comparer les potentialités fonctionnelles de différents biomes (Dinsdale *et al.*, 2008).

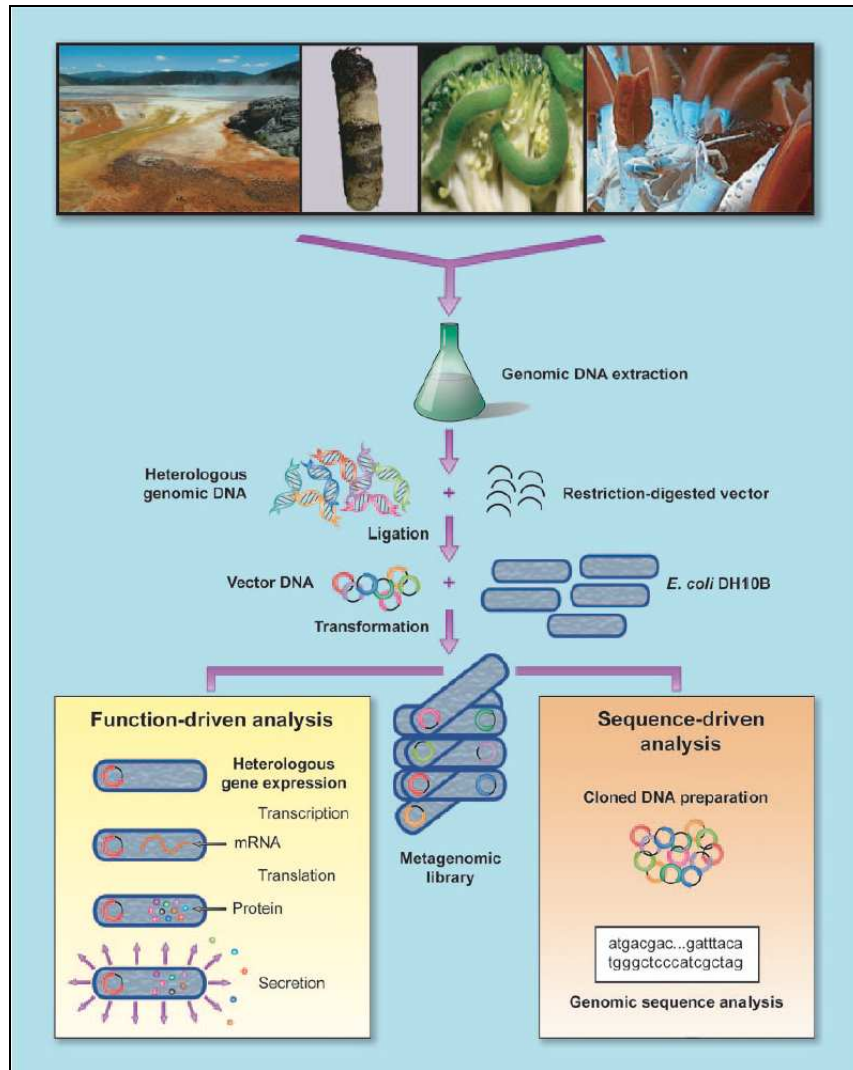


Figure 13.2 : Représentation schématique de la construction et du criblage de la banque métagénomique issue d'un échantillon environnemental. (Extrait de (Handelsman, 2004))

Les séquences sont ensuite assemblées en contigs (ensemble de séquences) grâce à des programmes d'assemblage (par exemple ARACHNE (Batzogluou *et al.*, 2002)) (MacLean *et al.*, 2009). En combinant les résultats d'analyse portant sur l'ADNr 16S, sur le pourcentage en GC, sur le taux de couverture moyen (nombre de lectures par unité de longueur d'ADN) des différents contigs et l'usage des codons, il est possible de reconstruire des super contigs, correspondant à des pseudo-génomes. Il s'agit de génomes plus ou moins complets d'organismes majoritaires, cependant le

regroupement de séquences provenant d'organismes différents, mais très proches ne peut être complètement exclu. La reconstruction de génome est basée pour partie sur le %GC; ainsi l'assemblage est extrêmement difficile quand les souches présentent de nombreux réarrangements génétiques, donc quand les souches ont un contenu génétique différent (îlots génomiques etc.).

2. L'analyse des pseudo-génomés

Une fois assemblés, il est nécessaire d'assigner les pseudo-génomés aux organismes présents par des analyses phylogénétiques des gènes marqueurs moléculaires comme 16S rADN, rpoA.

Les pseudo-génomés reconstruits sont ensuite intégrés dans une plateforme d'annotation et soumis à une annotation automatique. Dans un premier temps la prédiction des phases ouvertes de lecture (ORF : *open reading frame*) est défini, en déterminant les codons initiateurs (start) et terminateurs (stop) bordant des séquences suffisamment longues pour correspondre à un gène et en utilisant l'usage du codon (basé sur certains gènes de ménage). Puis une fonction est attribuée aux gènes par comparaison de séquences dans les différentes bases de données. Les séquences peuvent ensuite être analysées et annotées de façon experte grâce, par exemple, à l'interface d'annotation MaGe développée par le Génoscope (Vallenet et al., 2006) (<https://www.genoscope.cns.fr/agc/mage>) et illustré Figure I3.3.

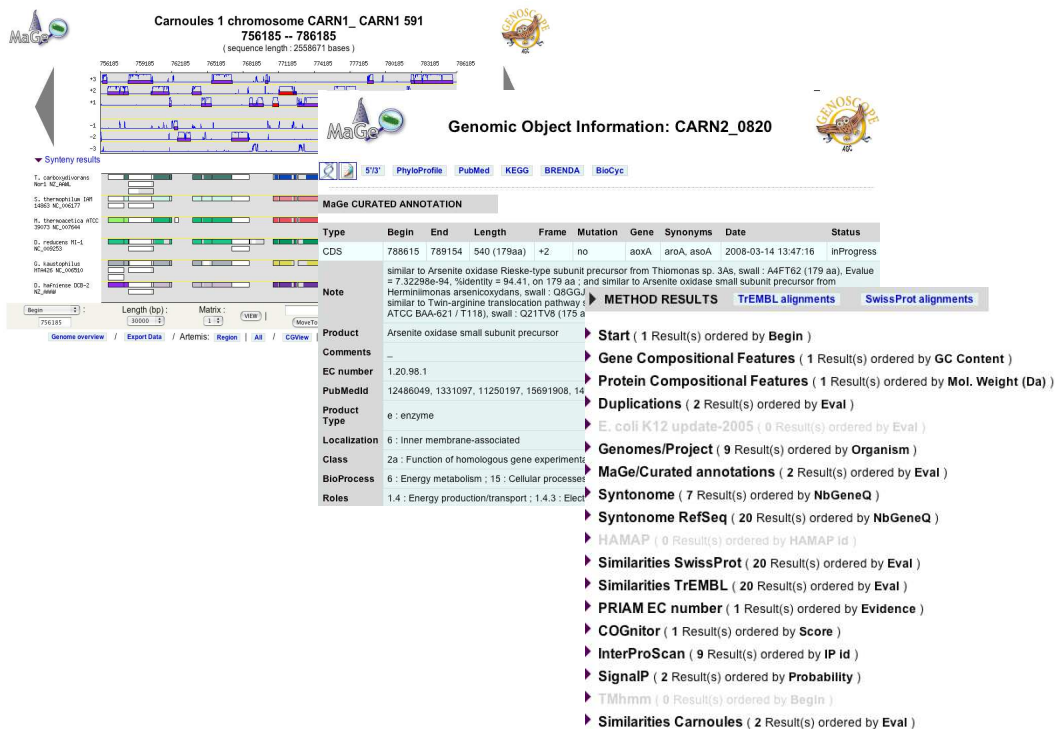


Figure I3.3 : Représentation de la plate-forme d'annotation de séquence MaGe développée par le Génoscope.

Grâce aux nombreux outils d'analyses bioinformatiques (comme InterProScan, TrEMBL, PRIAM EC number), il est possible d'assigner une fonction à chaque gène et ainsi d'explorer les capacités métaboliques de ces pseudo-génomés. Certains outils permettent de plus une analyse *in silico* des différentes voies métaboliques (Caspi *et al.*, 2008) et génèrent des analyses globales du métabolisme des différentes souches (*Microbial Genome Biochemical pathway*) (<https://www.genoscope.cns.fr/agc/mage/wwwpkgdb/MageHome/index.php?webpage=microcyc>).

Ce qui permet de mettre en évidence les interactions trophiques des micro-organismes et le rôle de chacun au sein de leur environnement.

Exemples d'analyses par métagénomique :

Ce type d'approche et ces différentes techniques ont été mises en œuvre dans différentes études (Tringe *et al.*, 2005) (Baker *et al.*, 2006) (Chivian *et al.*, 2008) (Zehr *et al.*, 2008) (Williamson *et al.*, 2008). Tyson et collaborateurs ont utilisé une technique de séquençage shotgun sur l'ADN issu des biofilm d'un AMD et ils ont ainsi pu reconstruire presque entièrement les génomes de *Leptospirillum* du groupe II et de *Ferroplasma* de type II et partiellement trois autres génomes (Tyson *et al.*, 2004). Le très bas pH du biofilm (0,5) combiné à la faible quantité de source d'énergie dans ce milieu a eu pour conséquence une faible diversité microbienne. Ceci couplé au faible taux de réarrangements génomiques, fait de cette communauté un échantillon idéal pour mettre en œuvre cette technique de métagénomique.

La reconstruction n'est cependant pas très aisée, même avec des outils informatiques puissants et une masse de séquences importante. Par exemple le séquençage massif par shotgun des fragments d'ADN de la mer des Sargasses a généré environ 1,36 Gb de séquence, l'estimation montre que ces valeurs correspondent à plus de 1 000 génomes, et à plus de 1,2 million de gènes. Malgré cette quantité de données, seuls quelques pseudo-génomés ont été reconstruits dont un organisme dominant, *Prochlorococcus* (Venter *et al.*, 2004).

Les analyses génomiques préalables sur des souches isolées du milieu d'intérêt permettent souvent d'aider à la reconstruction des pseudo-génomés. Ainsi en utilisant la séquence génomique de *Ferroplasma acidarmanus fer1*, Allen et Banfield ont pu reconstruire le génome de la souche environnementale à plus de 90 % (Allen & Banfield, 2005).

Les approches métagénomiques appliquées à l'étude des communautés microbiennes constituent une réelle avancée pour comprendre la complexité de ce type d'environnement et des interactions bactériennes au sein de leur niche écologique (Tyson *et al.*, 2004) (Streit & Schmitz, 2004).

III METATRANSCRIPTOMIQUE

La métagénomique permet l'analyse des potentialités génétiques d'une communauté bactérienne ; l'analyse métatranscriptomique quant à elle permet l'étude de l'expression des gènes de cette communauté. L'analyse des ARNs bactériens extraits directement d'un échantillon peut se faire d'une part par des reverse transcriptions (RT-PCR) couplées à des PCR : les ARNm sont convertis en ADNc et amplifiés grâce à des amorces spécifiques des gènes d'intérêt. L'étude des ARNs bactériens peut d'autre part être réalisée par des approches d'analyse par puces ou par séquençage complet des ADNc.

La métatranscriptomique par séquençage

L'extraction des ARNm doit être effectuée rapidement après l'échantillonnage afin de limiter sa dégradation. En général les ARNr sont éliminés du pool d'ARN afin d'augmenter la proportion des ARNm, les transcrits sont alors amplifiés et convertis en ADNc grâce à la reverse transcriptase. Puis ils sont séquencés par pyroséquençage (Figure 13.4) (Moran, 2009). Les données de séquences correspondent donc aux gènes exprimés par la communauté microbienne et les données de couverture correspondent au taux d'expression de ces gènes. Les séquences sont ensuite analysées de la même façon que les séquences issues de l'approche métagénomique. Cette approche a été réalisé sur différents environnements, comme des communautés microbiennes marines (Gilbert *et al.*, 2008) (Poretsky *et al.*, 2009).

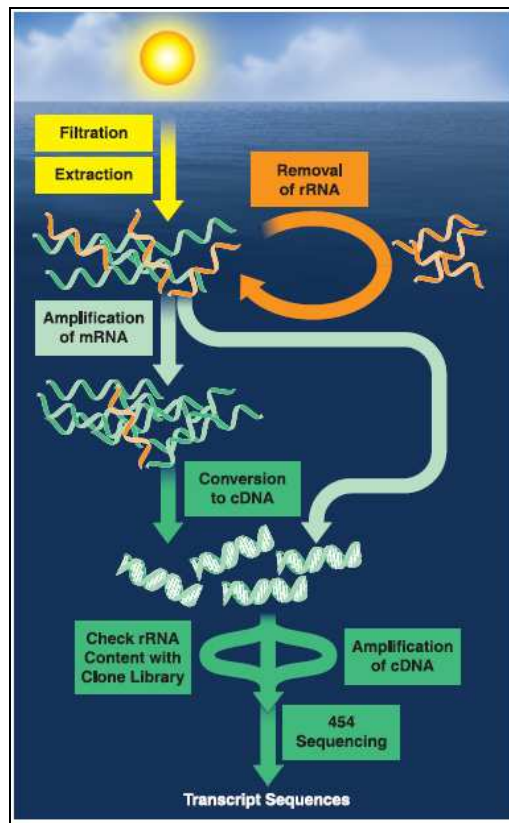


Figure 13.4 : Représentation schématique de la métatranscriptomique (Extrait de (Moran, 2009)).

Les limites de la métatranscriptomique résident dans la difficulté à extraire une quantité suffisante d'ARNm procaryotes d'une part et d'autre part dans l'absence de relation directe entre le nombre de transcrits et l'activité des protéines correspondantes. Pour compléter l'analyse il est donc préférable de réaliser une étude métagénomique.

IV METAPROTEOMIQUE

Coupler l'analyse d'un génome avec l'analyse protéomique permet de passer de l'étude des potentialités d'un génome à l'étude de ses fonctionnalités.

La technologie la plus utilisée est en général la migration de l'extrait protéique sur un gel d'électrophorèse polyacrylamide à deux dimensions (2D PAGE) combinée à la spectrométrie de masse. Cette technologie est de la même façon utilisée pour réaliser la caractérisation à grande échelle des protéines exprimées *in situ* par une communauté microbienne : on parle alors de la métagénomique.

Principe :

Les protéines sont extraites directement de l'échantillon ou après séparation des eucaryotes ou des acides humiques (Benndorf *et al.*, 2007). Après dénaturation, les protéines sont séparées sur un gel polyacrylamide d'électrophorèse à une dimension (en fonction de la masse de la protéine) ou en deux dimensions (une première dimension en fonction d'un gradient de pH qui sépare les protéines en fonction de leur point isoélectrique puis dans une deuxième dimension en fonction de leur masse). Elles sont ensuite révélées par coloration au bleu de Coomassie ou au nitrate d'argent par exemple. Tous les spots ou toutes les bandes seront alors séquencés par spectrométrie de masse (Wilmes & Bond, 2009) (Figure 13.5). Enfin les protéines peuvent être séparés par une chromatographies liquide (LC) ou sur des puces par HPLV-Chip/MS (Hardouin *et al.*, 2006).

La spectrométrie de masse est une technologie en constante évolution, son principe est basé sur le temps de vol (corrélé à la masse) des peptides issus de la digestion des protéines (par la trypsine en général). Ceci permet de donner un profil de poids moléculaire de l'ensemble des peptides de chaque protéine. Ce profil sera comparé *in silico* à celui issu des séquences métagénomique grâce à des logiciels comme Mascot, Profound (Wilmes & Bond, 2009).

L'analyse des protéines basée sur l'utilisation des gels d'électrophorèse 2D couplée à la spectrométrie de masse présente des limites : l'analyse ne permet pas de détecter les protéines de faible abondance, ni les protéines avec des points isoélectriques ou des masses extrêmes.

De nouvelles technologies sont aujourd'hui disponibles, en particulier l'analyse métagénomique réalisée sans gel d'électrophorèse. La première étape consiste en une digestion enzymatique des protéines extraites, les différents peptides sont ensuite séparés sur une colonne de chromatographie liquide à haute résolution (HPLC) puis analysés par spectrométrie de masse (ESI-MS/MS : *electron spray ionization MS*) (Valenzuela *et al.*, 2006).

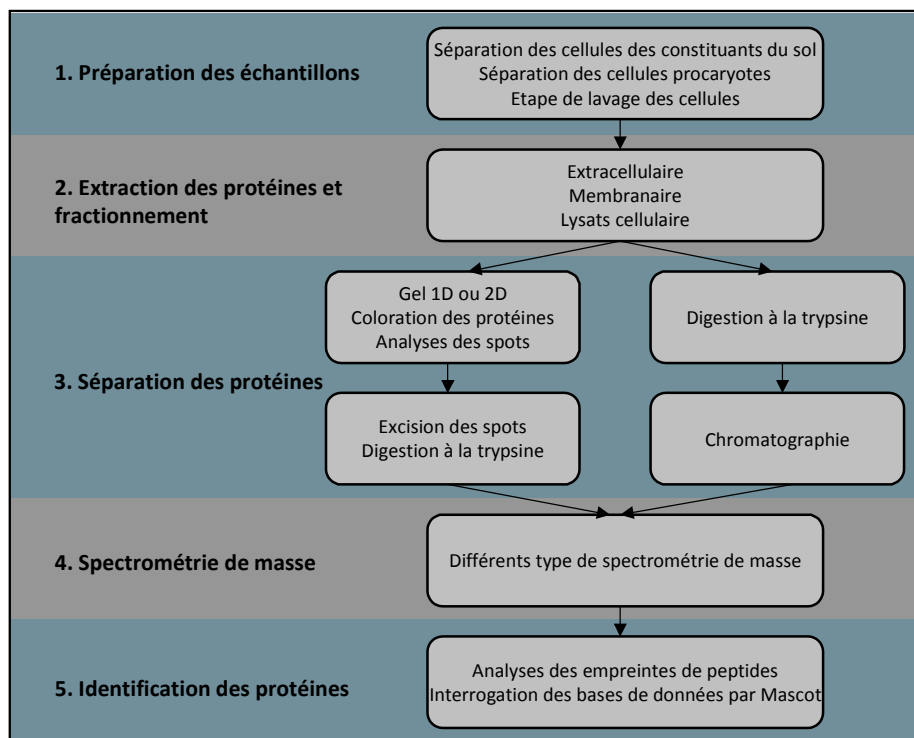


Figure 13.5 : Schéma synthétique de l'analyse métagénomique (A partir de (Wilmes & Bond, 2009)).

Les données de séquence de l'analyse métagénomique permettent de générer une base de données nécessaire à l'identification de protéines de l'analyse métagénomique du même environnement. La puissance de l'étude métagénomique est donc corrélée à la complétude de cette base de données.

Récemment l'analyse métagénomique correspondant à l'analyse métagénomique réalisée par Tyson et collaborateurs a été réalisée sur le biofilm d'AMDs (Tyson *et al.*, 2004). Les auteurs ont détecté 2033 protéines des souches les plus abondantes, en particulier de *Leptospirillum* (Ram *et al.*, 2005) (Banfield *et al.*, 2005b). Ils ont alors mis en évidence qu'une des protéines les plus abondantes était un cytochrome impliqué dans l'oxydation du fer et donc de la formation de l'AMD.

De nombreuses avancées techniques sont fréquemment disponibles en métagénomique et permettent d'affiner les analyses et de repousser sans cesse les limites de détection des protéines. D'autre part les bases de données sont de plus en plus vastes de par le nombre toujours plus grand d'analyses MétaOMICS. Ainsi l'analyse métagénomique est en constant développement et son utilisation est grandissante.

V CONCLUSIONS

De nombreuses autres approches dérivées de ces techniques Méta-OMICS ont vu le jour (Park *et al.*, 2008). Toutefois, les analyses Méta-OMICS ne permettent pas d'attribuer des fonctions aux gènes ou aux protéines de fonction inconnue. Ceux-ci sont pourtant en nombre très important dans les différentes études. Par exemple 42 % des gènes prédits dans l'étude menée par Tyson et collaborateurs portent une fonction inconnue (Tyson *et al.*, 2004) (Ram *et al.*, 2005).

Il est donc nécessaire de les coupler aux approches classiques d'étude des souches ou au criblage de fonction. L'isolement de souches sur des milieux de cultures variés reste essentiel et nécessaire à l'étude de la communauté microbienne et de ses interactions au sein d'un écosystème. Il est en effet nécessaire de réaliser des analyses physiologiques afin de comprendre le rôle précis de chaque organisme et de chaque gène.

Les approches Méta-OMICS sont un premier pas dans l'analyse des souches non-cultivées et des communautés complexes. De plus les résultats de ce type d'analyse apportent une connaissance plus précise des souches jusqu'alors non cultivées et pourraient ainsi permettre de générer de nouvelles approches ou de nouveaux milieux de culture pour les cultiver.

Chapitre IV

Objectifs des travaux

Les mécanismes d'adaptation et de résistance à l'arsenic développés par les procaryotes sont variés. Parmi eux, les plus étudiés sont la réduction de l'arséniate par le système Ars (très répandu dans l'environnement), la méthylation de l'arsenic, la production d'exopolysaccharides et l'oxydation de l'arsénite par les gènes *aox*. L'objectif de ce travail a donc été l'étude de la diversité des réponses au stress et les différents métabolismes de l'arsenic chez les procaryotes.

Le système d'oxydation codé par les gènes *aox* n'avait au début de mes travaux été mis en évidence que chez quelques souches dont *Herminiimonas arsenicoxydans*, *Rhizobium* NT26 ou *Alcaligenes faecalis*. Ainsi peu de séquences étaient disponibles dans les banques et peu de souches arsénite-oxydantes étaient caractérisées. Différentes questions se sont donc posées : Quelle diversité existe entre les différents gènes *aox* ? Contrairement aux gènes *ars*, les gènes *aox* peuvent-ils être considérés comme marqueurs moléculaires de souches arsénites-oxydantes ? Les souches correspondantes sont-elles répandues dans l'environnement et quels sont leurs mécanismes d'adaptation ?

D'autre part, le début de ma thèse coïncide avec les premières études de génomique environnementales (métagénomique et métaprotéomique). Celles-ci sont particulièrement adaptées aux environnements présentant une faible diversité microbienne, c'est notamment le cas des AMDs. L'arsenic y étant fortement présent, une telle approche pourrait-elle permettre d'élucider les interactions trophiques des différentes bactéries au sein d'un AMD arsénié ? De plus dans ces environnements, les arsénite-oxydases sont-elles nombreuses ?

Ce projet, concernant la diversité des réponses au stress et des différents métabolismes de l'arsenic chez les isolats et sur les communautés bactériennes, va permettre de mieux comprendre le rôle de ces bactéries dans ces environnements, mais aussi d'appréhender les mécanismes moléculaires mettant en jeu leur survie et leur développement. Le travail s'est principalement focalisé sur la diversité des gènes *aox* d'isolat en condition de laboratoire et de communautés complexes *in situ*.

La première partie de ce travail a donc porté sur l'étude des différentes capacités d'adaptation bactérienne à un milieu arsénié, et en particulier sur l'étude de l'arsénite-oxydase de souches cultivables isolées de divers milieux contaminés. Il s'agit de *Pseudomonas* S11 (Chapitre I-Partie I), *Leptothrix* S1.1 (Chapitre I-Partie II) et *Thiomonas* 3As (Chapitre I-Partie III).

L'étude s'est ensuite élargie à la diversité des gènes *aox* de bactéries appartenant à une collection de souches d'environnements arséniés et à l'analyse de l'utilisation du gène *aox* comme marqueur moléculaire (Chapitre II).

L'utilisation des données et des approches issues de l'analyse en condition de laboratoire a permis dans un second temps l'analyse *in situ* des capacités d'adaptation à l'arsenic, en particulier l'étude de la diversité des gènes *aox* de communautés bactériennes.

D'une part, l'étude de la diversité des gènes *aoxB*, couplée à une analyse phylogénétique et fonctionnelle d'un sédiment d'une ancienne mine d'argent à Sainte-Marie-aux-Mines, en Alsace a été mise en œuvre. Cette étude a permis de mettre en évidence la grande dispersion des gènes *aox* dans un environnement peu contaminé (Chapitre III).

D'autre part, une étude environnementale descriptive (métagénomique) et fonctionnelle (métaprotéomique) a été menée sur les sédiments miniers arséniés d'un AMD à Carnoulès dans le Gard. Celle-ci a permis d'élucider les interactions trophiques des bactéries présentes dans cet environnement extrême et le rôle clef d'une souche arsénite-oxydante (Chapitre IV).

L'ensemble des travaux que j'ai réalisés au cours de ces trois années de thèse font l'objet de publications scientifiques parues (Chapitre II), soumises (Chapitre I - Partie III) ou prêtes pour soumission (Chapitre I - Partie I et II, Chapitre III et Chapitre IV). J'ai donc choisi de présenter ce manuscrit sous forme d'articles scientifiques.

Personnels Travaux

Chapitre I

Caractérisation de bactéries
arsénite-oxydantes

première partie

Etude d’une souche arsénite-oxydante *Pseudomonas* S11

I INTRODUCTION

Le site minier de Salsigne proche de Carcassonne était, au vingtième siècle, la plus importante mine d’or en Europe, mais il fut aussi exploité pour sa richesse en arsenic (950 000 t d’As III produites). De cette exploitation, il reste aujourd’hui 15 millions de tonnes de déchets contenant jusqu’à 100 g/kg d’arsenic.

Une souche *Pseudomonas* sp. S11 a été isolée de ce site pour ses capacités d’oxydation mais aussi pour son niveau de résistance élevé à l’arsenic. Dans cette étude, cette souche a fait l’objet d’une caractérisation physiologique générale d’une part et d’une analyse des caractères d’adaptation à l’arsenic d’autre part.

En premier lieu sa capacité de résistance élevée à l’arsenic a été explorée grâce à une banque de complémentation fonctionnelle. Ceci a permis de mettre en évidence quatre gènes *arsB* fonctionnels et liés à des gènes *tnpA*. Ces gènes *ars* ont donc probablement été acquis par transfert horizontal. Puis l’exploration des gènes *aox* codant l’arsénite-oxydase a révélé un opéron *aox* classique lui aussi lié à un gène *tnpA*. Des tests biochimiques, en particulier l’analyse de l’activité enzymatique sur gel protéique natif, ont permis de mettre en évidence une activité arsénite-oxydase constitutive. Enfin une induction de la production d’exopolysaccharides en présence d’arsenic a été mise en évidence. Cette souche présente donc divers mécanismes de résistance et d’adaptation à l’arsenic.

Les informations obtenues lors de cette étude pourraient permettre de comprendre comment les bactéries peuvent faire face à un environnement fortement arsénié. Elles pourraient aussi fournir des outils pour appréhender le rôle des bactéries du sol adaptées à un milieu arsénié dans la mobilisation de ce métalloïde vers le compartiment aquatique.

Ma collaboration dans ce projet, a porté sur la réalisation des tests physiologiques et génétiques, la participation aux tests biochimiques, et la réalisation du manuscrit et des figures.

L'ensemble de ces résultats fait l'objet d'une publication en cours de préparation et est présentée ci-dessous.

II MANUSCRIT

Hypertolerance to arsenic and constitutive arsenite-oxidase activity in the *Pseudomonas* sp. strain S11 isolated from the gold and arsenic mine waste from Salsigne (France)

Audrey Heinrich-Salmeron, Florence Goulhen-Chollet, Philippe Hammann, Didier Lièvreumont,
Philippe N. Bertin, and Marie-Claire Lett.

En préparation.

Hypertolerance to arsenic and constitutive arsenite-oxidase activity in the *Pseudomonas* sp. strain S11 isolated from the gold and arsenic mine waste from Salsigne (France)

Audrey Heinrich-Salmeron¹, Florence Goulhen-Chollet¹, Philippe Hammann², Didier Lièvre¹, Philippe N. Bertin¹, and Marie-Claire Lett¹.

¹Génétique Moléculaire, Génomique et Microbiologie, UMR7156 CNRS & UdS, Strasbourg, France,

²Plateforme Protéomique de l'Esplanade, CNRS, Strasbourg, France

Correspondence: lett@unistra.fr

ABSTRACT

We investigated a new *Pseudomonas* sp. strain S11, which has been isolated from the highly arsenic-rich solid waste from the gold and arsenic mine of Salsigne (France). This strain is hypertolerant for arsenic (MIC: 33mM) and produces EPS when grown in the presence of arsenic. It can oxidize arsenite (As III) to arsenate (As V). The screening for arsenite resistance of a S11 gene bank, transformed in an As-sensitive *E.coli*, led to the identification of four functional *arsB* genes. Sequencing analysis indicated that all of them are located immediately upstream of a transposase *tnpA* gene. Two of them, *arsB1* and *arsB2*, showed 99% identity with the transposon located TnAtcArs *arsB* gene from *Acidithiobacillus caldus*. A DNA fragment of 4.4 kb, harboring the *aoxAB* operon, encoding the two subunits of the arsenite-oxidase, was sequenced and revealed the presence of a putative *tnpA* gene and a putative *upsX* gene, respectively upstream of *aoxA* and downstream of *aoxB*. Arsenite-oxidase activity was detected on native polyacrylamide gels in cells grown with or without arsenite, indicating for the first time that the *aox* operon can be expressed constitutively. This was confirmed by the presence of the AoxB protein in bacteria treated or not with arsenic, as detected by Western-blot analysis and peptides sequencing.

INTRODUCTION

Arsenic is widely spread in the environment due to natural sources or is generated by anthropogenic activities, including agriculture and mining. The average concentration of arsenic in earth crust is about 2 mg kg⁻¹ (0,0002 wt%), but in mining waste it can come up to very high values. The mining site of Salsigne, located near Carcassonne, in south of France, was the biggest gold mine in Europe during the 20th century. It was also exploited for arsenic. The cumulative production of arsenic trioxide amounted to nearly 950,000 tons. Ore processing stopped by the end of 2004. The generated waste represents a total amount of 15 million tons and still contains high concentrations of arsenic, but also other pollutants such as bismuth, sulfur, copper, iron, etc. Perez & Valiente, (2005), by studying 21 surfaces samples, selected in a random manner, in the polluted area and on the surrounding ditches, evaluated the median concentration of As as 6g kg⁻¹, with a minimum of 0.7g kg⁻¹ and a maximum of 55 g kg⁻¹. Effluents, resulting from rainwater percolating through this arsenic-rich ores generate important contaminations of the nearby Orbiel River.

The World Health Organisation (WHO) classified arsenic among the most powerful carcinogens (http://www.who.int/water_sanitation_health/diseases/arsenicosis/en/). Water contaminations by this metalloid have been considered by EU, WHO and EPA as a major threat to

human health. Human exposure to arsenic can result in various diseases going from intestinal disturbance to different cancers as reviewed recently by Duker *et al.*, 2005 (Duker *et al.*, 2005).

In natural aquatic environment, arsenic occurs mainly as oxyanions arsenate (H_2AsO_4^- , HAsO_4^{2-}) [As V] or arsenite (H_3AsO_3^0) [As III], this last one being the most toxic and most mobile. Bacteria play a key role in arsenic biogeochemical cycle (Oremland & Stolz, 2003). The major bacterial transformations consist in arsenate reduction (*ars* genes) and dissimilation (*arr* genes), arsenite oxidation (*aox* or *aso* genes) and methylation (*arsM*) as recently reviewed (Páez-Espino *et al.*, 2009; Rosen & Liu, 2009). These mechanisms control the speciation of arsenic in the environment and thus play a crucial role in the arsenic mobility and bioavailability.

Arsenic is lethal for most organisms, yet many bacteria are known to survive by adapting to arsenic exposure. Microorganisms have evolved a noteworthy variety of mechanisms to accommodate the presence of high concentration of arsenic. This was illustrated by the exploration of the whole genome of the β -Proteobacterium *H. arsenicoxydans*, the first fully characterized arsenic-metabolizing microorganism. The study of this bacterium, isolated from arsenic-contaminated industrial wastewater, demonstrated that in addition to the resistance mechanisms to arsenite and arsenate, multiple strategies are used by these bacteria to efficiently adapt to and colonize toxic environments including synthesis of exopolysaccharides (EPS) and arsenic scavenging (Muller *et al.*, 2007). It should be noted that, most knowledge concerning the mechanisms of bacterial tolerance to arsenic has been obtained from water environment (for review, see Lievremont *et al.*, 2009), while the characteristics of bacteria adapted to polluted soils are rather scarce (Achour *et al.*, 2007), (Cai *et al.*, 2009a), (Cai *et al.*, 2009b), (Drewniak *et al.*, 2008).

The main purpose of this study was to investigate the mechanisms involved in the hypertolerance of a bacterium adapted to a soil containing very high concentrations of arsenic. Among a collection of As-resistant bacteria isolated from solid waste samples of the mining area of Salsigne, we chose an arsenite-oxidizing bacterium, *Pseudomonas* sp S11, which showed the highest MIC for arsenite. The gained informations should allow us to 1) understand how bacteria can cope with a highly toxic terrestrial environment and 2) provide some tools to decipher the environmental significance of bacterial strains adapted to soils containing high concentrations of arsenic and evaluate their role in the mobilization of this metalloid toward the aquatic compartment.

METHODS

Bacterial strain, plasmids and growths conditions.

Strains used in this study are presented in Table 1. Plasmid pcDNA 2.1 (Invitrogen) was used for genomic DNA library construction. Bacterial strains were cultivated aerobically at 25°C in chemically defined medium (CDM) described by Muller et al, 2003 (Muller *et al.*, 2003) or in Luria-Bertani (LB) medium. Strain S11 was isolated on CDM agar supplemented with 100 ppm As III by streaking the culture obtained after the incubation of one gram of arsenic rich solid waste from the mining site of Salsigne for 5 day in CDM broth (25 mL) at 25°C. M63 (Miller, 1972) medium, which was used when testing the genomic DNA library, was prepared as follows: 200 ml of solution A (13.6 g K₂HPO₄, 2 g (NH₄)₂SO₄, 0.50 mg FeSO₄, 7H₂O), 2 ml of solution B (49.28 mg MgSO₄, 7H₂O) and 20 ml of solution C (glucose 20%) were mixed and make up to 1 litre with sterilized water. All solutions were prepared with purified water (Milli-Q system, Millipore). The pH of solution A were adjusted to 7 with KOH. Solutions A, B, C and water were individually sterilized by autoclaving (120°C for 20 min).

Determination of S11 strain phenotypic characters. Phenotypic data obtained by classical methods of carbohydrate utilization determined on various carbon substrates were tested by Biolog EcoPlates (Garland & Mills, 1991), or with the API 20NE systems. The minimal inhibitory concentration (MIC) to arsenic was determined on liquid media, 5µl bacterial suspension (Late-log-phase cells) was transferred in triplicate on liquid CDM (10 ml) non-amended (controls) or amended with increasing concentration of arsenic. CMI was determined as As concentration inhibiting the occurrence of turbidity in liquid medium after 3 days at 25°C. Cell mobility was observed on a low concentration agar plate (0.3% agar) after 24 h or 48 h growth as described in Muller *et al.*, 2007 (Muller *et al.*, 2007). Antibiotic resistance was tested by the disc diffusion method on Mueller Hinton Agar (Difco Laboratories) with susceptibility disks (Sanofi Diagnostics –Pasteur) (Courvalin *et al.*, 1985).

Transmission electron microscopy. Bacteria were negatively stained (2% (w/v) aqueous uranyl acetate) and observed by transmission electron microscopy (TEM Hitachi H-600 operated at 75 kV). A drop of 24 h grown bacterial culture was deposited on nickel grids coated with a polyvinyl formal (Formvar®) film. After 10 s, the grid was dried using a piece of filter paper, and finally uranyl acetate solution (2%, w/v) was added to stain and the grid was air-dried at ambient temperature.

Detection of the arsenite-oxidation activity. Strain S11 was transferred to CDM agar plates supplemented with arsenite (100 mg liter⁻¹, 1.33 mM), and grown for 24 h at 25°C. Bacteria were subsequently screened for oxidation of arsenite to arsenate by using the AgNO₃ method (Lett *et al.*, 2001), which is based on a differential reaction of AgNO₃ with arsenite or arsenate. The cells on agar plates were flooded with a solution of 0.1 M AgNO₃. A brownish precipitate of silverorthoarsenate

(Ag_3AsO_4) revealed the presence of arsenate in the medium (colonies expressing arsenite oxidase), while the presence of arsenite (bacteria lacking the arsenite oxidase activity) was detected by a bright yellow precipitate of silverorthoarsenite (Ag_3AsO_3)

Arsenic quantification and speciation in liquid cultures. Bacteria were grown in CDM broth supplemented with arsenite ($100 \text{ mg liter}^{-1}$ to $2200 \text{ mg liter}^{-1}$, 1.33 mM to 29.3 mM) 24 h at 25°C and agitated. Arsenic transformation was determined in culture supernatants obtained by filtration of a stationary phase bacterial suspension through a sterile $0.22\text{-}\mu\text{m}$ -pore-size Durapore filter (Millipore). Arsenic species were separated by high-performance liquid chromatography (HPLC) on a reversed-phase polymeric resin (PRP-X100; inner diameter, 250 by 4.1 mm; particle size, $10 \mu\text{m}$; Hamilton) and quantified by inductively coupled plasma-atomic emission spectrometry (ICP-AES) with a sequential Jobin Yvon JY 138 Ultrace spectrometer as described by Weeger *et al.*, 1999 (Weeger *et al.*, 1999).

Quantitative exopolysaccharides production assay. Exopolysaccharides (EPS) production was assessed in 96-well polystyrene microplates as follows: $150 \mu\text{L}$ growth medium containing 0, 50, 100 and 200 ppm (0, 0.665, 1.33 and 2.66 mM) of As III were inoculated with $5 \mu\text{L}$ of an exponential culture. Samples were incubated 72 hours at 25°C without shaking. EPS were visualized after removing of the growth media by staining with 0.05% crystal violet for 15 min. Following destaining with $150 \mu\text{L}$ 96% ethanol, the biofilm was quantified by absorbance at 595 nm compared to uninoculated controls. For each experiment, three separate cultures of each strain were analyzed in four replicates.

DNA manipulation. DNA extraction was performed using the Genome Wizard Promega Kit. Plasmids preparation was performed with Midiprep Qiagen®. PCR amplification of 16S rRNA genes, putative *ars* genes and putative *aox* genes, were performed with primers described in Table 1. The PCR reactions were carried out with a “Mastercycler ep gradient Eppendorf®”, with an initial denaturation step at 95°C for 1 min, followed by 35 cycles of denaturation (40 s at 95°C), annealing (40 s), and elongation (1.10 min or 2 min for ribosomal RNA 16S genes amplifications) steps at 72°C . Annealing temperatures was defined to the GC content of primers: 52°C for *aox* and *ars* genes and 54°C for ribosomal RNA 16S genes amplification. The products were separated by agarose gel electrophoresis and purified by PCR purification Kit Qiagen®. The S11 strain genomic library was constructed in plasmid pcDNA 2.1 as previously described (Hommais *et al.*, 2001). About 80,000 clones were selected on LB plates supplemented with $100 \mu\text{g ml}^{-1}$ ampicillin and pooled. Large-scale plasmid DNA isolation was carried out using the JETstar Kit 2.0 (genomed, Qbiogen®). This library was used to transform *E. coli* arsenic-susceptible strain AW3110. Clones of interest were selected on M63 medium supplemented with 50 ppm As III (0.665 mM).

Detection and identification of arsenite oxidase activity.

Preparation of soluble protein extracts of strain S11 and H. arsenicoxydans. Strain S11 and *H. arsenicoxydans* were grown to late exponential phase on CDM supplemented or not with As III (100 ppm, 1.33 mM). Cells were gently harvested by centrifugation (10,000 g, 10 min) and washed twice and homogenized in physiological serum (NaCl, 0.9%). Then they were disrupted by sonication at 100A (ten-times 30 s at maximum speed with 1 min interval on ice, 80% duty cycle). Unbroken cells were discarded by centrifugation (10,000 g, 10 min). Finally, protein concentration was determined according to the method of Lowry *et al.*, 1951 (Lowry *et al.*, 1951) using BSA as a standard.

Detection of arsenite oxidase activity on native gels. Soluble proteins were separated on a native gel using 12% polyacrylamide containing Triton X-100 0.1% (w/v). Electrophoresis was carried out at 50 mA per gel. After a 15 min equilibration in MES (50 mM pH 6) at 25°C, the gel was stained in the same buffer supplemented with 80 µl phenazine monosulfate (PMS) 100 mM and 20 µl of 2,4-dichloro phenol indophenol (DCPIP) 300 mM for 30 min in the dark. The presence of an arsenite oxidase activity resulted in the decoloration of the DCPIP after addition of 400 ppm (5.32 mM) As III. Spot of interest were cut out and stored at -20°C.

SDS-PAGE analysis of proteins. The protein extraction was performed by boiling the soluble extract (described above) for 5 min. The protein extracts and the proteins spots cut out from native gel were analysed on 12% slab gels by SDS-PAGE which was performed according to the method of Laemmli, 1970 (Laemmli, 1970). Electrophoresis was carried out at 50 mA per gel. Proteins were stained with Coomassie Blue reagent R250. Selected spots were cut out from gels and stored at -20°C until mass spectrometry analysis.

In gel digestion. The gel digestion procedure was carried out as described by Rabilloud *et al.*, 2001 (Rabilloud *et al.*, 2001). Preparation of the gel pieces before trypsin digestion was performed by a liquid handler robot (QuadZ215, Gilson International, France). Briefly, gel bands were washed alternately with 100 µl of 25 mM NH₄HCO₃ and 100 µl of acetonitrile (ACN) (3 min wash under shaking and the liquid was discarded before addition of the next solvent). This hydrating/dehydrating cycle was repeated twice and the pieces of gel were dried for 20 min before reduction (10 mM DTT/25 mM NH₄HCO₃ buffer at 56°C for 45 min) and alkylation (25 mM iodoacetamide/25 mM NH₄HCO₃ buffer for 45 min, room temperature). Afterwards, gel spots were again washed with 3 cycles of 25 mM NH₄HCO₃/ACN alternately. Following 20 min drying step, the gel pieces were rehydrated by three volumes of trypsin (Promega, V5111), 12.5 ng µl⁻¹ in 25 mM NH₄HCO₃ buffer (freshly diluted) and incubated overnight at room temperature. Tryptic peptides were extracted from gel by vigorous shaking for 30 min in adapted volume of 35% H₂O/60% ACN/5% HCOOH.

MALDI mass spectrometry MALDI mass measurement was carried out on an Autoflex III Smartbeam (Bruker-Daltonik GmbH, Bremen, Germany) matrix-assisted laser desorption/ionization

time-of-flight mass spectrometer (MALDI-TOF TOF) used in reflector positive mode. A prespotted anchorchip target (PAC system from Bruker Daltonik, technical note TN011) with HCCA matrix was used to analyze the tryptic digests. The resulting peptide mass fingerprinting data (PMF) and peptide fragment fingerprinting data (PFF) were combined by Biotoools 3 software (Bruker Daltonik) and transferred to the search engine MASCOT (Matrix Science, London, UK). Variable modifications (methionine oxidation and cysteine carbamidomethylation) and one tryptic missed cleavage were taken into account and the peptide mass error was limited to 50 ppm. Proteins were identified by searching data against NCBI non-redundant protein sequence database.

Western immunoblotting: Western immunoblotting was performed by using rabbit antisera raised against the purified AoxB (1:285 dilution) from *H. arsenicoxidans* (Koechler *et al.*, 2009). Following a 1-h incubation with this first antibody, the nitrocellulose membrane was exposed to goat anti-rabbit antibody coupled to alkaline phosphatase (1:800 dilution; Bio-Rad Laboratories) for 1 h and was then developed with the alkaline phosphatase color development reagent.

RESULTS

Isolation and phenotypic characteristics of strain S11.

Strain S11 was isolated on CDM agar, in aerobic conditions, from the As-contaminated soil of Salsigne. It represents the first strain characterized on this mining site, which contains up to 10%wt of arsenic. Microscopic observations showed that cells are Gram negative rods with rounded ends, approximately 1.4-2.6 μm long and 0.6-0.3 μm wide and harbor a single polar flagellum. A TEM observation is presented in Fig 1A. Colonies were ivory, notched and dry when cells were grown on CDM agar. When grown in liquid CDM medium, bacteria formed important aggregates, which were difficult to dislocate.

Phylogenetic analyses based on 16S rRNA gene sequence (1,418 nucleotides) further supported the conclusion that the strain belongs to the *Gammaproteobacteria* and that its nearest phylogenetic relatives are members of the family *Pseudomonadaceae* in the genus *Pseudomonas stutzeri* (99% sequence similarity on 100% coverage) with Ribosomal Database Project.

The biochemical characterization of the strain S11 was determined by carbon sources utilization and antibiotic resistance assays. Inoculating BIOLOG plates and using the API 20NE system S11 capabilities were tested. S11 is able to grow by oxidizing pyruvic acid methyl ester, D-Xylose, D-galacturonic acid, L-asparagine, Tween 40, Tween 80, D-mannitol, glycogen, itaconic acid and S11 was found able to reduce nitrate in N_2 , to assimilate D-glucose as well as D-mannitol. Physiological characteristics are presented in Table 2. Growth of the isolate on CDM agar was not inhibited by

chloramphenicol but by ampicillin, kanamycin, tetracycline, streptomycin, trimethoprim-sulfamethoxazol and rifampicin.

Resistance to arsenic and metals.

The MIC of strain S11 for arsenite (As[III]) was 33 mM on CDM liquid. Strain S11 was also found to be resistant to numerous heavy metals and metalloids such as As[V], Se[IV], Cu[II], Pb[II] (data not shown).

The resistance genes were investigated by using a gene bank constructed in the arsenic-susceptible *E. coli* strain AW3110, with total DNA of the strain S11. Transformants were selected for their ability to grow on arsenite 50 ppm (0.665 mM). Seventeen arsenite-resistant *E. coli* clones were identified and inserts of their plasmid were sequenced and mapped. Alignment and phylogenetic analysis of the sequences in the NCBI database revealed the presence in the strain S11 of four *arsB* genes, encoding a functional arsenite efflux pump. Based on nucleotide BLAST analysis, two of them, called *arsB1* and *arsB2*, found respectively in cluster 1 and cluster 2, showed 99% identity with TnAtcArs *arsB* gene from *Acidithiobacillus caldus* (De Groot *et al.*, 2003, Tuffin *et al.*, 2005) . Moreover, like in TnAtcArs, we found upstream, flanking *arsB1* and *arsB2* genes, two ORFs, one being similar to ORF8-1 coding a putative cystathione- β -synthase, a protein containing a [CBS] domain- and the other one to ORF7-1, a putative NADH-like oxido-reductase, none of them having a proven function in arsenite resistance (Tuffin *et al.*, 2005). In addition we detected downstream to *arsB1* and *arsB2*, a *tnpA1* gene coding a putative transposase. However, an unknown sequence, separated the ORFs *arsB2* and *tnpA2*. Disruption occurs on nucleotide 650 in *arsB2*, *tnpA2* begun at position 66, as compared with TnAtcArs. In addition two additional *arsB* genes, called *arsB3* and *arsB4*, were found in the strain showing respectively 82% and 85% identity with the TnAtcArs *arsB* gene. These two last ones were also adjacent to a transposase *tnpA* gene.

Identification of the arsenite oxidation characteristics of the strain S11

The strain *Pseudomonas* sp S11 was grown during 24 h in CDM supplemented with 100 ppm (1.33 mM) As III and the arsenite and arsenate concentrations were determined on the filtrated culture supernatants by using HPLC-ICP-AES analyses. Results confirmed complete oxidation after 24 h incubation. Moreover, incubation of the strain with growing concentrations of arsenite showed that the oxidation capacity was conserved up to 2,200 ppm (29.33 mM) As III. However, due to the important flocculation of the cultures in relation with the aggregation of the strain in the presence of arsenic, it was impossible to establish a growth curve in parallel and calculate the enzymatic activity.

To investigate the oxidation capacity a fragment of 1,435 pbs of the large subunit of arsenite oxidase was amplified from the genomic DNA of S11 with degenerate oligonucleotide primers (Quéméneur *et al.*, 2008) and sequenced. In addition, a fragment of 4,430 bps was sequenced around the *aoxB* gene with primers designed in the *aoxB* gene. The map of the *aox* operon is presented in Figure 2. The *aoxAB* genes are composed of 2,473 bps and 521 bps respectively and present 82% identity and 100% coverage with *aoxAB* from *Pseudomonas* sp. TS44 (Cai *et al.*, 2009b) (Genbank accession number EU311944) and 64% identity and 84% coverage with *aoxB* and 67% identity and 35% coverage with *aoxA* from *H. arsenicoxydans* (Muller *et al.*, 2003). Upstream of *aoxA*, we found a fragment of a *tnpA* gene coding a putative transposase showing 84% identity with the *tnpA* of Tn4652 from *Pseudomonas putida* pwwO. Downstream, we detected a fragment of a gene coding UxpB protein showing 72% identity with *P. stutzeri* A1501. UxpB has an unknown function, but is probably involved in phosphatase activity.

Examination of AoxA protein sequences alignment using clustalW showed about 90% identity on 179 amino acids with AoxA Rieske subunit from *H. arsenicoxydans*. A cristallographic study realized in *Alcaligenes faecalis* (Ellis *et al.*, 2001) showed that the [2Fe-2S] cluster of this subunit is coordinated by amino acids Cys60 and His62 from the β 4- β 5 loop and Cys78 and His81 from the β 6- β 7 loop. These two residues pairs coordinate the two Fe atoms. Figure 2 points out that all the active residues are highly conserved in strain S11 as well as the residues around and motifs, such as beta strand, hydrogen bound or helical region.

The amino acids alignment of AoxB sequences displayed about 70% identity on 829 amino acids with AoxB from *H. arsenicoxydans*. Based on *A. faecalis* AoxB sequence, (Ellis *et al.*, 2001) the large subunit of AoxB is divided into four domains : Domain I contain [3Fe-4S] cluster coordinated by the motif Cys21-X₂-Cys24-X₃-Cys28 highly conserved in all sequences studies. Residues His195, Gly203, Arg419 and His423 are residues where represent the substrate binding site and residues Ser238, Cys24, Asn704, Ser99 and Ser98 are probably involved in electron transfer All catalytic sites, binding sites or electron transfer sites of the both amino acids sequences harbor very high conservation sequence and present as well neighbors sequences highly conserved. However in strain S11 a loss of 2 regions of 6 and 8 resides (553 to 559 and 779 to 786) was observed.

The expression of the arsenite-oxidase of *Pseudomonas* sp strain S11 is constitutive

The arsenite oxidase capacity of the strain S11, was tested by investigating the enzymatic activity on native gel, using the redox mediator DCPIP as electron acceptor. Arsenite-oxidase activity was observed in bacterial cells grown up to stationary phase in the presence or in the absence of arsenate, suggesting that in S11 arsenite oxidase is constitutively produced and active (Figure 3).

Western-blotting analysis using rabbit antisera raised against the purified AoxB from *H. arsenicoxydans* were performed on strain S11 total proteins separated on a denaturing gel electrophoresis. A protein band showing the same size band and intensity was revealed in both cultures grown with and without arsenite (data not shown). Taken together, these results indicate the constitutive production of AoxB.

To confirm these results, bands showing arsenite-oxidase activity were cut out and run on a denaturing gel electrophoresis. The resulting bands were stained with Coomassie blue, cut out and analyzed by MS/MS mass spectrometry. This allows us to identify AoxB from *Pseudomonas* sp. at approximately 70 kDa (Table 3). Moreover, five peptides were detected to determine AoxB proteins, peptides A9-2 and B9-3 (i.e. liftgispmvr and SPEQNALNVDDFTTR, respectively) were found in the sequence of S11.

Remarkably, peptides from AoxB were detected in both native bands corresponding to cells grown with or without As III, confirming the constitutive production of AoxB in S11 as already detected by DCPIP assay. Finally, the proteomic analysis allowed the identification of a peptide corresponding to a c-cytochrome (peptide A2, Table 3).

Investigation of biofilm synthesis.

Colonization capacities linked to biofilm formation were also tested. Transmission electron microscopy (TEM) observations showed a clear exopolysaccharides (EPS) layer around the cells when they were cultivated in the presence of As III versus in absence (Figure 1). Biofilm assays were performed in order to quantify the EPS production. The biofilm production was clearly enhanced when bacteria were challenged with As III 100 ppm (1.33 mM) (about +20% +/-2%).

DISCUSSION

This work reports the characterization of a bacterial strain showing high adaptation capacities to an environment heavily contaminated with arsenic. It has been isolated from the former gold and arsenic mine of Salsigne, which is characterized by an outstanding high concentration of arsenic. To our knowledge, nothing is known so far about the microbial community of this very toxic environment.

The strain S11 has been identified as a *Pseudomonas* sp. Bacteria belonging to these species are widely distributed in the environment, occupying diverse ecological niches (Lalucat et al., 2006). They are largely found in contaminated environment. While many of these bacteria were shown to

be resistant to arsenite, only a few of exhibit also arsenite oxidation activity (Achour *et al.*, 2008) (Cai *et al.*, 2009b).

However when compared to other *Pseudomonas*, the strain S11 is one of the most resistant to arsenic ever described, since it exhibits a MIC of 33 mM As III. A similar strain, isolated from a mining environment and described by Chang *et al.*, 2008 (Chang *et al.*, 2008) arbors a resistance to 26 mM of arsenic, while arsenite-oxidizers, isolated from an industrial soil, presented MICs ranging between 13 and 23 mM (Achour *et al.*, 2008).

This high level of resistance may be in relation with the presence in strain S11 of four copies of the *arsB* gene, all of them coding an active arsenic-efflux pump, as revealed by the complementation of the *E. coli* arsenic-susceptible strain AW3110. These genes may have been acquired by transposition. This hypothesis is supported by the fact that the four copies of *arsB* are adjacent to a *tnpA* gene, coding a transposase. Moreover, two *arsB* sequences, called *arsB1* and *arsB2*, showed 99% identity with the *arsB* gene associated to the transposon TnAtcArs from *A. caldus* (De Groot *et al.*, 2003). It should be noted that the gene order between *tnpA*, *arsB*, ORF8-1 and ORF-7, initially described in TnAtcArs was conserved in the case of *arsB1* and *arsB2*. TnAtcArs is belonging to the Tn3 family and has been shown to be able to transposition (Kotze, 2006, Tuffin *et al.*, 2006).

In addition to the high resistance to arsenic, the strain S11 was found able to oxidize arsenite and conserves this activity up to 2,200 ppm (29.33 mM). By comparison, in *H. arsenicoxydans* the Aox activity was conserved to a concentration of 6 mM.

The Aox activity was detected, by using the DCIP test on a non-denaturing gel electrophoresis, on bacteria grown in the presence or in the absence of arsenic, suggesting that the arsenite-oxidase is constitutively produce. The production of an active arsenite-oxidase was confirmed both by Western-blotting, as well as by the isolation and the identification of AoxB signature peptides. By comparison, when performed with *H. arsenicoxydans* ULP, arsenite-oxidase activity was never observed with bacteria grown in the absence of arsenite. This is the case for most arsenite-oxidizers. In a very recent work, Koechler *et al.*, (submitted) demonstrated that, in *H. arsenicoxydans*, the expression of the *aox* operon is controlled by the binding of the sigma factor RpoN to the consensus sequence (TGGCACG(N4)TTG) which is located immediately upstream of *aoxAB*. This sequence was found also in 6 other arsenite-oxidizers. On the contrary, in the case of the strain S11 this consensus sequence is totally absent, and this may be the reason while there is no arsenic-dependant control of the expression of the enzyme.

The association of the *aoxAB* operon with a transposase gene suggests that it may have been acquired by horizontal transfer. The hypothesis of the mobility of the *aoxAB* operon will have to be

tested in the future, as well as that of the 4 *arsB* genes, described previously, which are also associated with *tnpA* genes.

The detection of a cytochrome-C₅₅₁ peptide, associated to the native Aox protein, represent a good argument for the involvement of this protein as a metabolic partner in arsenite oxidation by the strain S11. If many studies showed that the *aoxB* gene is contiguous to *aoxD*, coding a putative c-type cytochrome (Muller *et al.*, 2003) (Duquesne *et al.*, 2008) only Santini *et al.*, (2007), described the role of cytochrome c₋₅₅₂, coded by a *cytC* gene belonging to the arsenite oxidase cluster, in arsenite oxidation in the facultative chemolithoautotroph NT-26 strain (Santini *et al.*, 2007).

Finally, the induction of cellular aggregates in presence of arsenic as well as the synthesis of EPS, may constitute an efficient mechanisms to isolate the bacterial cell from the very toxic environment. The quantification and speciation of arsenic forms performed by HPLC-ICP-AES showed a loss of arsenic total (around 5%, with two repetitions) suggesting an arsenic sequestration. However, no biofilm scavenging of arsenic could be observed by TEM analysis, as it has been found in a previous study in *H. arsenicoxidans* (Muller *et al.*, 2007).

ACKNOWLEDGEMENTS

We acknowledge M. Erhardt for the transmission electron microscopy analysis performed at the microscopy platform of the Institut de Biologie Moléculaire des plantes (IBMP-CNRS, Strasbourg). We acknowledge Sandrine Koechler and David Halter for technical help. This work was support by the French Ministry of Research and the Université de Strasbourg. This work was performed in the frame of the Groupement de Recherche "Métabolisme de l'Arsenic chez les micro-organismes: de la résistance à la détoxification" (GDR2909-CNRS).

REFERENCES

- Achour, A. R., Bauda, P. & Billard, P. (2007).** Diversity of arsenite transporter genes from arsenic-resistant soil bacteria. *Res Microbiol* **158**, 128-137.
- Bayard, R., Chatain, V., Gachet, C., Troadec, A. & Gourdon, R. (2006).** Mobilisation of arsenic from a mining soil in batch slurry experiments under bio-oxidative conditions. *Water Res* **40**, 1240-1248.
- Cai, L., Liu, G., Rensing, C. & Wang, G. (2009a).** Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. *BMC microbiology* **9**, 4.
- Cai, L., Rensing, C., Li, X. & Wang, G. (2009b).** Novel gene clusters involved in arsenite oxidation and resistance in two arsenite oxidizers: *Achromobacter* sp. SY8 and *Pseudomonas* sp. TS44. *Appl Microbiol Biotechnol*, 1-11.
- Carlin, A., Shi, W., Dey, S. & Rosen, B. P. (1995).** The *ars* operon of *Escherichia coli* confers arsenical and antimonial resistance. *Journal of Bacteriology* **177**, 981-986.
- Casiot, C., Leblanc, M., Bruneel, O., Personné, J. C., Koffi, K. & Elbaz-Poulichet, F. (2003).** Geochemical processes controlling the formation of As-rich waters within a tailings impoundment (Carnoulès, France). *Aquatic Geochemistry* **9**, 273-290.
- Chang, J. S., Yoon, I. H. & Kim, K. W. (2007).** Isolation and *ars* detoxification of arsenite-oxidizing bacteria from abandoned arsenic-contaminated mines. *Journal of Microbiology and Biotechnology* **17**, 812-821.
- Chang, J. S., Kim, Y. H. & Kim, K. W. (2008).** The *ars* genotype characterization of arsenic-resistant bacteria from arsenic-contaminated gold-silver mines in the Republic of Korea. *Appl Microbiol Biotechnol* **80**, 155-165.
- Chatain, V., Bayard, R., Sanchez, F., Moszkowicz, P. & Gourdon, R. (2005).** Effect of indigenous bacterial activity on arsenic mobilization under anaerobic conditions. *Environ Int* **31**, 221-226.
- Courvalin, P., Goldstein, F., Philippon, A. & Sirot, J. (1985).** *L'antibiogramme*.
- De Groot, P., Deane, S. M. & Rawlings, D. E. (2003).** A transposon-located arsenic resistance mechanism from a strain of *Acidithiobacillus caldus* isolated from commercial, arsenopyrite biooxidation tanks. *Hydrometallurgy* **71**, 115-123.
- Dondon, M. G., De Vathaire, F., Quénel, P. & Fréry, N. (2005).** Cancer mortality during the 1968-1994 period in a mining area in France. *European Journal of Cancer Prevention* **14**, 297-301.

- 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**, 1069-1074.
- Duker, A. A., Carranza, E. J. M. & Hale, M. (2005).** Arsenic geochemistry and health. *Environ Int* **31**, 631-641.
- Duquesne, K., Lieutaud, A., Ratouchniak, J., Muller, D., Lett, M. C. & Bonnefoy, V. (2008).** Arsenite oxidation by chemoautotrophic moderately acidophilic *Thiomonas* sp.: From the strain isolation to the gene study. *Environmental Microbiology* **10**, 228-237.
- Ellis, P. J., Conrads, T., Hille, R. & Huhn, P. (2001).** Crystal structure of the 100 kDa arsenite oxidase from *Alcaligenes faecalis* in two crystal forms at 1.64 Å and 2.03 Å. *Structure* **9**, 125-132.
- Garland, J. L. & Mills, A. L. (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.
- Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J. P., Danchin, A. & Bertin, P. (2001).** Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Molecular Microbiology* **40**, 20-36.
- Koechler, S., Cleiss, J., Proux, C. & other authors (2009).** Transcriptional and post-transcriptional controls affects arsenite oxidase activity in *Herminiimonas arsenicoxydans*. *Molecular Microbiology* submitted.
- Kotze, A. A., Tuffin, I. M., Deane, S. M. & Rawlings, D. E. (2006).** Cloning and characterization of the chromosomal arsenic resistance genes from *Acidithiobacillus caldus* and enhanced arsenic resistance on conjugal transfer of *ars* genes located on transposon TnAtcArs. *Microbiology* **152**, 3551-3560.
- Laemmli, U. K. (1970).** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lalucat, J., Bennasar, A., Bosch, R., García-Valdès, E. & Palleroni, N. J. (2006).** Biology of *Pseudomonas stutzeri*. *Microbiology and Molecular Biology Reviews* **70**, 510-547.
- Lett, M. C., Paknikar, K. M. & Lievremont, D. (2001).** A simple and rapid method for arsenite and arsenate speciation. *Biohydrometallurgy: Fundamentals, Technology and Sustainable Development*, 541-546.
- Lievremont, D., Bertin, P. N. & Lett, M. C. (2009).** Arsenic in contaminated waters: Biogeochemical cycle, microbial metabolism and biotreatment processes. *Biochimie*.

- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951).** Protein measurement with the Folin phenol reagent. *J Biol Chem* **193**, 265-275.
- Miller, J. H. (1972).** Experiments in Molecular Genetics. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*.
- Morin, G., Juillot, F., Casiot, C., Bruneel, O., Personné, J. C., Elbaz-Poulichet, F., Leblanc, M., Ildelfonse, P. & Calas, G. (2003).** Bacterial formation of tooeleite and Mixed Arsenic(III) or Arsenic(V) - Iron(III) gels in the Carnoulès acid mine drainage, France. A XANES, XRD, and SEM study. *Environmental Science and Technology* **37**, 1705-1712.
- Muller, D., Lièvreumont, D., Simeonova, D. D., Hubert, J. C. & Lett, M. C. (2003).** Arsenite oxidase *aox* genes from a metal-resistant β -proteobacterium. *Journal of Bacteriology* **185**, 135-141.
- Muller, D., Médigue, C., Koechler, S. & other authors (2007).** A tale of two oxidation states: Bacterial colonization of arsenic-rich environments. *PLoS Genetics* **3**, 0518-0530.
- Ordóñez, E., Letek, M., Valbuena, N., Gil, J. A. & Mateos, L. M. (2005).** Analysis of genes involved in arsenic resistance in *Corynebacterium glutamicum* ATCC 13032. *Appl Environ Microbiol* **71**, 6206-6215.
- Oremland, R. S. & Stolz, J. F. (2003).** The ecology of arsenic. *Science* **300**, 939-944.
- Páez-Espino, D., Tamames, J., De Lorenzo, V. & C-novas, D. (2009).** Microbial responses to environmental arsenic. *BioMetals* **22**, 117-130.
- Pérez, G. & Valiente, M. (2005).** Determination of pollution trends in an abandoned mining site by application of a multivariate statistical analysis to heavy metals fractionation using SM&T-SES. *J Environ Monit* **7**, 29-36.
- Prasad, K. S., Subramanian, V. & Paul, J. (2009).** Purification and characterization of arsenite oxidase from *Arthrobacter* sp. *BioMetals*, 1-11.
- Quéméneur, M., Heinrich-Salmeron, A., Muller, D., Lièvreumont, D., Jauzein, M., Bertin, P. N., Garrido, F. & Joulain, C. (2008).** Diversity surveys and evolutionary relationships of *aoxB* genes in aerobic arsenite-oxidizing bacteria. *Appl Environ Microbiol* **74**, 4567-4573.
- Rabilloud, T., Strub, J. M., Luche, S., Van Dorsselaer, A. & Lunardi, J. (2001).** A comparison between *Sypro Ruby* and *ruthenium ii tris* (bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels. *Proteomics* **1**, 699-704.
- Rosen, B. P. & Liu, Z. (2009).** Transport pathways for arsenic and selenium: A minireview. *Environ Int* **35**, 512-515.

Santini, J. M., Kappler, U., Ward, S. A., Honeychurch, M. J., vanden Hoven, R. N. & Bernhardt, P. V. (2007). The NT-26 cytochrome c552 and its role in arsenite oxidation. *Biochim Biophys Acta* **1767**, 189-196.

Simonato, L., Moulin, J. J., Javelaud, B., Ferro, G., Wild, P., Winkelmann, R. & Saracci, R. (1994). A retrospective mortality study of workers exposed to arsenic in a gold mine and refinery in France. *American Journal of Industrial Medicine* **25**, 625-633.

Tuffin, I. M., de Groot, P., Deane, S. M. & Rawlings, D. E. (2005). An unusual Tn21-like transposon containing an *ars* operon is present in highly arsenic-resistant strains of the biomining bacterium *Acidithiobacillus caldus*. *Microbiology* **151**, 3027-3039.

Tuffin, I. M., Hector, S. B., Deane, S. M. & Rawlings, D. E. (2006). Resistance determinants of a highly arsenic-resistant strain of *Leptospirillum ferriphilum* isolated from a commercial biooxidation tank. *Appl Environ Microbiol* **72**, 2247-2253.

Vanden Hoven, R. N. & Santini, J. M. (2004). Arsenite oxidation by the heterotroph *Hydrogenophaga* sp. str. NT-14: The arsenite oxidase and its physiological electron acceptor. *Biochimica et Biophysica Acta - Bioenergetics* **1656**, 148-155.

Weeger, W., Lièvreumont, D., Perret, M., Lagarde, F., Hubert, J. C., Leroy, M. & Lett, M. C. (1999). Oxidation of arsenite to arsenate by a bacterium isolated from an aquatic environment. *BioMetals* **12**, 141-149.

Yan, Y., Yang, J., Dou, Y. & other authors (2008). Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 7564-7569.

Zawadzka, A. M., Crawford, R. L. & Paszczyński, A. J. (2007). Pyridine-2,6-bis(thiocarboxylic acid) produced by *Pseudomonas stutzeri* KC reduces chromium(VI) and precipitates mercury, cadmium, lead and arsenic. *BioMetals* **20**, 145-158.

LEGENDS

Figure 1: Transmission electron microscopic observations from *Pseudomonas* S11 grown without As III (A) or with 100 ppm (1.33mM) As III (B). Exopolysaccharids production can be observed on panel B.

Figure 2: Panel A: Organization of the *aox* Gene cluster in *Pseudomonas* S11.

PanelB: Protein sequence alignment and putative secondary structure of AoxA. Sequence alignment was carried out by CLUSTAL. Residues known to play a role in As(III) oxidase activity are framed in black. The beta-sheet and alpha helix are shown according to the three-dimensional structure of *A.faecalis*.

Figure 3: Arsenic activity from *Pseudomonas* S11 protein extract in native gel.

Native electrophoresis protein profiles from S11 grown on 100 ppm As III (1.33 mM) (lane 2 +As) or not (lane 1 -As) and resolved using 12% SDS-PAGE. The arrow shows the position of the decoloration of the blue stained gel with indicating an oxidation .

Table 1: Strains and primer used in the study

Names	Media	Sequences 5'-3'	Reference
Strains			
<i>Pseudomonas stutzeri</i> S11	CDM or LB		This study
<i>Herminiimonas arsenicoxydans</i> ULP	CDM		(Muller <i>et al.</i> , 2003)
M1 (<i>H. arsenicoxydans</i> disrupt in <i>aoxB</i> with a mini-Tn5::lacZ2)	CDM		(Muller <i>et al.</i> , 2003)
<i>Cupriavidus metallidurans</i> CH34	CDM		
<i>Escherichia coli</i> electromax DH5- α	LB or M63		Invitrogen
<i>E. coli</i> AW3110 (<i>E. coli</i> W3110 Δ arsRBC)	LB or M63		(Carlin <i>et al.</i> , 1995)
Primers			
16S			
rd1		5'- AAGCTTAAGGAGGTGATCCAGCC-3'	Universal primer
fd1		5'- AGAGTTTGATCCTGGCTCAG-3'	Universal primer
pcDNA			
rev-primer		5'-GGAAACAGCTATGACCATG-3'	Universal primer
M13-20-primer		5'- GTAAAACGACGGCCAGT-3'	Universal primer
Aox			
BM1-2F		5'-CACTTCTGCATCGTGGG(AGCT)TG(CT)GG(AGCT)TA-3'	(Quéméneur <i>et al.</i> , 2008)
BM1-2R		5'- TGCCCCAGATGATGCC(CT)TT(CT)TC(AG)(AT)A-3'	(Quéméneur <i>et al.</i> , 2008)
Aox11-3F		5'-GCCGGGGTAAGGACAATGTTTTG-3'	This study
Aox11-3R		5'-CCCCATGGTCAAAAAGATTG-3'	This study
Aox11-1		5'-CGTCATGGTGATCTGCATG-3'	This study
Aox11-2		5'-GCCGTCGTTATCCTTGATG-3'	This study
Aox11-3		5'-GATCTCTATCGCACGAAGTTC-3'	This study
Aox11-A		5'-GCTCCCCATGATCAAAAAG-3'	This study
Aox11-B		5'-CGAGGGGGTACTAGTTTTTC-3'	This study
Aox11-C		5'-TGTTAGCGGGCCACTTATATG-3'	This study
Aox11-D		5'-ACAATCGCCCGGTGAATATAC-3'	This study

Table 2: General physiological and metabolic properties of strain S11

Characteristics	S11
Cell morphology	Rods
Cell size (µm)	1.4-2.6 X 0.6-0.3
Colony aspect	ivory, notched and dry
Assimilation of:	
Beta methyl-D-glucoside	-
Galactonic acid gamma-lactone	-
Arginine	-
Pyruvic acid methyl ester	+
D-Xylose	+
D-galacturonic acid	+
L-asparagine	+
Tween 40	+
i-erythriol	-
4-hydroxy benzoic acid	-
L-serine	-
Tween 80	+
Alpha-cyclodextrin	-
N-acetyl-D-lucosamine	-
Gamma-hydroxybutyric acid	-
L-threonine	-
D-mannitol	+
Glycogen	+
D-glucosaminic acid	-
Itaconic acid	+
Glycyl-L-glutamic acid	-
D-cellobiose	-
Glucose-1-phosphate	-
Alpha-ketobutyric acid	-
Phenylethylamine	-
Alpha-D-lactose	-
D-alpha-glycerol phosphate	-
D-malic acid	-
Putescine	-
Nitrate reduction in N ₂	+
D-glucose assimilation	+
D-mannitol assimilation	+
Oxidase	+

Table 3: MS/MS identifications of 3 proteins bands showing arsenite oxidase activity.

Name ¹	Peptids	Identifications ²	peptids number	Mw/pI	Access N ^o NCBI	Mascot Score	Error (ppm)	coverage ³ %
A2	NSPTVLNAVFNAAQFWDGR	cytochrome c551 peroxidase precursor [<i>Pseudomonas</i> stutzeri A1501]	1	35905/5.59	gij146281333	75	7	5
A9	TNHIWQSAYHDQHLSFR	arsenite oxidase Mo-pterin subunit [<i>Pseudomonas</i> sp. TS44]	1	92741/5.78	170026536	78	12	10
	LIFTGISTPMVR	arsenite oxidase large subunit [<i>Pseudomonas</i> sp. 1]	1	41706/5.27	162568509	68	11	19
B9	EIDSQGGPTGHLATYER	arsenite oxidase Mo-pterin subunit [<i>Pseudomonas</i> sp. TS44]	2	92741/5.78	170026536	132	17	12
	TNHIWQSAYHDQHLSFR							
	SPEQNALNVDFTTR	arsenite oxidase large subunit [<i>Pseudomonas</i> sp. 1]	1	41706/5.27	162568509	118	11	23

1: Name: spot number

2: Identification and annotated function if applicable

3: The sequence coverage represents the % of the combined peptide masses covering the indicated sequence.

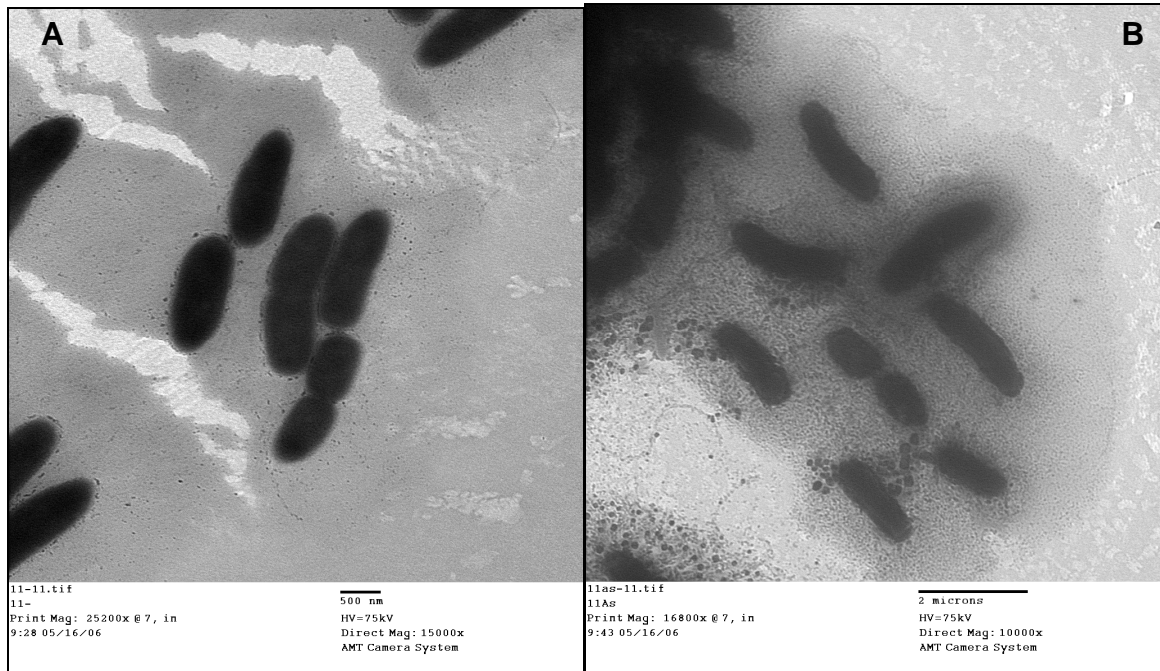


Figure 1

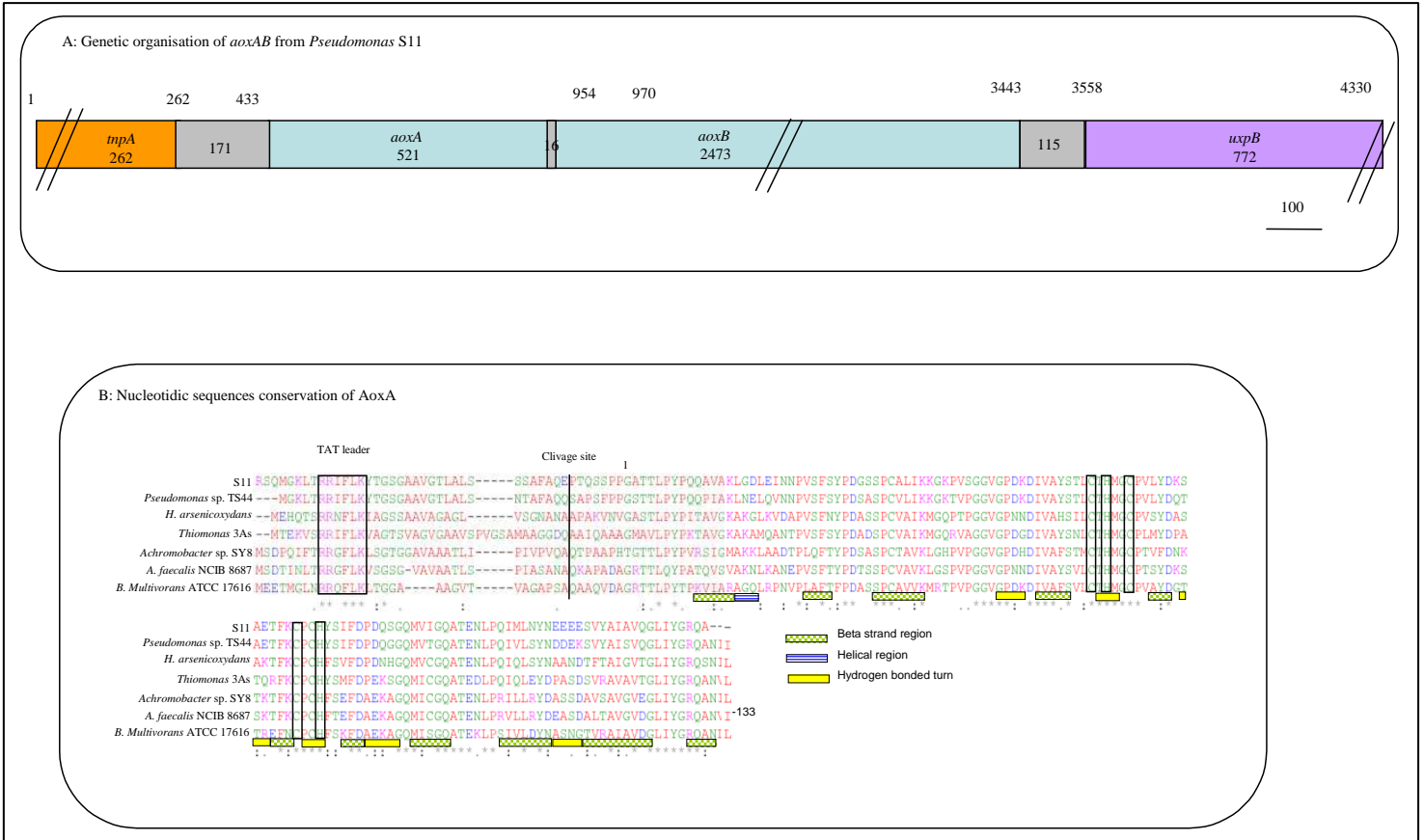


Figure 2



Figure 3

deuxième partie

Etude d'une souche arsénite-oxydante Leptothrix arsenicoxydans

I INTRODUCTION

Le site de Loperec dans le Finistère est une ancienne exploitation minière d'or et d'arsenic. Ainsi, il présente de fortes teneurs en métaux dans les eaux d'exhaure (sortie de la galerie), particulièrement en arsenic et en fer, teneurs dues à l'altération de la roche mère. Le fer précipite à l'exhaure sous forme d'hydroxydes, tandis que les concentrations en arsenic peuvent atteindre des valeurs très importantes : les concentrations d'arsenic et de fer dans les eaux de rejets se situent entre 0,2-2,0 mg/L et 5-12 mg/L respectivement. Contrairement au cas général de drainage minier, les eaux issues de la galerie ne présentent pas de caractère acide. En effet, de puissantes formations carbonatées permettent de neutraliser l'acidité générée par l'oxydation de l'arsénopyrite. La résurgence est de 10 m³/h et présente un pH compris entre 6 et 7 et un E_h de 20 à 50 mV.

Une nouvelle souche appartenant au genre *Leptothrix* a été isolée de ce site riche en arsenic (Battaglia-Brunet *et al.*, 2006a). Au vu de ces caractéristiques et des analyses phylogénétiques et moléculaires, une nouvelle espèce est proposée : *Leptothrix arsenicoxydans*. Cette souche présente la capacité d'oxyder l'As III en As V avec une activité spécifique de 23±18 mmol/min/mg. C'est une souche mobile, aérobie et hétérotrophe capable de croître à des gammes de température de 10-37°C et de pH de 6-8.

Sa capacité à oxyder l'arsénite a été testée en « fed-batch », ainsi les conclusions peuvent être extrapolées *in situ*. En conséquence la souche jouerait un rôle pivot dans la dépollution de l'arsenic *in situ*, l'As V pouvant aisément se complexer au fer et au manganèse présent sur le site. Sa capacité à oxyder l'As III en As V peut être considérée d'importance pour l'utilisation de cette souche dans la bioremédiation du site.

La caractérisation de cette souche à fort potentiel de bioremédiation a été menée par Fabienne Battaglia-Brunet au laboratoire du BRGM (Bureau des Recherches Géologiques et Minières) à Orléans. Ma collaboration dans ce projet a porté notamment sur l'analyse qualitative et quantitative des résistances aux antibiotiques et aux métaux et métalloïdes, ainsi que les observations en microscopie électronique à transmission. Ces travaux ont donné lieu à la rédaction d'une publication présentée dans le paragraphe qui suit.

II MANUSCRIT

**First evidence of As(III)-oxidation by a strain belonging to
Leptothrix genus: *Leptothrix arsenicoxydans* sp. nov.,
isolated from arsenic-rich mine water sediment**

Fabienne Battaglia-Brunet, Caroline Michel, Dominique Breeze, **Audrey Heinrich-Salmeron**, Marie-Claire Lett, Catherine Joulian

En préparation pour soumission à Antonie van Leeuwenhoek.

First evidence of As(III)-oxidation by a strain of the *Leptothrix* genus (*Leptothrix arsenicoxydans* sp. nov.) isolated from arsenic-rich mine-water sediment

Fabienne Battaglia-Brunet^{1,5}, Caroline Michel¹, Dominique Breeze¹, Audrey Heinrich-Salmeron², Marie-Claire Lett², Catherine Joulian¹

1 BRGM, Environment & Process Division, Ecotechnology Unit, 3 av. Claude Guillemin, 45060 Orléans, France

2 Laboratoire de Génétique Moléculaire, Génomique et Microbiologie, Université de Strasbourg, CNRS-UMR7156 Equipe Ecophysiologie Moléculaire des Microorganismes 28, rue Goethe, 67083 Strasbourg, France

Running title: As(III) oxidation by *Leptothrix arsenicoxydans* sp. nov.

Keywords: arsenic, *Leptothrix*, manganese, oxidation

⁵ Corresponding author: Fabienne Battaglia-Brunet. Address: BRGM, EPI/Ecotechnology, Bat G1, 3 avenue Claude Guillemin, BP36009, 45060 Orléans cedex 02, France. E-mail: f.battaglia@brgm.fr. Phone: +33(0)2 38 64 39 30. Fax: +33(0)2 38 64 36 80

ABSTRACT

A new arsenite [As(III)]-oxidizing bacterium, isolated from arsenic-rich mine-water sediment at near-neutral pH, is proposed as a new species of the *Leptothrix* group, even though unable to produce sheaths during growth. Like other members of the genus, *L. arsenicoxydans* sp. nov. is strictly aerobic, heterotrophic, and motile with a single polar flagellum. It exhibits the classical pH (6 – 8) and temperature (10 – 37 °C) ranges for the growth of *Leptothrix* species, and excretes a manganese-oxidizing factor in the culture medium. However, *L. arsenicoxydans* can use only a few simple organic substrates and, in contrast to its closest phylogenetic relative, *L. cholodnii*, is able to oxidize As(III) in synthetic mine water. This demonstration of As(III) oxidation by a member of the widely distributed *Leptothrix* genus will contribute to understanding the combined removal of manganese and arsenic in contaminated freshwater streams and water treatment plants.

INTRODUCTION

Arsenic causes, or increases the risk of, numerous illnesses (Hopenhayn, 2006) including keratosis and skin cancer, internal cancers (such as lung and bladder cancer), and diseases of the vascular system. It is probably the environmental contaminant responsible for the highest risks of morbidity and mortality worldwide, both because of its toxicity and because of the number of people exposed. Consequently, simple and cheap technologies enabling the removal of arsenic from water need to be developed order to provide people with safe drinking water and avoid the dispersion of this highly toxic element. The contamination of surface water and sediments by mine water is a serious concern, as illustrated in Europe by the Carnoules River responsible for the dispersion of 2 to 6 tons of arsenic per year (Morin *et al.*, 2003) and the Rio Tinto River with its daily discharge of 100 kg arsenic into the Atlantic Ocean (Elbaz-Poulichet *et al.*, 2001). Although a range of physicochemical processes is available for removing arsenic from water, most require a special oxidation step to completely oxidize As(III) when present (Montiel and Welté, 1996; Jekel, 1994). The drawbacks of chemical oxidation (such as cost of oxidants, undesirable byproducts) can be avoided by using biological treatments. Many bacteria are able to perform As(III) oxidation, either autotrophically with As(III) being used as the sole energy source (Santini *et al.*, 2000; Battaglia-Brunet *et al.*, 2006a; Nakazawa and Hareyawa, 2007), or heterotrophically (Weeger *et al.*, 1999; Macur *et al.*, 2001; Salmassi *et al.*, 2002; Santini *et al.*, 2002; Bruneel *et al.*, 2003). When the water to be treated contains iron or manganese together with As(III), arsenic may be removed in a filter reactor (Katsoyiannis and Zouboulis, 2006). As(III) is chemically oxidized during the oxidation of iron by dissolved oxygen, most likely by reactive intermediates such as Fe(IV), reactive oxygen species and/or

carbonate radicals (Leupin and Hug, 2005); manganese oxides can also participate in As(III) oxidation (Zhang *et al.*, 2007). However, oxidation rates are higher when bacteria are present, and a shift from abiotic to biotic iron and manganese removal can substantially increase water-treatment plant capacity (Katsoyiannis and Zouboulis, 2006). When this process is applied to produce drinking water, the main organisms present in the bacterial population are related to the genera *Gallionella* and *Leptothrix*. Katsoyiannis *et al.* (2004) suggest that *Leptothrix* may participate in As(III) oxidation; these bacteria are heterotrophic and aerobic organisms that are known to produce sheaths and precipitate oxides of iron and manganese (Spring *et al.*, 1996), but the ability of known *Leptothrix* strains to oxidize As(III) had not been demonstrated. Casiot *et al.* (2006) isolated a bacterial strain mediating As(III) oxidation from a groundwater iron-treatment pilot; the organism was one of the dominant bacteria in the bioreactor and was considered as a new genus in that it shared only 95% identity on the 16S rDNA sequence with the closest *Leptothrix* relative. In this present article, we describe a new As(III)-oxidizing strain (S1.1^T) that we have identified as belonging to the *Leptothrix* genus. This organism, which we propose as a new species of the genus, was isolated from iron- and arsenic-rich sediment of the Loperec mine site (Finistère), where a natural biogeochemical process removing arsenic from water was observed (Battaglia-Brunet *et al.*, 2006b).

MATERIALS AND METHODS

Strains and growth conditions

Strain S1.1T was isolated in sterile demineralized water by plating dilutions of a suspension of red Fe-rich settling-pond sediment (Battaglia-Brunet *et al.*, 2006b) onto a Modified CASO1 Selective Medium (MCSM) containing 100 mg.l⁻¹ As(III) (Battaglia-Brunet *et al.*, 2002), 0.2 g.l⁻¹ yeast extract and 15 g l⁻¹ agar. An isolated colony was then picked and dispersed in liquid MCSM supplemented with arsenic and yeast extract, and subcultured five times alternately on a solid and liquid medium to obtain an axenic culture.

Leptothrix cholodnii LMG 9467T, a duplicate of LMG 7171T, was obtained from the Belgian Co-ordinated Collection of Micro-organisms Medium, Bacteria Collection (BCCM/LMG). It was maintained on solid BCCM 187 at pH 7. For biomass production while avoiding the formation of Fe and Mn precipitates, the *L. cholodnii* was cultured on MCSM containing 1 g.l⁻¹ yeast extract and 1.5 g.l⁻¹ peptone.

Strain S1.1T has been deposited in the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) as DSM XXXX and in the BCCM Bacteria Collection as LMG XXXX.

16S rRNA gene phylogeny

The almost complete 16S rRNA gene sequence (1521 bp) of strain S1.1T was obtained, as described by Battaglia-Brunet *et al.* (2006b), and deposited in the Genbank database (Benson *et al.*, 1999) under accession number DQ241397. The sequence was aligned with related sequences obtained from Genbank using the sequence aligner software from the Ribosomal Database Project II (Maidak *et al.*, 2001), and the alignment refined manually with the sequence editor BioEdit (Hall, 1999). Pairwise evolutionary distances of 1248 unambiguous nucleotides were computed with the Jukes and Cantor (1969) method, and a phylogenetic tree was constructed with the neighbor-joining method (Saito and Nei, 1987). The robustness of the tree topology was tested by bootstrap analysis with 1000 resamplings (Felsenstein, 1985). Phylogenetic programs were implemented in the Treecon package (Van de Peer and De Wachter, 1994).

DNA-DNA hybridization and G + C content

A DNA base composition determination by HPLC and a DNA-DNA hybridization of strain S1.1^T against *L. cholodnii* LMG 9467^T were performed by the DSMZ, as described by De Ley *et al.* (1970) taking into consideration the modifications described by Huss *et al.* (1983).

Physiologic and chemotaxonomic characterization

The influence of temperature (at pH 6) and pH (at 25 °C) were studied in MCSM with As(III) and 0.2 g.l⁻¹ yeast extract; the media were buffered with 25 mM MOPS (pH <7), phosphate buffer (pH 7 and 8) or glycine buffer (pH 9 and 10) for the pH tests. The yeast extract influence was determined in MCSM with As(III) and 0, 0.1, 0.2, 0.5 or 1 g.l⁻¹ yeast extract. Incubations were performed in Erlenmeyer flasks under static conditions.

The influence of aeration was studied at 25 °C in 50 ml penicillin flasks filled with 20 ml of MCSM with As(III) containing 0.5 g.l⁻¹ yeast extract. The following conditions were applied: a) agitated with a cotton stopper; b) static with a cotton stopper; c) static with a rubber stopper and 100% air; d) static with a rubber stopper, 99% N₂ and 1% O₂; and e) static with a rubber stopper and 100% N₂.

Growth on organic compounds was tested under static conditions at 25 °C in cotton-stoppered glass tubes filled with 10 ml MCSM, but lacking As(III) and yeast extract, supplemented with 20 mM of one of the following compounds: acetate, galactose, L-glutamate, pyruvic acid, succinic acid, glycogen, L-xylose, butyrate, alanine, asparagine, arginine, glycine, lysine, leucine, tryptophane, casein, D-ribose, DL-3-hydroxybutyrate, glutarate, quinate, elemental sulfur, thiosulfate. Additional substrates were tested using API strips (Biotype 100) incubated for 1 to 3 weeks at 25 °C.

The ability of strain S1.1^T to respire nitrate or chlorate was tested in 50 ml rubber-stoppered penicillin flasks filled with 40 ml MCSM, with or without As(III), containing 0.2 g.l⁻¹ yeast extract and 10 mM NaClO₃ or NaNO₃. The flasks were flushed with O₂-free N₂ before autoclaving. MCSM with 0.2 g.l⁻¹ yeast extract served as an aerobic control in a) cotton-stoppered flasks, and b) N₂-flushed rubber-stoppered flasks.

Growth was followed by measuring the optical density at 600 nm (Agilent 8453 UV-visible spectrophotometer, Palo Alto, CA, USA) or by direct counting under an optical microscope (Nikon Optiphot, x400 magnification) using a Thoma cell.

Resistance to metals and antibiotics

The Minimal Inhibitory Concentration (MIC) of metals was determined according to the method of Lim and Cooksey (1993). Briefly, bacterial suspensions were transferred to MCSM-agar plates containing 0.2 g.l⁻¹ yeast extract and increasing concentrations of metals. The MIC was defined as the metal concentration that inhibited confluent growth after 3 days at 30 °C. Antibiotic resistance was tested on MCSM-agar plates by the disc diffusion method (Courvalin *et al.*, 1985).

Microscopy

Bacteria were observed by transmission electron microscopy (TEM) using a Hitachi H-600 TEM operated at 75 kV after negative staining with 2% aqueous uranyl acetate solution. Nickel grids coated with a polyvinyl formal (Formvar[®]) film were covered with 10-20 µl of bacterial culture grown for 24 h. After 10 s, the excess culture was blotted by contact with a piece of filter paper, 5 µl uranyl acetate solution were added for staining, and the grid was air-dried for 5 min at ambient temperature.

Determination of As(III)-oxidizing activity

The bacteria were grown, until the early stationary growth phase, at 25 °C on MCSM, both with and without As(III), containing yeast extract (0.2 g.l⁻¹) and sodium thiosulfate (1 mM). For total protein extraction, the bacteria were harvested by centrifugation (15,000xg for 20 min at 4 °C), re-suspended in 800 µl MES (2[-N-morpholino]ethanesulfonic acid) buffer (50 mM, pH 6) containing a protease inhibitor (Complete Mini EDTA-free Protease inhibitor cocktail tablets, Roche), and broken in a Lysing Matrix B tube (Bio101) by shaking with a FastPrep (FP) 120 instrument (Bio 101) at a speed of 4.5 m.s⁻¹ for 30 s. The total proteins were harvested by centrifugation (14,000xg, 1 min.), then concentrated using Vivaspin 500 tubes (cutoff: 5000 Da). The protein content was determined by the Lowry method (Bio-Rad DC Protein Assay). The As(III)-oxidase activity was then measured as described by Anderson *et al.* (1992).

Determination of manganese-oxidizing activity

Bacteria were grown under agitation for 48 hours (early stationary growth phase) at 25 °C on BCCM or MCSM. The Mn²⁺-oxidizing activity was then measured on the spent culture media obtained by culture centrifugation (10,000xg for 10 min. at 4 °C) followed by filtration at 0.2 µm. The activity was measured using the TMPD (*N, N, N', N'*-tetramethyl-*p*-phenylenediamine) essay described by De Vring-De Jong *et al.* (1990). In this essay, the MnO₂ formed following the oxidation of MnCl₂ by the Mn²⁺-oxidizing factor reacts with TMPD to form Wurster blue.

Fed-batch study using a synthetic mine water

Fed-batch experiments were performed in 500 ml rubber-stoppered flasks filled with 100 g of 3-5 mm pozzolana (bacterial growth support sterilized by autoclaving three times, at 24 h intervals, for 40 min. at 120 °C). The flasks were filled under sterile air with 150 ml of synthetic water (prepared as described by Battaglia-Brunet *et al.*, 2006b) containing 20 mg.l⁻¹ yeast extract, and kept in a rubber stoppered bottle flushed with CO₂. The final pH was close to the water pH at the Loperec site (pH 6). Each flask was inoculated with 15 ml of strain S1.1^T culture prepared in MCSM containing 10 mg.l⁻¹ As(III) and 0.2 g.l⁻¹ yeast extract. Control flasks received 15 ml of strain S1.1^T culture filtered three times at 0.2 µm in order to remove all bacteria. The flasks were incubated in static conditions at 25 °C. Every day, a fraction of the liquid phase (30 ml per day the first week, then 100 ml per day) was removed and replaced by fresh synthetic water.

Analysis

For arsenic concentrations higher than 10 mg.l⁻¹, As(III) and total As were analyzed with a flame atomic absorption spectrophotometer (AAS, Varian SpectrAA 300). The As(III) was extracted in methyl isobutyl ketone with pyrrolidine dithiocarbamate as extractant, as described by Battaglia-Brunet *et al.* (2002). For arsenic concentrations lower than 10 mg.l⁻¹, As(III) and As(V) were separated using an ion exchange method as described by Battaglia-Brunet *et al.* (2006b). Total manganese and iron were analyzed by AAS. Fe(II) was quantified by the colorimetric method using 5-ortho-phenantroline and an UV-visible spectrophotometer (Agilent 8453, Palo Alto, CA, USA).

RESULTS AND DISCUSSION

Strain S1.1^T was selected among several strains able to grow in the presence of arsenic because of its ability to rapidly oxidize As(III). Individual cells of isolate S1.1^T are motile gram-negative rods, 1.5 to 2.0 µm long and 0.5 µm wide, with a single polar flagellum (Figure 1).

Its optimal growth temperature is 25 °C. No growth was observed below 9 °C or above 37 °C. The strain grew as a biofilm on the glass side of the flasks below 15 °C, and formed pellets above

30 °C. Free cells were observed between 15 and 30 °C. Strain S1.1^T grows optimally at pH 7. No growth was observed at pH lower than 6 or higher than 8. Strain S1.1^T is not able to grow autotrophically with As(III) as sole energy source. When cultured in MCSM containing 100 mg.l⁻¹ As(III) and 0.2 g.l⁻¹ yeast extract, the As(III) was completely oxidized in 30 h during the exponential growth phase (Figure 2).

Growth rate and As(III) oxidation rate increased with the addition of a yeast extract concentration up to 0.5 g.l⁻¹ (Table 1). Above this concentration, the growth rate was maintained but the As(III) oxidation rate decreased.

Specific As(III) activity at the early stationary growth phase was 23 ± 16 nmol.min⁻¹.mg⁻¹ in the presence of As(III), whereas it was not easily detectable in the absence of As(III), suggesting that the synthesis of the As(III)-oxidase is induced by As(III), as is the case for many As(III)-oxidizing microorganisms (Muller *et al.*, 2003; Silver and Phung, 2005; Michel *et al.*, 2007).

The MIC of strain S1.1^T for different metals is: Cu²⁺, 80 mg.l⁻¹; Cr(VI), <5 mg.l⁻¹; Zn²⁺, <5 mg.l⁻¹; Pb²⁺, >40 mg.l⁻¹; Mn²⁺, 100 mg.l⁻¹; As(III), 200 mg.l⁻¹; As(V), >1600 mg.l⁻¹; and Sb(III), <20 mg.l⁻¹. Strain S1.1^T presents approximately the same MIC for these different metals as the As(III)-oxidizing strain *Herminiimonas arsenicoxydans* (Muller *et al.*, 2006), apart from As(III) for which *H. arsenicoxydans* is twice as resistant. Growth of the new isolate on MCSM-agar was inhibited by tetracycline, streptomycin, rifampicin and chloramphenicol, but not by ampicillin and gentamycin.

Phylogenetic analysis on 1248 nucleotides of the 16S rRNA gene sequence of strain S1.1^T showed the strain's affiliation to the *Leptothrix* group within the β -subclass of the *Proteobacteria*. The strain's closest characterized species are *L. cholodnii* strains with % identities, just above the 97% limit for species definition (Wayne *et al.*, 1987). It is also close to the species *Ideonella dechloratans*, a facultative anaerobe able to reduce chlorate (Malmqvist *et al.*, 1994). Strain S1.1^T is only distantly related (95.5% identities) to another *Leptothrix*-related As(III)-oxidizing isolate, "strain B2", originating from a biological groundwater treatment process for Fe and As removal (Casiot *et al.*, 2006). DNA-DNA hybridization results showed that strain S1.1^T shares only 6.7% similarities with its closest phylogenetic relative *L. cholodnii* LMG 9467^T, which is less than the 24% to 33% similarities obtained between *L. cholodnii*, *L. mobilis* and *L. discophora* species (Spring *et al.*, 1996), as well as the 70% threshold value for the definition of bacterial species (Wayne *et al.*, 1987).

Interestingly, the 16S rRNA gene sequence of strain S1.1^T shares 98% identities with environmental clones retrieved from an As(III)-oxidizing biofilm at Hot Creek (California, USA) and published only in Genbank (accession numbers AY168737 and AY168738). Together, these findings suggest that, in addition to their proven implication in iron oxidation-reduction in As-contaminated sites, *Leptothrix*-related microbes play an important role in biological As(III) oxidation in the

environment. Also of note is that strain S1.1^T was isolated from an iron-rich environment where Fe(II) oxidation occurs in an oxic-anoxic transition zone; a habitat typical of the environments where the sheathed bacteria of the genus *Leptothrix* are usually found (Spring *et al.*, 1996).

Phenotypic characteristics of strain S1.1^T and its phylogenetic relatives are compared in Table 2. The GC content of strain S1.1^T is 66.2%, lower than those of the main *Leptothrix* species and *I. dechloratans*. Note that although the substrate profile of strain S1.1^T differs from those of *I. dechloratans* and the *Leptothrix* species, the ability to use substrates also differs strongly from one *Leptothrix* species to another, and even between strains of *L. cholodnii*. *L. cholodnii*^T was unable to either oxidize As(III) or grow in the presence of 100 mg.l⁻¹ of As(III).

Like the organisms of the *Leptothrix* genus which are all strictly aerobic, strain S1.1^T was unable to grow in either anaerobic or micro-aerophilic (1% O₂) conditions, although it was able to grow in a non-agitated rubber-stoppered flask with air as the gas phase. Strain S1.1^T was also unable to grow anaerobically using chlorate or nitrate as electron acceptors, unlike *I. dechloratans* which was reported to grow in these conditions. Cells of strain S1.1^T are motile with a single polar flagellum – a common feature of all currently available *Leptothrix* strains (Spring *et al.*, 1996). Conversely, *I. dechloratans* cells possess two or several polar or subpolar flagella.

One general feature of the *Leptothrix* species is their ability to produce a sheath of tube-like extracellular polymeric structures (Emerson and Ghiorse, 1993). Although the production of such a structure was never observed when growing strain S1.1^T, it should be noted that the ability to produce a sheath is not a permanent character for many *Leptothrix* strains after isolation.

Some strains of the *Leptothrix* genus, such as *Leptothrix discophora* SS-1, have been shown to excrete a manganese-oxidizing factor in their culture medium (Adams and Ghiorse, 1987; Boogerd and De Vring, 1987; De Vring-De Jong *et al.*, 1990; Emerson and Ghiorse, 1992). We therefore investigated the ability of strain S1.1^T to oxidize Mn²⁺. The Mn²⁺-oxidizing activity of spent culture media from strain S1.1^T showed that this strain secreted a manganese-oxidizing factor. However, the composition of the growth medium seems to influence the Mn²⁺-oxidizing activity; both strain S1.1^T and *Leptothrix cholodnii* LMG 9467^T exhibited a weaker (1.97 ± 0.25 nmol.ml⁻¹.min⁻¹ for *L. cholodnii*) or no (S1.1^T) activity in the BCCM compared to that quantified in the MCSM (5.32 ± 0.46 and 2.27 ± 0.59 nmol.ml⁻¹.min⁻¹, respectively). The main difference between the two media is the higher concentration of Mn²⁺ and Fe³⁺ in the BCCM. Van Veen *et al.* (1978) and Boogerd and De Vring (1987) had already observed that the manganese-oxidizing factor is able to bind to precipitated manganese dioxide, resulting in the loss of its activity -- a phenomenon that could explain the lower Mn²⁺-oxidizing activity in BCCM.

A fed-batch experiment with synthetic water showed that strain S1.1^T was able to oxidize As(III) in conditions close to those of the Loperec site; As(III) concentration was significantly lower in the inoculated flasks than in the sterile flasks (Figure 3A). As(V) concentration decreased during the course of experiment, with As(V) adsorbing onto precipitated iron hydroxides. Strain S1.1^T thus has the physiological traits for oxidizing arsenic in situ, despite being isolated and subcultured with a significantly higher arsenic concentration. There was no evidence of biological manganese (Figure 3B) and iron (Figure 3C) removal.

As strain S1.1^T was isolated from an oxic-anoxic transition zone, we tested its ability to grow with limited oxygen diffusion; the gas phase was sterile air, the flasks were rubber-stoppered and not agitated, and the fed medium was deprived of oxygen. The fed-batch experiment showed that strain S1.1^T with limited oxygenation was able to catalyze As(III) oxidation, but not iron or manganese oxidation even though the isolate had been able to synthesize a Mn²⁺-oxidizing factor in well oxygenated conditions. It had already been shown that an efficient oxygenation is necessary for the production of a Mn²⁺-oxidizing factor (Okazaki *et al.*, 1997), and that a completely aerobic environment is required for biological Mn²⁺ removal in bioreactor (Katsoyiannis and Zouboulis, 2006). This suggests that strain S1.1^T participates in As(III) oxidation, but may not be able to influence on-site iron or manganese speciation and concentrations.

CONCLUSIONS

Strain S1.1^T has been shown by its 16S rRNA gene phylogeny to belong to the *Leptothrix* genus, and it shares several physiological characters with species of this genus; i.e. pH and temperature growth range, growth only in aerobic and heterotrophic conditions, motile rod-shaped cells with one polar flagellum, and the ability to produce a Mn²⁺-oxidizing factor. In addition, strain S1.1^T presents original features of which the main one is the ability to oxidize As(III) at both high and low concentrations. This feature is not shared by its closest phylogenetic relative *L. cholodnii*. Strain S1.1^T also presents genetic divergences with *L. cholodnii*, as indicated by DNA-DNA hybridization similarities of only 6.7% between the two strains. Therefore, we propose the definition of a new species, *Leptothrix arsenicoxydans*, with isolate S1.1^T as the type strain. Because iron and manganese oxides are known to efficiently trap As(V), the demonstration of As(III)-oxidation by a member of the *Leptothrix* genus is of major importance for understanding the ecology both in As-impacted environments and in water-treatment plants using iron and manganese removal systems.

Description of *Leptothrix arsenicoxydans* sp. nov. (N.L. adj. ar.se.nic'o.xy.dans. element arsenic from Gr.n. arsên, male; Gr. adj. oxys, sour; N.L. v. oxydo, to make acid, to oxidize; N.L. part. adj. arsenicoxydans, arsenic-oxidizing).

L. arsenicoxydans has straight motile rods, 1.5 to 2.0 µm long and 0.5 µm wide, with a single polar flagellum, does not produce a sheath, and stains gram-negative. Growth takes place between 10 and 37 °C with an optimum at 25 °C, at within a pH range of 6 to 8. The bacterium is strictly aerobic and heterotrophic, and oxidizes As(III) and Mn(II). Glycogen, arginine, lysine and casein support growth. The G + C content of the DNA is 66 mol%.

ACKNOWLEDGEMENTS

This work was funded by BRGM (contribution 05721) and performed in the framework of GDR 2909 "Métabolisme de l'arsenic chez les prokaryotes". TEM analysis was performed by Mathieu Erhardt at the microscopy platform of the Institut de Biologie Moléculaire des Plantes (IBMP-CNRS, Strasbourg).

REFERENCES

- Adams L & Ghiorse W (1987) Characterisation of extracellular Mn²⁺-oxidizing activity and isolation of an Mn²⁺-oxidizing protein from *Leptothrix discophora* SS-1. *J Bacteriol* 169: 1279-1285.
- Anderson G, Williams J & Hille R (1992) The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum-containing hydroxylase. *J Biol Chem* 267: 23674-23682.
- Battaglia-Brunet F, Dictor M-C, Garrido, F, Crouzet C, Morin D, Dekeyser K, Clarens M & Baranger P (2002) An As(III)-oxidizing bacterial population: selection, characterization, and performance in reactors. *J Appl Microbiol* 93: 656-667.
- Battaglia-Brunet F, Jouliau C, Garrido F, Dictor M-C, Morin D, Coupland K, Johnson DB, Hallberg KB & Baranger P (2006a) Oxidation of arsenite by *Thiomonas* strains and characterization of *Thiomonas arsenivorans* sp. nov. *Ant Van Leeuw Intl J* 89: 99-108.
- Battaglia-Brunet F, Itard Y, Garrido F, Delorme F, Crouzet C, Greffie C & Jouliau C (2006b) A simple biogeochemical process removing arsenic from a mine drainage water. *Geomicrobiol J* 23: 201-211.
- Benson DA, Boguski MS, Lipman DJ, Ostell J, Ouellette BF, Rapp BA, Wheeler DL (1999) *Nucleic Acids Res. J* 27: 12-7
- Boogerd F & De Vring J (1987) Manganese oxidation by *Leptothrix discophora*. *J Bacteriol* 169: 489-494.
- Bruneel O, Personné J-C, Casiot C, Leblanc M, Elbaz-Poulichet F, Mahler BJ, Le Flèche A & Grimont PAD (2003) Mediation of arsenic oxidation by *Thiomonas* sp. in acid-mine drainage (Carnoules, France). *J Appl Microbiol* 95: 492-499.
- Casiot C, Pedron V, Bruneel O, Duran R, Personné J-C, Grapin G, Drakidès C & Elbaz-Poulichet F (2006) A new bacterial strain mediating As oxidation in the Fe-rich biofilm naturally growing in a groundwater Fe treatment pilot unit. *Chemosphere* 64: 492-496.
- Courvalin P., Goldstein F., Philippon A., Sirot J. (1985) *L'antibiogramme*, Paris: MPC-Videom (in French)
- De Ley J, Tijtgat R. (1970) Evaluation of membrane filter methods for DNA-DNA hybridization. *Ant Van Leeuw J* 36: 461-74
- De Vring-De Jong E, Corstjens P, Kempers E, Westbroek P & De Vrind J (1990) Oxidation of manganese and iron by *Leptothrix discophora*: use of *N, N, N', N'*-tetramethyl-*p*-phenylenediamine as an indicator of metal oxidation. *Appl Environ Microbiol* 56: 3458-3462.

- Elbaz-Poulichet F, Morley NH, Beckers JM & Nomerange P (2001) Dissolved metal fluxes through the strait of Gibraltar : the influence of the Tinto and Odiel rivers (SW Spain). *Mar Chem* 73: 193-213.
- Emerson D & Ghiorse W (1992) Isolation, cultural maintenance, and taxonomy of a sheath-forming strain of *Leptothrix discophora* and characterization of manganese-oxidizing activity associated with the sheath. *Appl Environ Microbiol* 58: 4001-4010.
- Emerson D & Ghiorse W (1993) Ultrastructure and chemical composition of the sheath of *Leptothrix discophora* SP-6. *J Bacteriol* 175: 7808-7818.
- Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783-791
- Hall, T. A. (1999). *BioEdit: a user-friendly biological sequence alignment editor and analysis*. <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>
- Hopenhayn C (2006) Arsenic in drinking water: impact on human health. *Elements* 2: 103-107.
- Huss, V. A. R., Festl, H. & Schleifer, K.-H. (1983). Studies on the spectrophotometric determination of DNA hybridization from renaturation rates. *Syst Appl Microbiol* 4: 184-192.
- Jukes TH & Cantor CR (1969) Evolution of protein molecules. In Munro HN, editor, *Mammalian Protein Metabolism*, pp. 21-132, Academic Press, New York.
- Katsoyiannis IA, Zouboulis AI & Jekel M (2004) Kinetics of bacterial As(III) oxidation and subsequent As(V) removal by sorption onto biogenic manganese oxides during groundwater treatment. *Ind Eng Chem Res* 43: 486-493.
- Katsoyiannis IA & Zouboulis AI (2006) Use of iron- and manganese-oxidizing bacteria for the combined removal of iron, manganese and arsenic from contaminated groundwater. *Water Qual Res J Canada* 41(2): 117-129.
- Jekel MR (1994) Removal of arsenic in drinking water treatment. *Arsenic in Environment, Part I: Cycling and Characterization* (Nriagu JO ed.), pp. 119-132.
- Leupin OX & Hug SJ (2005) Oxidation and removal of arsenic (III) from aerated groundwater by filtration through sand and zero-valent iron. *Wat Res* 39: 1729-1740.
- Lim CK & Cooksey DA (1993) Characterisation of chromosomal homologs of the plasmid-borne copper resistance operon of *Pseudomonas syringae*. *J Bacteriol* 175, 4492-4498.

- Macur RE, Wheeler JT, Mc Dermott TR & Inskeep WP (2001) Microbial populations associated with the reduction and enhanced mobilization of arsenic in mine tailings. *Environ Sci Technol* 35: 3676-3682.
- Maidak BL, Cole JR, Libum TG, Parker CT Jr, Saxman PR, Farris RJ, Garrity GM, Olsen GJ, Schmidt TM & Tiedje JM (2001) The RDP-II (Ribosomal Database Project). *Nucleic Acids Res.* J29: 173-4
- Malmqvist A, Welander T, Moore E, Ternstrom A, Moin G & Stenstrom IM (1994) *Ideonella dechloratans* gen. nov. sp. nov., a new bacterium capable of growing anaerobically with chlorate as an electron acceptor. *System Appl Microbiol* 17: 58-64.
- Michel C, Jean M, Coulon S, Dictor M-C, Delorme F, Morin D & Garrido F (2007) Biofilms of As(III)-oxidizing bacteria: formation and activity studies for bioremediation process development. *Appl Microbiol Biotechnol* 77: 457-467.
- Montiel A & Welté B (1996) Elimination de l'arsenic dans les eaux destinées à la consommation humaine. SAGEP- Eau de Paris. Paper given at the specialized AIDE congress "Micropolluants d'origine naturelle – arsenic et autres composés", Vienne, France.
- Morin G, Juillot F, Casiot C, Bruneel O, Personné J-C, Elbaz-Poulichet F, Leblanc M, Ildefonse P & Calas G (2003) Bacterial formation of tooeleite and mixed arsenic(III) or arsenic(V)-iron(III) gels in the Carnoulès acid mine drainage, France. A XANES, XRD, and SEM study. *Env Sci Technol* 37: 1705 - 1712.
- Muller D, Lièvreumont D, Simeonova D, Hubert J-C & Lett M-C (2003) Arsenite oxidase *aox* genes from a metal-resistant β -proteobacterium. *J Bact* 185: 135-141.
- Muller D, Simeonova DD, Riegel P, Mangenot S, Koechler S, Lièvreumont D, Bertin PN & Lett M-C (2006) *Herminiimonas arsenicoxydans* sp. nov., a metalloresistant bacterium. *Int J Syst Evol Microbiol* 56: 1765-1769.
- Nakazawa H & Hareyama W (2007) Biological oxidation of arsenite in strong acid water. *Res Proc* 54: 182-186.
- Okazaki M, Sugita T, Shimizu M, Ohode Y, Iwamoto K, De Vring-De Jong E, De Vring J & Corstjens P (1997) Partial purification and characterization of manganese-oxidizing factors of *Pseudomonas fluorescens* GB-1. *Appl Environ Microbiol* 63: 4793-4799.
- Saitou N & Nei M (1987) The neighbor-joining methods: a new method for reconstruction phylogenetic trees. *Mol. Biol. Evol.* J4:406-25

- Salmassi TM, Venkateswaren K, Satomi M, Nealson KH, Newman DK & Hering JG (2002) Oxidation of arsenite by *Agrobacterium albertimagni* AOL15 sp. nov., isolated from Hot Creek, California. *Geomicrobiol J* 19: 53-66.
- Santini JM, Sly LI, Schnagl RD & Macy JM (2000) A new chemolithoautotrophic arsenite-oxidizing bacterium isolated from a gold mine: phylogenetic, physiological, and preliminary biochemical studies. *Appl Environ Microbiol* 66, 92-97.
- Santini JM, Wen LISA, Comrie D, de Wulf-Durand P & Macy JM (2002) New arsenite-oxidizing bacteria isolated from Australian gold mining environments – phylogenetic relationships. *Geomicrobiol J* 19: 67-76.
- Silver S & Phung L. (2005) Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. *Appl Env Microbiol* 71: 599-60.
- Spring S, Kämpfer P, Ludwig W & Schleifer KH (1996) Polyphasic characterisation of the genus *Leptothrix*: new descriptions of *Leptothrix mobilis* sp. nov. and *Leptothrix discophora* sp. nov. nom. rev. and emended description of *Leptothrix cholodnii* emend. *System Appl Microbiol* 19: 634-643.
- Van de Peer Y & De Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput. Appl. Biosci.* J10:569-70
- Van Veen WL, Mulder EG & Deimema MH (1978) The *Sphaerotilus-Leptothrix* group of bacteria. *Microbiol Rev* 42: 329-356.
- Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP & Trüper HG (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Intl J Syst Bacteriol* 37: 463-464.
- Weeger W, Lièvreumont D, Perret M, Lagarde F, Hubert J-C, Leroy M & Lett M-C (1999) Oxidation of arsenite to arsenate by a bacterium isolated from an aquatic environment. *Biometals* 12: 141-149.
- Zhang G, Qu J, Liu H, Liu R & Wu R (2007) Preparation and evaluation of a novel Fe-Mn binary oxide adsorbent for effective arsenite removal. *Water Res* 41: 1921-1928.

Table 1. Influence of yeast extract concentration on growth and As(III) oxidation of strain S1.1^T.
Uncertainty in brackets represents the standard error of the mean (triplicates).

Yeast extract concentration (g.l ⁻¹)	0.1	0.2	0.5	1.0
Growth rate (h ⁻¹)	0.085 (+/- 5.10 ⁻³)	0.098 (+/- 2.10 ⁻³)	0.121 (+/- 1.10 ⁻³)	0.123 (+/- 4.10 ⁻⁴)
As(III) oxidation rate (mg l ⁻¹ h ⁻¹)	4.12 (+/- 0.67)	4.66 (+/- 0.32)	7.14 (+/- 0.37)	3.61 (+/- 0.06)

Table 2. Phenotypic characteristics of strain S1.1^T compared to those of phylogenetic relatives of the *Leptothrix* group.(*) (Malmqvist *et al.*, 1994); (**) (Spring *et al.*, 1996); ND: not determined; W: weak growth.

	Strain S 1.1 ^T	<i>I. dechloratans</i> ^T (*)	<i>Leptothrix</i> <i>cholodnii</i> LMG7171 (**)	<i>Leptothrix</i> <i>cholodnii</i> LMG8142 (**)	<i>Leptothrix</i> <i>cholodnii</i> LMG8143 (**)	<i>Leptothrix</i> <i>mobilis</i> ^T LMG17066 (**)	<i>Leptothrix</i> <i>discophora</i> ^T LMG8141 (**)
GC %	66.2	68.1	70	69	68	68	71
Temperature range (°C)	9 - 37	12 - 42	10 - 35	10 - 35	10 - 35	10 - 37	15 - 33
pH range	6.0 – 8.0	ND	6.5 – 8.5	6.5 – 8.5	6.5 – 8.5	6.5 – 8.5	6.0 – 8.0
As(III) oxidation	+	ND	-	ND	ND	ND	ND
Denitrification	-	+	ND	ND	ND	ND	ND
Chlorate reduction	-	+	ND	ND	ND	ND	ND
Acetate	-	+	+	-	-	-	-
Citrate	-	ND	-	-	-	-	-
Glucose	-	+	-	+	+	-	-
Arabinose	-	ND	ND	ND	ND	ND	ND
Cellobiose	-	ND	ND	ND	ND	ND	ND
Galactose	W	ND	-	+	+	-	-
Sucrose	-	ND	-	+	+	-	-
Mannitol	-	ND	-	+	-	-	-
Sorbitol	-	ND	ND	ND	ND	ND	ND
Glycerol	-	ND	-	+	+	-	-
Lactate	-	+	+	+	+	-	-
L-Glutamate	-	ND	+	+	+	-	-
Pyruvate	W	+	-	+	+	-	-
D-Fructose	-	+	-	+	-	-	-
Gluconate	-	ND	ND	ND	ND	ND	+
Succinate	-	+	ND	ND	ND	ND	+

Fumarate	-	-	-	-	-	+	-
Malate	-	ND	-	+	+	-	-
Glycogen	+	ND	-	-	-	-	-
L-xylose	W	ND	-	-	-	-	-
Malonate	-	ND	-	-	-	-	-
Tartrate	-	ND	-	-	-	-	-
Propionate	-	+	ND	ND	ND	ND	ND
DL-3-hydroxybutyrate	-	-	+	+	+	-	-
Butyrate	-	+	-	-	-	-	-
Alanine	W	+	-	-	-	-	-
Asparagine	-	+	ND	ND	ND	ND	ND
Arginine	+	ND	-	-	-	-	-
Glycine	-	ND	-	-	-	-	-
Lysine	+	ND	-	-	-	-	-
Proline	-	ND	+	+	+	-	+
Leucine	-	ND	+	+	+	-	-
Tryptophane	W	ND	-	-	-	-	-
Casein	+	ND	-	-	-	-	-
Benzoate	-	ND	-	-	-	-	-
D-Ribose	W	ND	+	-	-	-	-
D-Fucose	-	ND	-	-	-	-	+
Glutarate	-	ND	-	-	-	+	-
Phenylpyruvate	-	ND	+	+	+	-	-
Quinate	W	ND	+	-	-	-	+
Protocatechuate	-	ND	+	+	+	-	-
Elemental sulfur	-	ND	ND	ND	ND	ND	ND
Thiosulfate	W	ND	ND	ND	ND	ND	ND

FIGURE LEGENDS

Figure 1. Cell of strain S1.1^T observed by TEM.

Figure 2. Oxidation of As(III) during the growth of strain S1.1^T. Error bars represent the standard error of the mean (triplicates).

Figure 3. Evolution of arsenic (A), manganese (B) and iron (C) concentrations in the liquid phase during a fed-batch experiment using synthetic mine water. Error bars represent the standard error of the mean (triplicates).

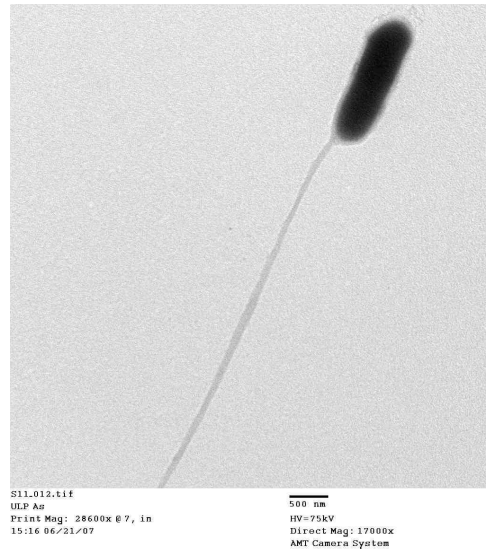


Figure 1. Cell of strain S1.1^T observed by TEM

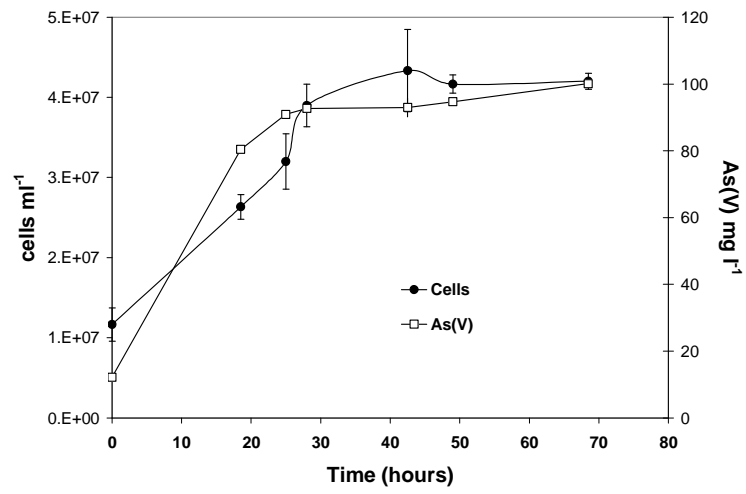


Figure 2. Oxidation of As(III) during the growth of strain S1.1^T. Error bars represent the standard error of the mean (triplicates).

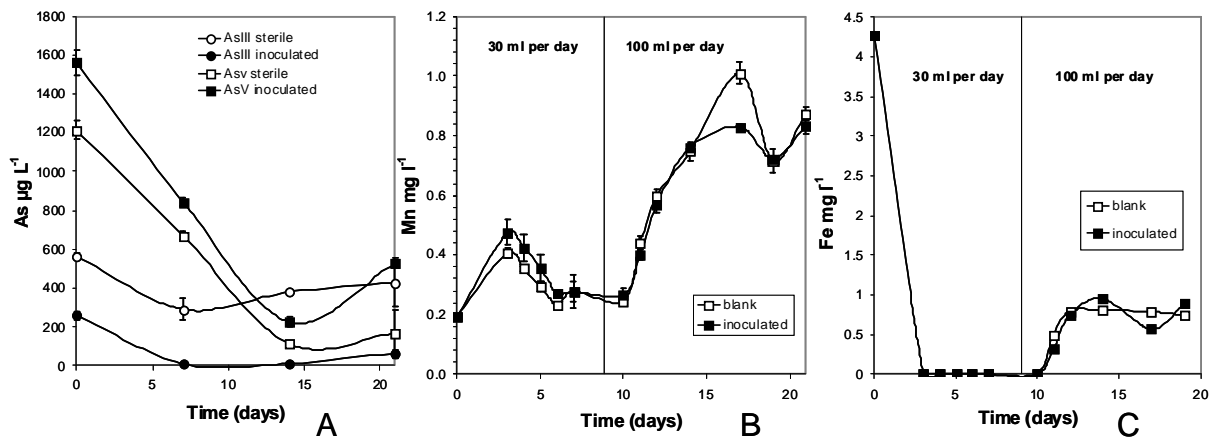


Figure 3. Evolution of arsenic (A), manganese (B) and iron (C) concentrations in the liquid phase during a fed-batch experiment using synthetic mine water. Error bars represent the standard error of the mean (triplicates).

troisième partie

Génomique descriptive et comparative d'une souche arsénite-oxydante *Thiomonas* 3As

I INTRODUCTION

Les souches *Thiomonas* sont distribuées ubiquitairement dans les AMDs et en particulier dans les sites contaminés par l'arsenic. Ainsi plusieurs souches du genre *Thiomonas* ont été isolées des sites miniers de Carnoulès dans le Gard (Bruneel *et al.*, 2003) (Duquesne *et al.*, 2003) (Battaglia-Brunet *et al.*, 2006b), de Chéni dans le Limousin (Bryan *et al.*, 2009), etc. Différents auteurs mettent en évidence le rôle clef de ces souches de *Thiomonas* dans les cycles biogéochimiques des drainages miniers acides (Baker & Banfield, 2003) (Morin *et al.*, 2003).

Parmi les différentes souches isolées de l'AMD de Canoulès (décrit dans le chapitre II de l'introduction), le génome de la souche *Thiomonas* 3As a été entièrement séquencé, annoté expertement et une analyse fine des potentialités métaboliques a été réalisée. Ce travail couplé aux analyses physiologiques (Bryan *et al.*, 2009) a permis dans un premier temps une description exhaustive du génome de *Thiomonas* 3As. Brièvement, le génome est composé d'un chromosome circulaire de 3,7 Mpb et d'un plasmide de 46,8 kpb. *Thiomonas* 3As est autotrophe, hétérotrophe facultative, elle peut oxyder des composés soufrés inorganiques (thiosulfate, sulfite, etc.) et résister à de nombreux métaux/métalloïdes (identiques à ceux trouvés sur le site), comme le cadmium, zinc, argent, cuivre, etc. La bactérie possède également la capacité de résister à l'arsenic. Deux opérons *ars* (*arsRBC* et *arsRDABC*) ont été retrouvés dans son génome. Elle est en outre capable d'oxyder l'As III en As V grâce à la présence de l'opéron *aox*. Il est important de noter le rôle de cette oxydase dans la bioremédiation du site puisque l'activité arsénite-oxydase est une clef du mécanisme d'atténuation naturelle observée à Carnoulès.

L'analyse du génome a, en outre, permis de mettre en évidence de très nombreux éléments extrachromosomiques et des îlots génomiques qui correspondent à des signatures de variabilité génétique. Afin d'analyser les similarités et les divergences génomiques, mais aussi de mettre en évidence l'évolution de ces génomes, une analyse par génomique comparative a été réalisée sur cinq souches de la même espèce (CB1, CB2, CB3 et CB6) et trois espèces de *Thiomonas* différentes (*T. arsenivorans*, *T. perometabolis* et *Thiomonas* sp. Ynys1). L'analyse des îlots génomiques grâce à l'approche CGH (Comparative Genomic Hybridization) a permis de conclure que la souche *Thiomonas* 3As a acquis une grande partie de ses capacités particulières, qui lui permettent de se développer et de coloniser cet AMD, par des transferts latéraux et des acquisitions d'îlots génomiques.

Cette étude a été réalisée dans le cadre du Groupement de Recherche (GDR) 2909 « Métabolisme de l'arsenic chez les micro-organismes : de la résistance à la détoxification » et notamment en association avec le Génoscope. Ma collaboration dans ce projet a porté sur l'annotation experte de gènes et en particulier sur l'exploration des gènes *aox*. Cette exploration a d'autre part été réalisée dans le cadre de l'analyse métagénomique du site de Carnoulès (Chapitre IV).

L'ensemble de ces analyses a donné lieu à une publication Arsène-Ploetze *et al.*, 2009 PloS Genetics, sous press, présentée ci-dessous.

II MANUSCRIT

Structure, Function and Evolution of *Thiomonas* spp.

Genome

Florence Arsène-Ploetze, Sandrine Koechler, Marie Marchal, Jean-Yves Coppee, Michael Chandler, Violaine Bonnefoy, Céline Brochier-Armanet, Mohamed Barakat, Valérie Barbe, Fabienne Battaglia-Brunet, Odile Bruneel, Christopher G. Bryan, Jessica Cleiss-Arnold, Stéphane Cruveiller, Mathieu Erhardt, **Audrey Heinrich-Salmeron**, Florence Hommais, Catherine Joulian, Evelyne Krin, Aurélie Lieutaud, Didier Lièvreumont, Caroline Michel, Daniel Muller, Philippe Ortet, Caroline Proux, Patricia Siguier, David Roche, Zoé Rouy, Gregory Salvignol, Djamila Slyemi, Emmanuel Talla, Stéphanie Weiss, Jean Weissenbach, Claudine Médigue, Philippe N. Bertin

PloS Genetics accepté

Structure, Function and Evolution of the *Thiomonas* spp. Genome

Florence Arsène-Ploetze¹, Sandrine Koechler¹, Marie Marchal¹, Jean-Yves Coppee², Michael Chandler³, Violaine Bonnefoy⁴, Céline Brochier-Armanet⁴, Mohamed Barakat⁵, Valérie Barbe⁶, Fabienne Battaglia-Brunet⁸, Odile Bruneel⁹, Christopher G. Bryan^{1†}, Jessica Cleiss-Arnold¹, Stéphane Cruveiller^{6,7}, Mathieu Erhardt¹⁰, Audrey Heinrich-Salmeron¹, Florence Hommais¹¹, Catherine Joulian⁸, Evelyne Krin¹², Aurélie Lieutaud⁴, Didier Lièvre¹, Caroline Michel⁸, Daniel Muller^{1,§}, Philippe Ortet⁵, Caroline Proux², Patricia Siguier³, David Roche^{6,7}, Zoé Rouy⁶, Gregory Salvignol⁷, Djamila Slyemi⁴, Emmanuel Talla⁴, Stéphanie Weiss¹, Jean Weissenbach^{6,7}, Claudine Médigue^{6,7}, Philippe N. Bertin¹

¹Génétique Moléculaire, Génomique et Microbiologie, UMR7156 CNRS and Université de Strasbourg, 28 rue Goethe, 67000 Strasbourg, France; ²Institut Pasteur, Genopole, Plate forme puces à ADN, F 75015 Paris, France; ³Laboratoire de Microbiologie et Génétique Moléculaire (UMR 5100), 118 Route de Narbonne, 31062 Toulouse CEDEX, France; ⁴Laboratoire de Chimie Bactérienne, Institut de Microbiologie de la Méditerranée, CNRS/Université Aix-Marseille, UPR9043, 31 chemin Joseph Aiguier 13402, Marseille Cedex 20; ⁵CEA, DSV, IBEB, LEMiRE, CNRS, Université Aix-Marseille II, CEA Cadarache, 13108 Saint-Paul-lez-Durance, France; ⁶Commissariat à l'Énergie Atomique (CEA), Direction des Sciences du Vivant, Institut de Génomique, Génoscope, 2 rue Gaston Crémieux, 91057 Evry Cedex, France; ⁷CNRS UMR8030, Génomique Métabolique, Laboratoire de Génomique Comparative, 2 rue Gaston Crémieux, 91057 Evry Cedex, France; ⁸BRGM, Environnement et Procédés, Unité Ecotechnologie, Avenue Claude Guillemin, 45060 Orléans, France; ⁹Laboratoire Hydrosociétés Montpellier, UMR 5569 (CNRS - IRD - Universités Montpellier I et II), Université Montpellier II, CC MSE, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France; ¹⁰Institut de Biologie Moléculaire des Plantes, CNRS, Université de Strasbourg, 12, rue du général Zimmer, 67000 Strasbourg; ¹¹Unité Microbiologie, Adaptation, Pathogénie CNRS-INSA-UCB UMR 5240, Université Lyon 1, Bâtiment A. Lwoff, 10 rue Raphaël Dubois, F-69622 Villeurbanne CEDEX, France; ¹²Génétique des Génomes Bactériens, URA2171, Institut Pasteur, 28 rue du Docteur Roux, 75015 Paris, France; [†]present address: Centre for Bioprocess Engineering Research (CeBER), Dept. Chemical Engineering, University of Cape Town, Rondebosch 7701, South Africa. [§]present address : Université Lyon 1, CNRS, UMR5557, Ecologie Microbienne, Villeurbanne, France.

Running head: *Thiomonas* Genome Diversity.

Correspondence: P. N . Bertin

E-mail: philippe.bertin@unistra.fr

Fax: +33 (0)3 68 85 18 15

Abbreviations: COG, Cluster of Orthologous Groups of proteins; MW, Molecular Weight; CAI, Codon Adaptation Index; CDS, CoDing Sequence; GEI, GEnomic Island; CGH, Comparative Genomic Hybridization; IS, Insertion Sequence; AMD, Acid Mine Drainage; ICE, Integrative and Conjugative Elements

Keywords: arsenic, horizontal gene transfer, evolution, genomic island, phage.

Abstract

Bacteria of the *Thiomonas* genus are ubiquitous in extreme environments, such as arsenic-rich acid mine drainage (AMD). The genome of one of these strains, *Thiomonas* sp. 3As, was sequenced, annotated and carefully examined, revealing specific adaptations allowing this bacterium to survive and grow in its highly toxic environment. In order to explore genomic diversity as well as genetic evolution in *Thiomonas* spp., a robust comparative genomic hybridization (CGH) approach was used on eight different strains of the *Thiomonas* genus, including five strains of the same species. The results obtained suggest that the *Thiomonas* genome has evolved through the gain or loss of genomic islands and that this evolution is probably influenced by the specific environmental conditions the strains live in.

Author summary

Thiomonas strains are ubiquitous in arsenic-contaminated environments. Recent advances revealed that closely related bacteria differ substantially in their response to arsenic. The genome of one of these strains, *Thiomonas* sp. 3As, was deciphered and compared to the genome of several other *Thiomonas* strains. We found that *Thiomonas* spp. evolved to survive and grow in their particular extreme environment, in part through the acquisition of multiple genomic islands. These islands confer heavy metal resistance and metabolic idiosyncrasies. Finally, our comparative genomic analyses suggest that the natural environment may influence the genomic evolution of these bacteria.

Introduction

In environments such as those impacted by acid mine drainage (AMD), high toxic element concentrations, low levels of organic matter and low pH make growth conditions extreme. AMD is generally characterized by elevated sulfate, iron and other metal concentrations, in particular, inorganic forms of arsenic such as arsenite (As(III)) and arsenate (As(V)) [1,2]. While these waters are toxic to the majority of prokaryotic and eukaryotic organisms, a few Bacteria and Archaea are not only resistant to but also able to metabolize some of the toxic compounds present [1]. Members of the *Thiomonas* genus are frequently found in AMD and AMD-impacted environments, as *Thiomonas* sp. 3As and "*Thiomonas arsenivorans*" [2-6]. These *Betaproteobacteria* have been defined as facultative Chemolithoautotrophs which grow optimally in mixotrophic media containing RISCs (Reduced Inorganic Sulfur Compounds) and organic supplements. Some strains are capable of autotrophic growth and others are capable of organotrophic growth in the absence of any inorganic energy source [5,7,8]. Recently described species and isolates include "*Tm. arsenivorans*" [9], *Tm. delicata* [10], *Thiomonas* sp. 3As [5] and Ynys1 [11]. *Thiomonas* sp. 85 3As as well as other recently isolated strains from AMD draining the Carnoulès mine site (southeastern France) containing a high arsenic concentration (up to 350 mg L⁻¹) [3,12], present interesting physiological and metabolic capacities, in particular an ability to oxidize As(III).

Over the past few years an increasing number of genomes has been sequenced, revealing that bacterial species harbour a core genome containing essential genes and a dispensable genome carrying accessory genes [13]. Some of these accessory genes are found within genomic islands (GEIs) [14] and have been acquired by horizontal gene transfer (HGT). These GEIs are discrete DNA segments (from 10 to 200 kbp) characterized by nucleotide statistics (G+C content or codon usage) that differ from the rest of the genome, and are often inserted in tRNA or tRNA-like genes. Their boundaries are frequently determined by 16-20 bp (up to 130 bp) perfect or almost perfect direct repeats (DR 96 s). These regions often 97 harbour functional or cryptic genes encoding integrases or factors involved in plasmid 98 conjugation or related to phages. GEIs encompass other categories of elements such as 99 integrative and conjugative elements (ICE), conjugative transposons and cryptic or defective 100 prophages Such GEIs are self-mobile and play an important role in genome plasticity [14]. In 101 almost all cases, GEIs have been detected *in silico*, by the comparison of closely related 102 strains. Nevertheless, the role of GEIs in genome plasticity has also been experimentally 103 demonstrated in several pathogenic bacteria such as *Staphylococcus aureus* or *Yersinia* 104 *pseudotuberculosis* [15,16] or in *Pseudomonas* sp. strain B13 isolated from a sewage treatment plant [17].

Deciphering dispensable genomes has revealed that the loss or gain of genomic islands may be important for bacterial evolution [18]. Indeed, these analyses allow the determination of the genome sequence, called pan-genome or supragenome, not just of individual bacteria, but also of entire species, genera or even bacterial kingdom [19,20]. These data result in debates on taxonomic methods used to define the bacterial species [21,22], e.g. pathogens such as *Streptococcus agalactiae* [21,23] or environmental bacteria such as *Prochlorococcus* [24,25] or *Agrobacterium* [26]. However, beyond these well-known and cultivable microorganisms, the diversity of bacteria, in particular, those found in extreme environments, has hitherto been comparatively poorly studied. Genome analysis of such extremophiles may yet reveal interesting capacities since these bacteria may express unexpected and unusual enzymes [27]. Since the role of GEIs in extremophiles has not been yet well explored, little is known about their evolution.

In the present study, the genome of *Thiomonas* sp. 3As was sequenced and analyzed. It was next compared to the genome of other *Thiomonas* strains, either of the same species or of other species of the same genus. This genome exploration revealed that *Thiomonas* sp. 3As evolved to survive and grow in 121 its particular extreme environment, probably through the acquisition of GEIs.

Results

General features of the *Thiomonas* sp. 3As genome

The genome of *Thiomonas* sp. 3As comprises a 3.7 Mbp circular chromosome and a 46.8 kbp plasmid (Table 1). The single circular chromosome contains 3,632 CDS (Table 1, Fig. S1a). The mean G+C content of the *Thiomonas* sp. 3As genome is 63.8%. However, the distribution along the genome revealed several regions with a G+C content clearly divergent from this mean value (Fig. S1). This suggests that several genomic regions are of exogenous origin. Indeed, 196 genes having mobile and extrachromosomal element functions were identified in the genome, among which a total of 91 ISs (Fig. S1a, Table S1) representing 2.5% of total CDS. None of these ISs were found as part of composite transposons, while several were identified as neighbours of phage-like site-specific recombinases.

The plasmid, pTHI, contains 68 predicted CDS. 21 genes were found in synteny with genes carried by the *R. eutropha* JMP134 plasmid pJP4, and among them, *par/trf/pem* genes necessary for plasmid partitioning, stability and replication (Fig. S1b). These observations suggest that pTHI, as JMP134, belongs to the IncP-1 α group [28]. pTHI contains 13 of the 14 genes involved in conjugation (*vir* and *tra* genes) and genes that could fulfill the function of the missing components were found on the chromosome. Therefore, *Thiomonas* sp. 3As may be able to express a complete Type IV secretory system (T4SS) of the Vir/Tra type required for pTHI conjugal transfer. IncP-1 α members are known to carry multi-resistance determinants and degradative cassettes [28], and plasmid pTHI indeed carries a Tn3-related transposon. This transposon contains part of a mercury resistance operon found in many other Gram negative bacterial transposon such as Tn21, Tn501 and Tn5053 [29].

Carbon and energy metabolism.

Thiomonas sp. 3As is able to use organic compounds as a carbon source or electron donor [5,8]. However, under certain conditions this bacterium may also be able to grow autotrophically [5]. A complete set of *cbb/rbc/cso* genes involved in carbon fixation *via* the Calvin cycle, and genes involved in glycogen, starch and polyhydroxybutyrate (PHB) biosynthesis pathways were identified (Fig. 1 and [5]). Fructose, glucose, several amino acid, C4-dicarboxylates, propionate, acetate, lactate, formate, ethanol and glycerol are potential carbon sources or electron donors, since genes involved in their import or degradation *via* the glycolysis, the Entner-Doudoroff, the TriCarboxylic Acid (TCA) or the "rubisco shunt" pathways are present in the genome. The presence of all genes involved in the oxidative phosphorylation pathway (Fig. 1) suggests that *Thiomonas* sp. 3As has a respiratory metabolism. Moreover, since several genes coding for terminal oxidases (i.e. *cbb3*, *bd* or *aa3*) were found, this respiratory metabolism may occur over a wide range of oxygen tensions. Finally, the presence of a nitrate reductase and of several formate dehydrogenases suggests that *Thiomonas* sp. 3As is able to use nitrate anaerobically as an electron acceptor and formate as electron donor. In the

absence of carbohydrates, *Thiomonas* sp. 3As is a chemolithotroph and may use reduced inorganic sulfur compounds (RISCs) as an electron donor [5]. The presence of *soxRCDYZAXB*, *dsr*, *sorAB*, *sqr* and *fccAB* genes revealed that *Thiomonas* can oxidize thiosulfate, sulfite, SO or H₂S to sulfate (Fig. 1) [30,31].

Adaptive capacities of *Thiomonas* sp. 3As to its extreme environment.

Thiomonas sp. is a moderate acidophile. Its optimum pH is 5 but this bacterium can withstand to pH as low as 2.9 (Slyemi, Johnson and Bonnefoy, personal communication). *Thiomonas* sp. 3As pH homeostasis mechanisms may therefore be strictly controlled as previously described [32,33]. First, genes encoding a potassium-transporting P type ATPase (*kdpABC*) are present in the *Thiomonas* sp. 3As genome. This ATPase could be involved in the generation of a positive internal potential produced by a greater influx of potassium ions than the outward flux of protons. Second, to strengthen the membrane, likely by lowering membrane proton conductance, *Thiomonas* sp. has cyclopropane fatty acids [5]. Accordingly, two putative *cfa* genes encoding cyclopropane fatty acid synthase have been detected. Third, cytoplasmic buffering can be mediated either by amino acid decarboxylation and/or by polyphosphate granules. 176 Genes encoding decarboxylases for lysine (4 CDS), phosphatidyl serine and glycine are present on *Thiomonas* sp. 3As genome. Moreover, urea (formed from arginine by arginase) may be degraded by urease (*ure* genes) or urea carboxylase and allophanate hydrolase. Urease encoding genes are known to be involved in acid tolerance in *Helicobacter pylori* [34]. Protons may be captured during polyphosphate synthesis. Polyphosphate granules have indeed been observed in electron micrographs of thin sections of *Thiomonas* sp. 3As [5]. Genes involved in such mechanisms (*ppk*, *pap*, *ppx*) were found in 3As genome. Fourth, primary and secondary proton efflux transporters were predicted by genome sequence analysis, including four putative Na⁺/H⁺ exchangers and voltage gated channels for chloride involved in the extreme acid resistance response in *E. coli* (*clcAB*) [35]. Finally, the elimination of organic acids can lead to pH homeostasis. Some organic acid degradation pathways have been detected in *Thiomonas* sp. such as an acetyl-CoA synthetase like. Moreover, formate oxidation was observed (Slyemi, Johnson and Bonnefoy, personal communication) and two formate dehydrogenases are encoded by the *Thiomonas* sp. 3As genome, these enzymes could convert acetate to acetyl-CoA and formate to CO₂ and hydrogen, respectively.

The Carnoulès AMD contains a high concentration of heavy metals such as zinc or lead. To resist to heavy metals, bacteria usually develop several resistance mechanism including toxic compounds extrusion pumps [36] or biofilm synthesis [37]. Flagella are important for the first steps of biofilm formation and all genes involved in motility, twitching and chemotaxis, were found in its genome. *Thiomonas* sp. 3As is motile but, unlike *H. arsenicoxydans*, this motility was not affected by

arsenic concentration (Table S2, [38]). Finally, *Thiomonas* sp. 3As is able to synthesize exopolysaccharides (Table S2), and one *eps* operon involved in their synthesis was identified in the genome, as well as two *mdoDG* clusters involved in glucan synthesis. Several genes conferring resistance to cadmium, zinc, silver, (*czc*, *sil* and *cad* genes), chromium (*chr* genes), mercury (2 *mer* operons encoding both MerA reductase but no organomercury lyase M 202 erB), copper (*cop* and *cus* genes) and tellurite (transporters THI_0898-0899) are likely involved in *Thiomonas* sp. 3As heavy metal resistance (Fig. 1). Arsenic resistance in bacteria is partly due to the expression of *ars* genes, among which, *arsC* encodes an arsenate reductase, *arsA* and *arsB* encode an arsenite efflux pump, *arsR* encodes a transcriptional regulator [39]. *Thiomonas* sp. 3As is resistant to up to 50 mM As(V) and up to 6 mM As(III) (Table S2, [8]). The analysis of the *Thiomonas* sp. 3As genome revealed the presence of two copies of the *ars* operon, an *arsRBC* operon (*ars1*) and an *arsRDABC* operon (*ars2*). *Thiomonas* sp. 3As is able to oxidize As(III) to As(V) [5] and this metabolism involves the arsenite oxidase encoded by *aoxAB* genes [5](Fig. 1). It has been shown that arsenite is imported *via* the aquaglyceroporin GlpF in *E. coli* [40]. However, as in *H. arsenicoxydans* [38], no homologue of GlpF was identified in the *Thiomonas* sp. 3As genome, suggesting that As(III) is imported *via* an unknown component.

As(III) is known to induce DNA damage and oxidative stress [41,42]. 24 genes involved in such stress responses were identified in the *Thiomonas* sp. 3As genome (Fig. 1). Moreover, this genome carries 54 genes involved in DNA repair. However, this strain lacks some genes present in *H. arsenicoxydans*, such as *alkB*, whereas two genes involved in mismatch repair were duplicated. Orthologs of genes that have been shown to be induced in response to arsenic in *H. arsenicoxydans* [38] were found in *Thiomonas* sp. 3As, i.e. *radA*, *recQ*, *ruvA*, *recA*, *xseA*, *polA*, *holB-like*, *dinB-like*, *parC*. The expression of *polA* has been previously shown to be induced in the presence of arsenic [8], suggesting that the *Thiomonas* sp. 3As response to arsenic include the expression of genes involved in DNA repair.

Comparative genomic analyses allowed the definition of 19 genomic islands (GEIs).

Several *Thiomonas* strains called CB1, CB2, CB3 and CB6 were isolated from the same environmental site as *Thiomonas* sp. 3As. The 16S rRNA/*rpoA*-based phylogeny of these isolates (>97% nucleotide identity), as well as DNA-DNA hybridization experiments (Fig. 2a) revealed that they represent different strains of the same species. All these strains are able to oxidize As(III) and are resistant to As(III) (Table S2). Nevertheless, subtle physiological differences were observed (Table S2).

The existence of both phylogenetical relationships and physiological differences between these strains prompted us to perform a comparative genome analysis in order to address the

evolution of *Thiomonas* strains. Therefore, genome variability was searched for by investigating genetic similarities and diversities among these closely related *Thiomonas* strains, using a Comparative Genomic Hybridization (CGH) approach (Fig. 3). These experiments revealed the presence of a flexible CDS (duplicated, absent or highly divergent) pool in CB1, CB6 CB3 and CB2 (Fig. 2b, Fig. 3, ArrayExpress database, accession number E-MEXP-2260) representing 2.5%, 3.2%, 24.1% and 23.1% of the genome of strain 3As (Fig. 2b), respectively. Altogether, these experiments led to the definition of 919 dispensable CDS, i.e. absent or highly divergent in at least one strain, accounting for 25.3% of strain 3As genes (Fig. 2d). The remaining conserved CDS (2713 CDS, 74.7% of the genome of strain 3As) represent a common backbone of the “core” genes of this species.

In order to enlarge our comparative analysis, genomic similarities were similarly searched for in other *Thiomonas* species: an arsenite-oxidizing strain, “*Tm. arsenivorans*”, and two closely related strains that are unable to oxidize arsenite, *Tm. perometabolis* and *Thiomonas* sp. Ynys1 (Table S2, Fig. 2, ArrayExpress database, accession number E-MEXP-2260). No significant hybridization was observed with oligomers corresponding to the plasmid, suggesting that pTHI is absent in all these strains. 18.4, 37.9 and 53.6 % of the 3As CDS were flexible in “*Tm. arsenivorans*”, Ynys1 and *Tm. perometabolis*, respectively. Altogether, 1571 CDS accounting for at least 43.3 % of the *Thiomonas* sp. 3As genome were found in the *Thiomonas* genus dispensable genome (Fig. 3d). Finally, these CGH experiments revealed that the *Thiomonas* core genome contains 2,061 CDS (56.7% of the *Thiomonas* genome). Interestingly, almost all genes involved in acid resistance described above, were found in this core genome, as for example 254 genes involved in polyphosphate granule synthesis, *cfa* and *kdp* genes, genes encoding ion transporter amino acid decarboxylase, formate dehydrogenase and other hydrogenases. One *ars* operon involved in arsenic resistance, i.e. *ars1*, and almost all genes involved in DNA repair were also conserved in all strains.

Among the flexible pool, 19 regions (ThGEI-A - ThGEI-S) had similarities with GEIs found in other bacterial genomes, suggesting that they were possibly acquired by horizontal gene transfer: (i) an abnormal deviation of the codon adaptation index (CAI) and the GC content at the 3rd nucleotide position of each codon (GC3) was observed in these regions as compared to the rest of the genome (Fig. 3), (ii) many of their genes formed syntenic blocks that differed from the general synteny observed in the rest of the genome (Table S3), (iii) genes with mobile and extrachromosomal element functions such as those coding for integrases were localized within these regions, (iv) these regions were present at the 3'-end of tRNA or miscRNA genes, and/or (v) the borders of five deletions were verified in CB strains, by PCR and direct repeats (10 to 112 bp-long) bordering these GEIs were found (Table S3).

Genetic content of the 19 genomic islands found in *Thiomonas* sp. 3As genome.

Genes found in the 19 *Thiomonas* sp. 3As GEIs and the synteny they share with genes in other bacteria are shown in Table S3. Interestingly, 70 (76.9%) of the 91 complete and partial ISs identified in the genome were located in genomic islands which represent only 21.5% of the genome (Fig. S1a). In addition to the high numbers of ISs found in these GEIs, many hypothetical proteins as well as modification/restriction enzymes were encoded by these regions. In ten GEIs, accessory genes are involved in a particular metabolism such as acetoin, atrazin, benzoate, ethyl tetra-butyl ether (ETBE) hydroxyisobutyrate phenylacetic acid and urea degradation (ThGEI-E, ThGEI-C, ThGEI-S or ThGEI-R), or heavy metal resistance (ThGEI-J, ThGEI-L, ThGEI-O).

Interestingly, several genes found in distinct GEIs shared high amino acid identity (>70 %, Fig. S2). In addition, 47 genes found in the two regions ThGEI-C and ThGEI-S shared 100 % identity. Because of this duplication, a 7 kbp region in ThGEI-S could not be sequenced and this gap may correspond to duplicated genes of ThGEI-C. These observations suggest that genomic rearrangements occurred between several GEIs. Moreover, several islands seem to be composite, since some fragments of such islands are deleted or duplicated in *Thiomonas* strains. Such composite structure may originate from insertion or excision of DNA elements in these GEIs, which involve integrase or excisionase. This hypothesis is strengthened by the observation that 32 integrases were found in almost all GEIs except for ThGEI-B and ThGEI-R. Some of such integrases are similar to phage integrases. In addition, 2 excisionases are present in ThGEI-H and ThGEI-P and such genes were localized in the vicinity of tRNA, an additional phage-like character.

One GEI, ThGEI-J, contains a prophage region (55.6 kbp) and a cluster of 6 heavy metal resistance genes (39.4 kbp, i.e., *cad*, *cus*, *sil* and *czc* genes involved in resistance to Cd, Cu, Zn, Co and Ag) (Fig. 4). The prophage region comprises 27 phage-related genes coding for structural and capsid or tail assembly proteins, replication, lysis and virulence factors. No conserved synteny with any previously described prophage could be observed. However, filamentous phage-like particles with icosahedral symmetry (capsid diameter of approximately 100 nm) and a various length tail (> 600 nm), were observed by TEM from *Thiomonas* sp. 3As liquid cultures exposed to the phage lytic phase inductor mitomycin C. Similar phage-like particles were observed in growth culture supernatants from CB1, CB3, CB6 and "*Tm. arsenivorans*" (Fig. 4) but not from CB2, Ynys1 and *Tm. perometabolis* (data not shown), in agreement with CGH results showing that the ThGEI-J is absent in these strains (Fig. 3, Table S3, ArrayExpress database, accession number E-MEXP-2260). These observations suggest that this prophage-like region may be functional in 3As, CB1, CB3, CB6 and "*Tm. arsenivorans*" under stress conditions, resulting in the formation of phage-like particles.

ThGEI-L and ThGEI-O gene content and their probable acquisition by HGT.

GEIs contribute to the adaptation of 306 microorganisms to their ecological niches and participate in genome plasticity and evolution [14]. Therefore, the environmental conditions may influence the loss or conservation of GEIs. Such hypothesis was checked by searching for genome similarities between strains originated from similar environments, i.e. AMD. To this aim, a hierarchical clustering was established based on genomic comparisons (Fig. 2C). Interestingly, the clustering obtained was different from that of the 16S rRNA/*rpoA*-based phylogenetic trees (Fig. 2A). Indeed, all strains that originated from AMD, i.e. "*Tm. arsenivorans*" and strains 3As, CB1, CB2, CB3 and CB6, grouped together, whereas Ynys1 and *Tm. perometabolis* formed a distinct group. Genes possibly dispensable for AMD survival were therefore searched for and we identified 2541 CDS conserved in all strains originated from AMD, and these CDS may constitute the "AMD" core genome of *Thiomonas*. Interestingly, several genes present in the ThGEI-L and ThGEI-O were conserved in AMD originated strains but absent in the other strains (Fig. 5).

The ThGEI-L carries genes involved in panthotenate and biotine synthesis, and may confer auxotrophy to the strains carrying this island. Moreover, genes encoding Co/Zn/Cd efflux pump were present in this GEI. In addition, this island is particularly rich in proteins with GGDEF and EAL domains. The GGDEF or EAL domain proteins are involved in either synthesis or hydrolysis of bis-(3'-5') cyclic dimeric GMP (c-di-GMP), an ubiquitous second messenger in the bacterial world that regulates cell-surface-associated traits and motility [43,44]. Because of the presence of such genes in the vicinity of 2 genes involved in chemotaxis, this island may be important for *Thiomonas* strains to form biofilm, a cellular process involved in resistance to toxic compounds [37]. Indeed, some of these genes are duplicated in CB2 and CB3 and these two strains were shown to develop better biofilm synthesis capacities (Table S2). This island also carries several genes encoding integrases and components of T4SS, such as *virB1*, *virB4*, *trbBCD*, *traCEFGI*, *mob* and a *pilE*-like gene. The presence of such genes suggests that this island originated from an integrative and conjugative element (ICE) that dissem 332 inates via conjugation [45]. These observations suggest that this island may be still mobile.

ThGEI-O contains the *aox* and *ars2* genes (Fig. 5). In addition, several other genes were found such as *mer*, *cop*, *cus* and *cad* genes involved in mercury, copper and cadmium resistance, respectively, *cys* involved in sulfate assimilation, and *moe/moa* genes involved in molybdenum cofactor biosynthesis as well as genes involved in exopolysaccharide production. The synteny of the genes found in *Thiomonas* sp. 3As ThGEI-O is not conserved in other arsenic-oxidizing bacteria (Fig. 5). Several genes present in this region are duplicated in CB1 and CB6 (i.e. the *cop* and *aox* genes), or in CB3 (i. e. *mer*, *cop*, *cus*, *dsb*, *cys*, *ars*, *moe/moa*, *aox*, and *ptxB* genes). Only a single copy of this region is present in CB2 and 3As. PCR amplification and sequencing revealed that this ThGEI-O island

is located in a different genomic region in CB2 as compared to 3As. Moreover, the *aox*, *ars* genes found in the ThGEI-O are duplicated in "*Tm. arsenivorans*" but absent or divergent in Ynys1 and *Tm. perometabolis*. Indeed, these two strains were unable to oxidize As(III), their As(III) resistance was lower than that of the other strains, and gene PCR amplification of *aox* and *ars2* failed with DNA extracted from these strains (Table S2). Altogether, the presence of at least one copy of these genes in all six strains isolated from arsenic-rich environments (i.e. 3As, CB strains and "*Tm. arsenivorans*") suggests that this GEI is of particular importance for the growth of *Thiomonas* strains in their toxic natural environment, AMD.

The evolutionary origin of the ThGEI-L and -O was investigated using two different approaches. First, we performed the phylogenetic analysis of the 196 genes contained in these two islands (Table S4, S5). The resulting trees revealed that these genes have very different evolutionary histories suggesting that the formation of ThGEI-L and -O islands occurred through the recruitment of genes from various origins by HGT (Table S4, columns 2-5). Interestingly, the closest homologue of 30/75 and 22/121 3As genes, in ThGEI-L and -O respectively, is found in other *Thiomonas* species (mainly *T. intermedia*), suggesting that the formation of these islands occurred prior to 358 the diversification of the *Thiomonas* genus and is thus relatively ancient. This hypothesis is supported by the global correspondence analysis (COA) performed on the entire genome. Our results did not reveal any particular codon usage bias, strengthening the hypothesis that these ThGEIs are ancient in *Thiomonas* genus (Fig. S3). This may explain why the major genes of these two islands are present in 3As, CB1, CB2, CB3, CB6 and "*Thiomonas arsenivorans*", as for example, the *ars2* operon and *aox* genes of the ThGEI-O. The phylogenetic analysis of *aox* genes revealed that all *Thiomonas* sequences grouped together with relationships that are very similar to organism relationships inferred with RpoA (Fig. S3a and b). This indicates that these genes were already present in the *Thiomonas* ancestor and vertically transmitted in this genus, but lost in Ynys1 and *Tm. perometabolis*. The phylogenetic analysis of the *arsB* genes, revealed that all *Thiomonas* sequences found in the ThGEI-O (i.e. *arsB2* from 3As, CB1, CB2, CB3, CB6 and "*Tm. arsenivorans*"), grouped together but not with *arsB1* genes that are part of the core genome of *Thiomonas*. Moreover, the evolutionary histories of these two proteins are different: ArsB1 proteins belong in a group containing mainly Alpha-Proteobacteria, whereas ArsB2 seems more closely related to Gamma-Proteobacteria (Fig. S3c). These observations revealed that the *ars1* and *ars2* operons were not acquired from the same source nor at the same time.

Discussion

The exploration of the *Thiomonas* sp. 3As genome suggests that this strain has a widerange of metabolic capacities at its disposal. Many of them may make this bacterium particularly well suited to survive in its extreme environment, the acidic and arsenic-rich waters draining the Carnoulès mine tailings, as for example biofilm formation and heavy metal resistance. Moreover, some metabolic capacities are unique as compared to another arsenic-resistant bacterium, whose genome has been recently sequenced and annotated, *H. arsenicoxydans*, a strict chemoorganotroph, isolated from activated sludge [38]. The first metabolic idiosyncrasy of *Thiomonas* sp. 3As is its particular carbon and energy metabolic capacities. Indeed, several organic or inorganic electron donors, such as reduced inorganic sulfur compounds [31], could be used. Second, some *Thiomonas* strains, i.e. CB1, CB3, CB6 and *T. arsenivorans*, carry two copies of the *aox* operon. As far as could be ascertained, this is the first example of *aox* gene duplication. Finally, *Thiomonas* sp. 3As is able to grow at pH3. Several genes potentially involved in acid resistance were found in *Thiomonas* genome. In addition, the Carnoulès toxic environment may cause severe DNA damage in *Thiomonas* sp. 3As, since arsenic is a co-mutagen that inhibits the DNA repair system [41]. DNA repair genes that have been previously shown to be induced in the presence of arsenic in *H. arsenicoxydans* were all found in *Thiomonas* sp. 3As genome, and the expression of *polA* has been shown to be induced in the presence of arsenic [8]. These observations suggest that this bacterium may respond to DNA damage. Nevertheless, we can hypothesize that these stressful conditions may lead to genomic rearrangements in *Thiomonas* genome. This could explain the important genomic diversity observed among the members of both the 3As species and the *Thiomonas* genus.

At the intra-species level, the dispensable genome defined by comparison of the CB strains with the 3As genome corresponds to 25.3% of *Thiomonas* sp. 3As genome. By comparison, this value is higher than that observed, with 402 the same approach, in other bacteria such as *S. agalactiae* (18 %) [46], lower than values calculated in the case of a pathogenic *E. coli* (32.4 %) [47], and similar to the value obtained in *Bacillus subtilis* (27 %) [48]. The value calculated for *Thiomonas* 3As and CB strains is very high, considering that these strains were isolated from the same site, closely related, and appear to share a recent common ancestor, as illustrated by our phylogenetical analyses. Consequently, we observed that despite strong sequence identities of housekeeping genes such as 16S rRNA or *rpoA*, the whole genome DNA-DNA hybridization value was relatively low, close to or less than 70%, for strains CB2 and CB3. Conventionally, this should indicate that these bacteria belong to separate evolutionary lineages and must be considered as different species [49]. However, the 16S rRNA-*rpoA* based analysis and CGH experiments revealed that the low DNA-DNA hybridization value correlates with the duplication or absence of several GEIs in these strains.

Consequently, we proposed that despite low DNA-DNA hybridization values, these five strains do indeed belong to the same species. Similarly, the DNA-DNA hybridization values obtained with *Thiomonas* sp. 3As as compared to strains Ynys1 and *Tm. perometabolis* were very low, as previously observed [5,10]. Altogether, the great genetic diversity observed in the present study by CGH experiments revealed that DNA-DNA hybridization method may not be appropriate to evaluate evolutionary lineages in *Thiomonas* strains. In this respect the CGH approach seems to be a reliable phylogenetic tool for typing these strains, as suggested in previous studies on other bacteria [47,50].

19 GEIs constitute a large flexible pool of accessory genes that encode adaptative traits. Some of these genes are not required for survival in AMD, since they were not found in all AMD-originated strain genome and correspond therefore to the dispensable gene pool. On the other hand, CGH-based clustering analysis revealed a significant relationship between 3As, CB1-6 and "*Tm. arsenivorans*", which originate from geographically distinct but similarly arsenic-rich environments. The *Thiomonas* sp. 3As strain and "*Tm. arsenivorans*" form two distinct groups on the basis of phylogenetical, physiological and genetic analyses. Nevertheless, the percent of flexible CDS of *Thiomonas* sp. 3As with "*Tm. arsenivorans*", is relatively low (18.4%), as compared to the value obtained with Ynys1 and *Tm. perometabolis* (37.9% and 53.6%, respectively). This value obtained with "*Tm. arsenivorans*" was in the same order of magnitude as the value obtained with CB2 and CB3 (23% and 23.6%, respectively). Altogether, 70% of the *Thiomonas* sp. 3As genome was conserved among all strains originated from AMD. Interestingly, two GEIs were conserved or duplicated in all these strains originated from AMD, i.e. ThGEI-O that carries the arsenic-specific operons *ars2* and *aox*, and genes involved in heavy metal resistance, and ThGEI-L that carries several genes involved in heavy metal resistance, biofilm formation and/or motility. Therefore, these GEIs shared by these species are presumably part of the AMD-originated *Thiomonas* core genome. This observation suggests that the acquisition or loss of these GEIs contributes to the evolution of this subgroup of the *Thiomonas* genus and that the evolution of *Thiomonas* strains has been influenced by their similar environments.

Several observations suggest that *Thiomonas* genome evolved by acquiring GEIs through horizontal gene transfer or by genome rearrangement. In the case of two islands, ThGEI-L and -O, an in-deep phylogenetic analysis revealed that these islands have a composite structure probably due to secondary acquisition or losses/rearrangements of some genes. In the case of other GEIs, the existence of HGT is suggested by the fact that genes form syntenic blocks and their GC% were divergent from the rest of the genome. Three mechanisms, i. e. conjugation, transduction and natural transformation, known to be involved in HGT in bacteria [14] may explain GEIs acquisition in *Thiomonas*. One prophage was found in the *Thiomonas* genome and may contribute to horizontal

gene transfer, as previously shown in pathogenic bacteria such as *Vibrio cholera*, *Yersinia pseudotuberculosis*, *Bartonella* [51-54], Cyanobacteria [55,56] or for the transfer of pathogenic island from *S. aureus* to *Listeria monocytogenes* [57]. In addition, genes encoding Type IV secretion systems (T4SS) were carried by the pTHI plasmid and the ThGEI-L. It has been recently proposed that GEI type T4SS are involved in the propagation of GEIs [45,58]. Therefore, it could also be possible that *Thiomonas* acquired such islands by conjugation. Orthologs of *Neisseria gonorrhoeae* genes involved in natural transformation [59] were also found in *Thiomonas* sp. 3As genome, i.e. the *pil* genes encoding a typeIV pili components and *comALMP*. This suggests that *Thiomonas* strains are able to acquire exogenous DNA. Finally, several observations suggest that *Thiomonas* genome undergone genomic rearrangements contributing to its evolution, as illustrated for the two GEIs, ThGEI-L and -O. Such rearrangements may be promoted by repeat sequences or duplications, that are at the origin of recombination [60]. Indeed, repeats sequences represent 7.62% of the *Thiomonas* sp. 3As genome, and some of the loci found in the GEIs are duplicated with high sequence identities, as ThGEI-C and ThGEI-S that are almost identical. In addition, several IS elements are highly similar, sharing more than 70% nucleotide identity. Interestingly, the majority of the ISs present in the *Thiomonas* sp. 3As genome are found in GEIs. This observation suggests that ISs duplication may have played a significant role in both assembly and evolution of these islands, or participated in GEI reshuffling. Altogether, conjugation, transduction, natural transformation and recombination may be at the origin of the high genomic content divergence observed among *Thiomonas* strains.

In conclusion, the evidence presented here suggests that *Thiomonas* sp. 3As has acquired some of its particular capacities that contribute to its survival and proliferation in AMD by horizontal gene transfer and genomic rearrangement. Furthermore, these data revealed a high degree of genetic variability within the *Thiomonas* genus, even at the intra species level. Indeed, the analysis of duplications and deletions of GEIs in several strains revealed the huge significance of these GEIs in the evolution of the *Thiomonas* genus, as well as the influence of the natural environment on the genomic evolution of this extremophile. The majority of intra or inter-species comparisons carried out thus far have concerned pathogens. Our analysis shows that GEIs play also an important role in the evolution of environmental isolates exposed to toxic elements.

Materials and Methods

Bacterial strains. *Thiomonas* sp. 3As was obtained from the acidic waters draining the Carnoulès mine tailings, southeastern France [5]. *Thiomonas* strains CB1, CB2, CB3 and CB6 were isolated from the same site: briefly, the isolates were purified by repeated single colony isolation on either R2A medium (Difco; strains CB1, CB2 and CB3) or 100:10 medium ([61]; strain CB6). Physiological, phylogenetic and genetic analyses of these four strains were performed as described previously (Table S2 and S6, [8]). Strains *Thiomonas* Ynys1 [62], *Tm. perometabolis* [63] and "*Tm. arsenivorans*" [9] were cultivated as previously described [8]. DNA-DNA hybridization was carried out as described by [64] under consideration of the modifications described by [65] using a model Cary 100 Bio UV/VIS spectrophotometer equipped with a peltier-thermostated 6x6 multicell changer and a temperature controller with *in situ* temperature probe (Varian).

DNA preparation, sequencing and annotation. DNA was extracted and purified from liquid cultures of pure isolates as previously described [8]. The complete genome sequence of *Thiomonas* sp. 3As was determined using the whole-genome shotgun method. Three libraries were constructed, two plasmids and one BAC to order contigs, as previously described [38]. From these libraries, 26,112 7,680 and 3,840 clones were end-sequenced, and the assembly was performed with the Phred/Phrap/Consed software package (www.phrap.com), as described previously [38]. An addition of 3,292 sequences was needed for the finishing phase. Coding sequences were predicted as previously described [38]. Putative orthology relationships between two genomes were defined by gene pairs satisfying either the Bidirectional Best Hit criterion or an alignment threshold (at least 40% sequence identity over at least 80% of the length of the smallest protein). These relationships were subsequently used to search for conserved gene clusters (synteny groups) among several bacterial genomes using an algorithm based on an exact graph-theoretical approach [66]. This method allowed for multiple correspondences between genes, detection of paralogy relationships, gene fusions, and chromosomal rearrangements (inversion, insertion/deletion). The 'gap' parameter, representing the maximum number of consecutive genes that are not involved in a synteny group, was set to five. Manual validation of automatic annotations was performed in a relational database (ArsenoScope, <https://www.genoscope.cns.fr/agc/mage/wwwpkgdb/MageHome/index.php?webpage=mage>) using the MaGe web interface [67]. The EMBL (<http://www.ebi.ac.uk/embl>) accession numbers for the genome of *Thiomonas* sp. 3As are FP475956 (chromosome) and FP475957 (plasmid).

Comparative genome hybridization (CGH) array. A custom 385K array for the *Thiomonas* sp. 3As chromosome and plasmid was designed and constructed by NimbleGen Systems. This DNA array encompasses 3,645 CDS of the 3As genome. Probe length was 50 nt and current mean probe spacing was 7 nt. Genomic DNAs from all strains were extracted with the Wizard® Genomic DNA Purification

Kit (Promega). DNA samples were labeled and purified using the BioPrime® Array CGH Genomic Labeling System protocol (Invitrogen). Test (Cy[®]3-labeled) and reference (*Thiomonas* sp. 3As genomic DNA, Cy[®]5-labeled) genomic DNAs were combined (400 pmol fluorescent dye each) and were co-hybridized to the array for 16 h at 42°C in a MAUI Hybridization System (BioMicro System) and slides were washed according to NimbleGen's recommendations. Dye swap experiments were used to compare *Thiomonas* sp. 3As and *Thiomonas* sp. CB1 genomic DNAs. Arrays were scanned with an Axon 4000B scanner. Data were acquired and analysed using NimbleScan 2.0 and SignalMap 1.9 software (Roche, NimbleGen) and analysed using the Partek Genomics Suite software (Partek Incorporated, St. Louis, Missouri, U.S.A.). Briefly, log₂-ratios (Cy[®]3/Cy[®]5) were calculated using the segMNT algorithm and gains and losses of genomic material were identified using Partek Genomics Suite Software, as follows: the files were imported and normalized with the qspline normalization [68] by NimbleScan. These data were then imported into the Partek Genomics Suite Soft. The segmentation was performed using the circular binary segmentation algorithm 535 from Olshen *et al.* [69]. Permutations are used to provide the reference distribution to check a second time. 1000 permutations are run using the Partek software. If the resulting p-value is below the threshold (p=0.01), then a breakpoint is added. To verify deletions, PCRs were performed as described in supplementary materials, using primers designed to anneal at the borders of the expected deletions. The CGH data are available in the database ArrayExpress, with the accession number E-MEXP-2260.

Phage excision and electron microscopy. Phage formation was induced by treating exponential cultures with mitomycin C (0.5 µg/mL) for 24h. The suspension was negatively stained with 16% ammonium molybdate for 10 seconds and dried over Formvar-coated nickel grids. Grids were examined at 40,000-fold magnification using a Hitachi 600 transmission electron microscope at 75 kV and photographed using a Hamamatsu ORCA-HR camera (Hamamatsu City, Shizuoka, Japan) with the AMT software (Advanced Microscopy Techniques Corp., Danvers, MA).

Phylogenetic and correspondence analysis.

For Each CDS, homologues were searched in NCBI databases. The 300 sequences with the best score were aligned using ClustalW [70]. Alignments were checked by hand and positions with more than 5% of gaps were automatically removed. Neighbour-Joining trees were constructed and analyzed to determine the evolutionary origin of each CDS (Table S4). The correspondence analysis (COA) [71] was performed using the library FactoMineR (<http://factominer.free.fr>) from the statistical package R (<http://www.r-project.org>). For all annotated genes of *Thiomonas* sp. 3As, we determined all the relative synonymous codon usage values [72] obtaining a matrix where the rows represent the genes and the 57 columns are the RSCU values for individual codons. As usual the 3 TER codons were excluded from the analysis. Codons corresponding to Cystein (TGC/TGT) and the

duet of Arginin(AGG/AGA) were also removed from 560 the analysis as they induce systematic artefactual biases [73].

Acknowledgments

Tm. perometabolis was obtained from the Pasteur Institute, Paris, France. Ynys1 was kindly donated by Kevin Hallberg and D. B. Johnson (Bangor University, North Wales, U.K.). The study was financed by the ANR 07-BLAN-0118 project (Agence Nationale de la Recherche). Annotation and database management were supported by a grant from ANR 07- PFTV MicroScope project. CGB was supported by a grant from the University of Strasbourg, MM by a grant from ANR COBIAS project (PRECODD 2007, Agence Nationale de la Recherche), JC-A, A H-S, SW and DS by a grant from the Ministère de la Recherche et de la Technologie and AL by a grant from ADEME-BRGM. This work was performed within the framework of the French research network "Arsenic metabolism in microorganisms" (GDR2909-CNRS).

Authors' contributions

FA-P explored the genome, analysed the CGH results and wrote the manuscript. SK and CP performed the CGH experiments that were coordinated by J-YC. MM carried out motility tests, phage observation and helped to draft the manuscript. CGB carried out the physiological and molecular genetic studies. OB isolated CB strains. CB-A carried out the phylogenetic analyses. PB conceived the study, coordinated it, and helped to write the manuscript. All other authors annotated and explored part of the genome and read and commented on the manuscript.

References

1. Johnson DB, Hallberg KB (2003) The microbiology of acidic mine waters. *Res Microbiol* 154: 466-473.
2. Hallberg KB, Johnson DB (2005) Microbiology of a wetland ecosystem constructed to remediate mine drainage from a heavy metal mine. *Sci Total Environ* 338: 53-66.
3. Bruneel O, Personné JC, Casiot C, Leblanc M, Elbaz-Poulichet F, et al. (2003) Mediation of arsenic oxidation by *Thiomonas* sp. in acid-mine drainage (Carnoulès, France). *J Appl Microbiol* 95: 492-499.
4. Battaglia-Brunet F, Dictor MC, Garrido F, Crouzet C, Morin D, et al. (2002) An arsenic(III)-oxidizing bacterial population: selection, characterization, and performance in reactors. *J Appl Microbiol* 93: 656-667.
5. Duquesne K, Lieutaud A, Ratouchniak J, Muller D, Lett MC, et al. (2008) Arsenite oxidation by a chemoautotrophic moderately acidophilic *Thiomonas* sp.: from the strain isolation to the gene study. *Environ Microbiol* 10: 228-237.
6. Coupland K, Battaglia-Brunet F, Hallberg KB, Dictor MC, Garrido F, et al. (2004) Oxidation of iron, sulfur and arsenic in mine waters and mine wastes: an important role of novel *Thiomonas* spp. In: Tsezos AHM, Remondaki E, editors. *Biohydrometallurgy: a sustainable technology in evolution*. Zografou, Greece: National Technical University of Athens. pp. 639-646.
7. Moreira D, Amils R (1997) Phylogeny of *Thiobacillus cuprinus* and other mixotrophic thiobacilli: proposal for *Thiomonas* gen. nov. *Int J Syst Bacteriol* 47: 522-528.
8. Bryan CG, Marchal M, Battaglia-Brunet F, Klugler V, Lemaitre-Guillier C, et al. (2009) Carbon and arsenic metabolism in *Thiomonas* strains: differences revealed diverse adaptation processes. *BMC Microbiol* 9: 127.
9. Battaglia-Brunet F, Joulian C, Garrido F, Dictor MC, Morin D, et al. (2006) Oxidation of arsenite by *Thiomonas* strains and characterization of *Thiomonas arsenivorans* sp. nov. *Antonie van Leeuwenhoek* 89: 99-108.
10. Katayama Y, Uchino Y, Wood AP, Kelly DP (2006) Confirmation 610 of *Thiomonas delicata* (formerly *Thiobacillus delicatus*) as a distinct species of the genus *Thiomonas* Moreira and Amils 1997 with comments on some species currently assigned to the genus. *Int J Syst Evol Microbiol* 56: 2553-2557.
11. Hallberg KB, Johnson DB (2003) Novel acidophiles isolated from moderately acidic mine drainage waters. *Hydrometallurgy* 71: 139-148.

12. Casiot C, Morin G, Juillot F, Bruneel O, Personné JC, et al. (2003) Bacterial immobilization and oxidation of arsenic in acid mine drainage (Carnoulès creek, France). *Water Res* 37: 2929-2936.
13. Bentley S (2009) Sequencing the species pan-genome. *Nat Rev Microbiol* 7: 258-259.
14. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, et al. (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* 33: 376-393.
15. Maiques E, Ubeda C, Tormo MA, Ferrer MD, Lasa I, et al. (2007) Role of staphylococcal phage and SaPI integrase in intra- and interspecies SaPI transfer. *J Bacteriol* 189: 5608-5616.
16. Lesic B, Carniel E (2005) Horizontal transfer of the high-pathogenicity island of *Yersinia pseudotuberculosis*. *J Bacteriol* 187: 3352-3358.
17. Sentchilo V, Ravatn R, Werlen C, Zehnder AJ, van der Meer JR (2003) Unusual integrase gene expression on the *clc* genomic island in *Pseudomonas* sp. strain B13. *J Bacteriol* 185: 4530-4538.
18. Hacker J, Kaper JB (2000) Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* 54: 641-679.
19. Tettelin H, Riley D, Cattuto C, Medini D (2008) Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* 11: 472-477.
20. Medini D, Donati C, Tettelin H, Massignani V, Rappuoli R (2005) The microbial pan-genome. *Curr Opin Genet Dev* 15: 589-594.
21. Achtman M, Wagner M (2008) Microbial 635 diversity and the genetic nature of microbial species. *Nat Rev Microbiol* 6: 431-440.
22. Fraser C, Alm EJ, Polz MF, Spratt BG, Hanage WP (2009) The bacterial species challenge: making sense of genetic and ecological diversity. *Science* 323: 741-746.
23. Tettelin H, Massignani V, Cieslewicz MJ, Donati C, Medini D, et al. (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pangenome". *Proc Natl Acad Sci U S A* 102: 13950-13955.
24. Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, et al. (2007) Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet* 3: e231.
25. Coleman ML, Sullivan MB, Martiny AC, Steglich C, Barry K, et al. (2006) Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* 311: 1768-1770.
26. Costechareyre D, Bertolla F, Nesme X (2009) Homologous recombination in *Agrobacterium*: potential implications for the genomic species concept in bacteria. *Mol Biol Evol* 26: 167-176.
27. Rothschild LJ, Mancinelli RL (2001) Life in extreme environments. *Nature* 409: 1092-1101.

28. Schluter A, Szczepanowski R, Puhler A, Top EM (2007) Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. *FEMS Microbiol Rev* 31: 449-477.
29. Nascimento AM, Chartone-Souza E (2003) Operon *mer*: bacterial resistance to mercury and potential for bioremediation of contaminated environments. *Genet Mol Res* 2: 92-101.
30. Friedrich CG, Bardischewsky F, Rother D, Quentmeier A, Fischer J (2005) Prokaryotic sulfur oxidation. *Curr Opin Microbiol* 8: 253-259.
31. Frigaard NU, Dahl C (2009) Sulfur metabolism in phototrophic sulfur bacteria. *Adv Microb Physiol* 54: 103-200.
32. Slonczewski JL, Fujisawa M, Dopson M, Krulwich TA (2009) Cytoplasmic pH measurement and homeostasis in bacteria and archaea. *Adv Microb Physiol* 55: 1-79, 317.
33. Baker-Austin C, Dopson M (2007) Life in acid: pH homeostasis 660 in acidophiles. *Trends Microbiol* 15: 165-171.
34. Stingl K, Altendorf K, Bakker EP (2002) Acid survival of *Helicobacter pylori*: how does urease activity trigger cytoplasmic pH homeostasis? *Trends Microbiol* 10: 70-74.
35. Iyer R, Iverson TM, Accardi A, Miller C (2002) A biological role for prokaryotic ClC chloride channels. *Nature* 419: 715-718.
36. Silver S, Phung le T (2005) A bacterial view of the periodic table: genes and proteins for toxic inorganic ions. *J Ind Microbiol Biotechnol* 32: 587-605.
37. Harrison JJ, Ceri H, Turner RJ (2007) Multimetal resistance and tolerance in microbial biofilms. *Nat Rev Micro* 5: 928-938.
38. Muller D, Médigue C, Koechler S, Barbe V, Barakat M, et al. (2007) A tale of two oxidation states: bacterial colonization of arsenic-rich environments. *PLoS Genet* 3: e53.
39. Silver S, Phung LT (2005) Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. *Appl Environ Microbiol* 71: 599-608.
40. Meng YL, Liu Z, Rosen BP (2004) As(III) and Sb(III) uptake by GlpF and efflux by ArsB in *Escherichia coli*. *J Biol Chem* 279: 18334-18341.
41. Hartwig A, Blessing H, Schwerdtle T, Walter I (2003) Modulation of DNA repair processes by arsenic and selenium compounds. *Toxicology* 193: 161-169.
42. Schwerdtle T, Walter I, Mackiw I, Hartwig A (2003) Induction of oxidative DNA damage by arsenite and its trivalent and pentavalent methylated metabolites in cultured human cells and isolated DNA. *Carcinogenesis* 24: 967-974.
43. Schirmer T, Jenal U (2009) Structural and mechanistic determinants of c-di-GMP signalling. *Nat Rev Microbiol* 7: 724-735.

44. Hengge R (2009) Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* 7: 263-273.
45. Burrus V, Waldor MK (2004) Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* 155: 376-386.
46. Tettelin H, Massignani V, Cieslewicz MJ, Eisen JA, Peterson S, 686 et al. (2002) Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *Proc Natl Acad Sci U S A* 99: 12391-12396.
47. Fukiya S, Mizoguchi H, Tobe T, Mori H (2004) Extensive genomic diversity in pathogenic *Escherichia coli* and *Shigella* strains revealed by comparative genomic hybridization microarray. *J Bacteriol* 186: 3911-3921.
48. Earl AM, Losick R, Kolter R (2007) *Bacillus subtilis* genome diversity. *J Bacteriol* 189: 1163-1170.
49. Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, et al. (1987) Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* 37: 463-464.
50. Binnewies TT, Motro Y, Hallin PF, Lund O, Dunn D, et al. (2006) Ten years of bacterial genome sequencing: comparative-genomics-based discoveries. *Funct Integr Genomics* 6: 165-185.
51. Li M, Kotetishvili M, Chen Y, Sozhamannan S (2003) Comparative genomic analyses of the *Vibrio* pathogenicity island and cholera toxin prophage regions in nonepidemic serogroup strains of *Vibrio cholerae*. *Appl Environ Microbiol* 69: 1728-1738.
52. Derbise A, Chenal-Francisque V, Pouillot F, Fayolle C, Prevost MC, et al. (2007) A horizontally acquired filamentous phage contributes to the pathogenicity of the plague *Bacillus*. *Mol Microbiol* 63: 1145-1157.
53. Brussow H, Canchaya C, Hardt WD (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* 68: 560-602.
54. Berglund EC, Frank AC, Calteau A, Vinnere Pettersson O, Granberg F, et al. (2009) Run-off replication of host-adaptability genes is associated with gene transfer agents in the genome of mouse-infecting *Bartonella grahamii*. *PLoS Genet* 5: e1000546.
55. Dammeyer T, Bagby SC, Sullivan MB, Chisholm SW, Frankenberg-Dinkel N (2008) Efficient phage mediated pigment biosynthesis in oceanic cyanobacteria. *Curr Biol* 18: 442-448.
56. Zeidner G, Bielawski JP, Shmoish M, Scanlan DJ, Sabehi 711 G, et al. (2005) Potential photosynthesis gene recombination between *Prochlorococcus* and *Synechococcus* via viral intermediates. *Environ Microbiol* 7: 1505-1513.
57. Chen J, Novick RP (2009) Phage-mediated intergeneric transfer of toxin genes. *Science* 323: 139-141.

58. Juhas M, Crook DW, Hood DW (2008) Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell Microbiol* 10: 2377-2386.
59. Hamilton HL, Dillard JP (2006) Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol Microbiol* 59: 376-385.
60. Treangen TJ, Abraham AL, Touchon M, Rocha EP (2009) Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS Microbiol Rev* 33: 539-571.
61. Schrader JA, Holmes DS (1988) Phenotypic switching of *Thiobacillus ferrooxidans*. *J Bacteriol* 170: 3915-3923.
62. Hallberg KB, Johnson DB (2003) Novel acidophiles isolated from moderately acidic mine drainage waters. *Hydrometallurgy* 71: 139-148.
63. London J, Rittenberg SC (1967) *Thiobacillus perometabolis* nov. sp., a non-autotrophic thiobacillus. *Arch Mikrobiol* 59: 218-225.
64. De Ley J, Cattoir H, Reynaerts A (1970) The quantitative measurement of DNA hybridization from renaturation rates. *Eur J Biochem* 12: 133-142.
65. Huss VAR, Festl H, Schleifer KH (1983) Studies on the spectrophotometric determination of DNA hybridization from renaturates rates. *Syst Appl Microbiol* 4: 184-192.
66. Boyer F, Morgat A, Labarre L, Pothier J, Viari A (2005) Syntons, metabolons and interactons: an exact graph-theoretical approach for exploring neighbourhood between genomic and functional data. *Bioinformatics* 21: 4209-4215.
67. Vallenet D, Labarre L, Rouy Z, Barbe V, Bocs S, et al. (2006) MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res* 34: 53-65.
68. Workman C, Jensen LJ, Jarmer H, Berka R, Gautier L, et al. (2002) A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biol* 3: research0048.0041-0048.0016.
69. Olshen AB, Venkatraman ES, Lucito R, Wigler M (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 5: 557-572.
70. Thompson JD, Plewniak F, Thierry J, Poch O (2000) DbClustal: rapid and reliable global multiple alignments of protein sequences detected by database searches. *Nucleic Acids Res* 28: 2919-2926.
71. Benzécri JP (1984) L'analyse des correspondances. Paris: Dunod.
72. Sharp PM, Tuohy TM, Mosurski KR (1986) Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res* 14: 5125-5143.
73. Perriere G, Thioulouse J (2002) Use and misuse of correspondence analysis in codon usage studies. *Nucleic Acids Res* 30: 4548-4555.

Figure Legend

Figure 1. Schematic representation of the major metabolic pathways in *Thiomonas* sp.

3As. These metabolic pathways were ascertained using genome and physiological data ([5], this study). Light blue, carbon metabolism; Red, oxidative phosphorylation; Green, stress response; Yellow, heavy metal resistance; Orange, iron metabolism; Dark blue, phosphate import (*pst* genes).

Figure 2. Genomic diversity among *Thiomonas* strains. **a.** phylogenetic dendrogram of the SuperGene construct of both the 16S rRNA and *rpoA* genes of the *Thiomonas* strains used in this study. *Ralstonia eutropha* H16 served as the outgroup. The DNA-DNA hybridization percentage is included in brackets. nd: not determined (CB1 and CB6 were almost identical at the genetic content level. For this reason, DNA-DNA hybridization analyses were performed only with CB1). Numbers at the branches indicate percentage bootstrap support from 500 resamplings for ML analysis. NJ analyses produced the same branch positions (not shown). The scale bar represents changes per nucleotide; **b.** number of variable CDS exhibited by each strain determined by CGH, expressed as a percent of the total number of 3As CDS; **c.** hierarchical clustering established based on CGH experiments, represented as a composite view of genome diversity in *Thiomonas* strains compared to 3As. The column represents genes as they are found along 3As genome, starting from the origin of replication (THI0001). Results for each strain are shown in each row. Red colour indicates absence or strong divergence leading to reduced hybridization efficiency as compared to the corresponding *Thiomonas* sp. 3As gene ($\text{Log}_2(\text{A635}/\text{A532}) \leq -1$); Green colour indicates a $\text{Log}_2(\text{A635}/\text{A532}) \geq 1$, suggesting duplication of this gene. The regions corresponding to ThGEI-A, B, C, E, F, G, H, I, J, K, O, P, Q, R and S are indicated; **d.** Percent of CDS found in core (in grey) or in dispensable (in red) genome, in 3As and CB strains (1st column), in AMD-originated strains (3As, CB strains and "*Tm. 779 arsenivorans*", 2nd column), or in all *Thiomonas* strains (3rd column).

Figure 3. Deleted or duplicated regions in CB strains, obtained by CGH experiments.

The codon adaptation index (CAI) and the GC content at the 3rd nucleotide position of each codon (GC3) of *Thiomonas* sp. 3As is shown in the upper part. Red stars: variable regions found in CB strains and corresponding to GEIs; Blue stars: other variable regions.

Figure 4. Prophage-like region and phage-like particle. **a.** schematic representation of the prophage-like region (55.6 kbp) and the contiguous metal resistance gene cluster (39.4 kbp). The total region is flanked by two ISs (light grey rectangles), and a partial transposase gene (THI_1986, dark grey

rectangle) delimiting the two clusters. Genes in italics are duplicated in the *Thiomonas* sp. 3As genome; **b.** organisation of the prophage-like region. The 71 putative CDS, coloured according to their category, are shown by arrows representing their direction. This cluster corresponds to the unique prophage-like region of the genome, containing notably structural, lysis and virulence associated phage genes; **c.** transmission-electron micrographs of tailed bacteriophage-like particles from *Thiomonas* species. No bacteriophage-like particles were observed in *Thiomonas* sp. CB2 culture supernatant. Cultures were amended with mitomycine C (0.5 µg/mL). Bars : 100 nm. The 3As phage-like particle is hypothetically coded by the prophage-like region described above.

Figure 5. Genomic Island ThGEI-O containing genes involved in arsenic metabolism, and ThGEI-L.

The best synteny results of the genes was obtained with genes of *Polaromonas naphthalenivorans* CJ2, *Leptothrix cholodnii* SP-6, *Rhodopseudomonas palustris* BisB18, *Achromobacter denitrificans* (ThGEI-L) or *Polaromonas* sp. JS666, *Rhodobacter ferrireducens*, *H. arsenicoxydans*, *Acidithiobacillus ferrooxidans* ATCC23270, and *Cupriavidus metallidurans* 805 CH34 pMOL30 (ThGEI-O). Blue circles: genes originated from phage; Blue stars: transposases; Blue arrows: direct repeats. The genes conserved in 3As and CB strains or in AMD-originated strains (3As, CB strains and "*Tm. arsenivorans*"), are indicated by pink and red lines, respectively.

Table 1. General genome and plasmid features

Molecule	Category	Feature	Value			
Plasmid	General characteristics	Size (bp)	46,756			
		GC content (%)	60.49			
		Coding density (%)	88.19			
		Predicted CDSs	68			
	Proteins with predicted function	with predicted	Percent of total CDSs	48.52		
			Secretion (%)	11.76		
			Partitioning	2.94		
			Replication, recombination	8.82		
			Inorganic ion transport and metabolism (Hg)	2.94		
			Proteins without predicted function	without predicted	Conserved hypothetical proteins (%)	13.24
					No homology to any previously reported sequences (%)	38.24
	Percent of total CDSs	51.48				
	Genome	General characteristics	Size (bp)	3,738,778		
			GC content (%)	63.8		
Coding density (%)			90.01			
16S-23S-5S rRNA operons			1			
tRNAs			43			
Predicted CDSs			3,632			
Proteins with predicted function		with predicted	Percent of total CDSs	74.2		
			Heavy metal resistance (%)	1.6		
			Related to arsenic metabolism/transport (%)	0.6		
Proteins without predicted function		without predicted	Conserved hypothetical proteins (%)	13.54		
			Hypothetical proteins (%)	11.95		
			Percent of total CDSs	25.8		
Mobile and extrachromosomal element functions				196		
			Transposases (nb CDSs)	101		
	Phage related (nb CDSs)		61			
Repeated regions (%)			7.62			

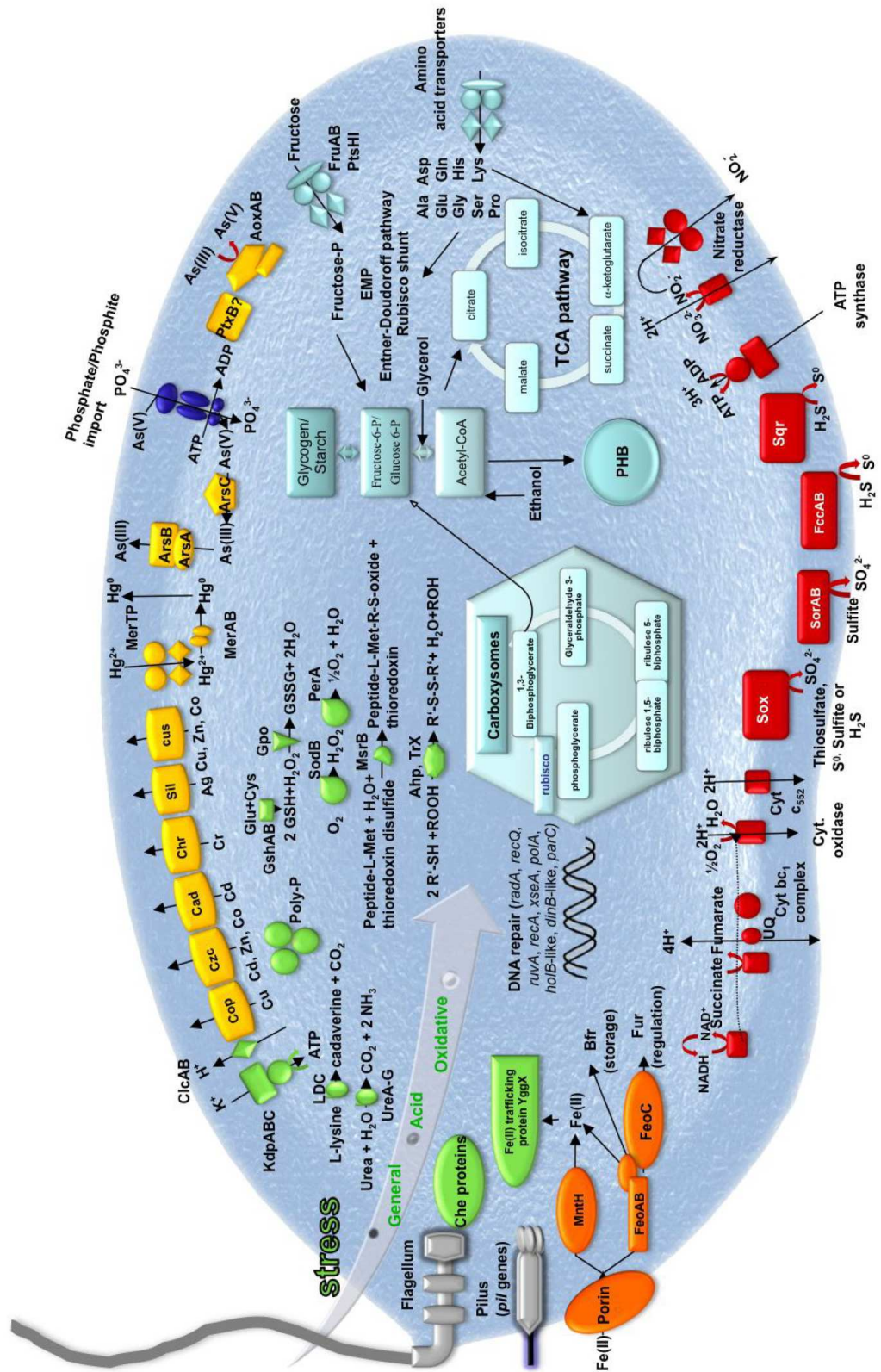


Figure 1

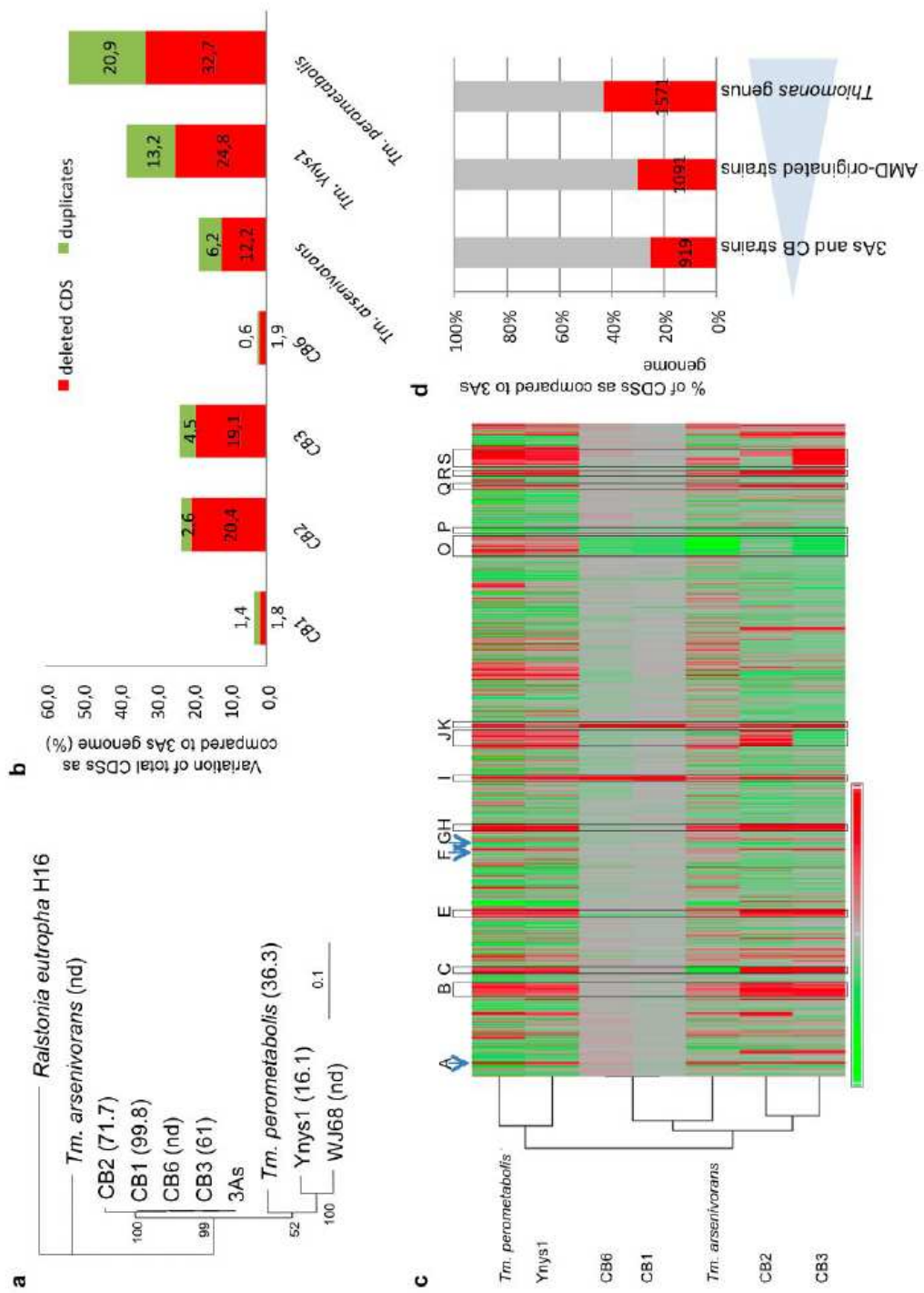


Figure 2

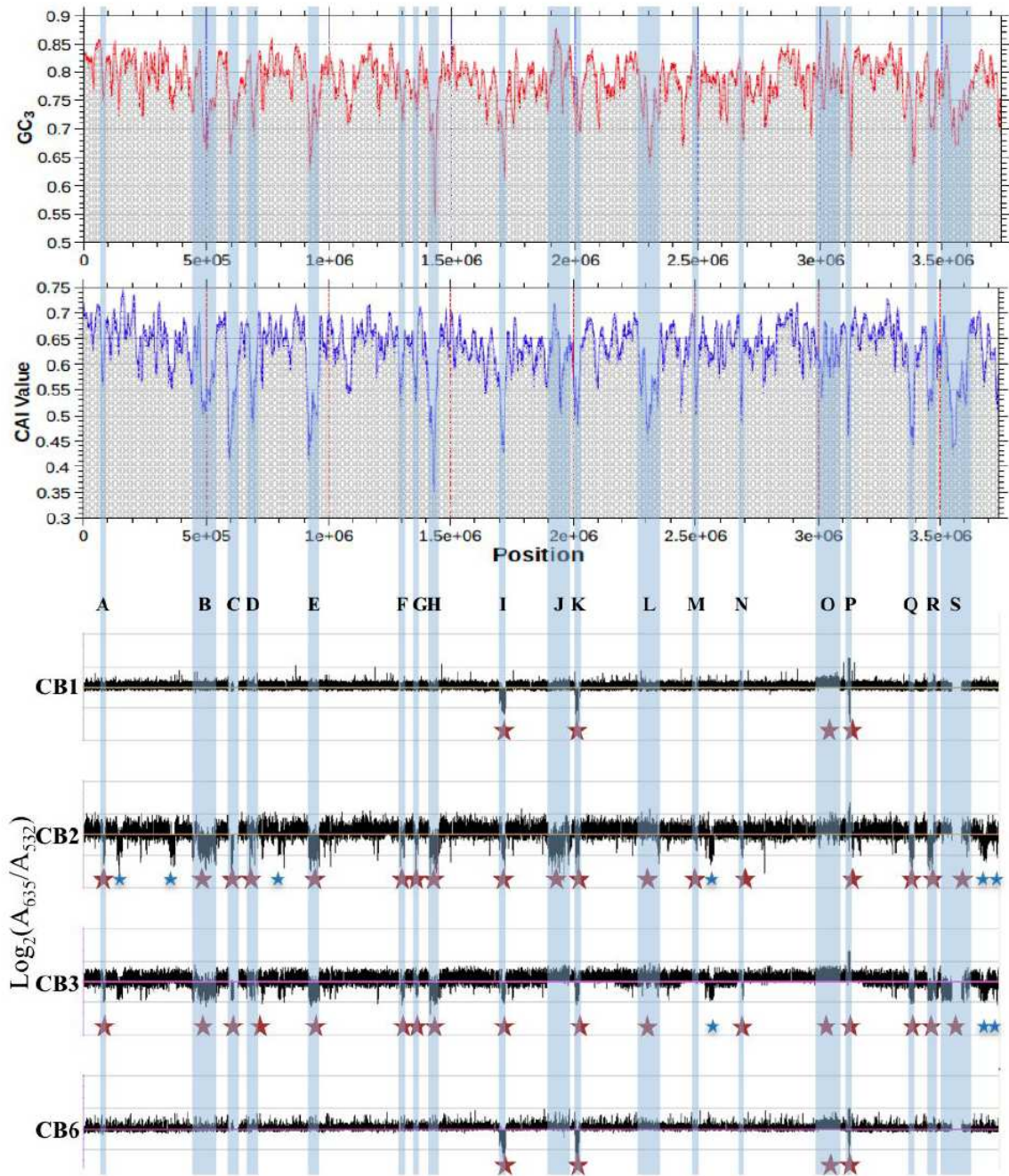


Figure 3

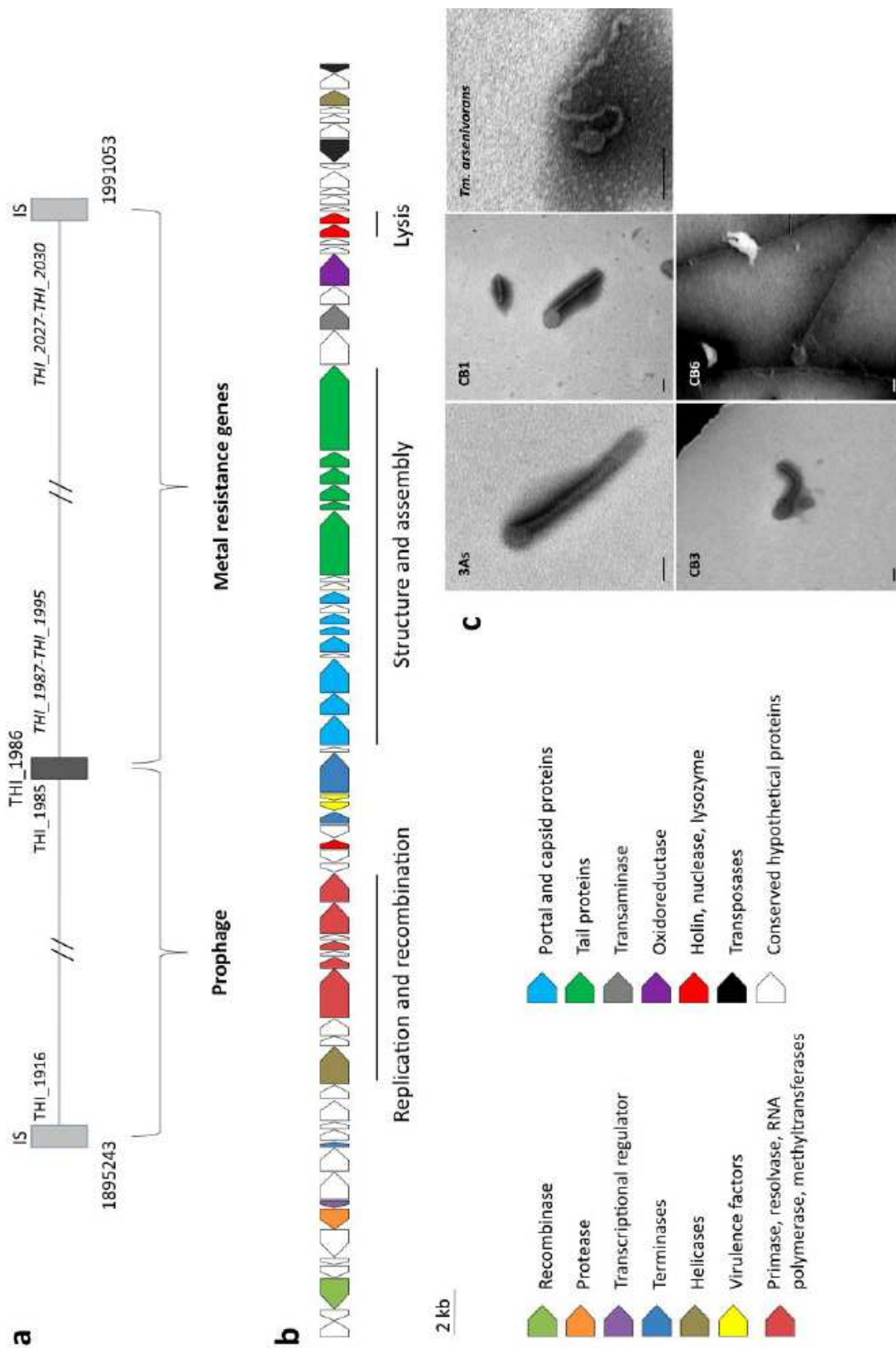


Figure 4

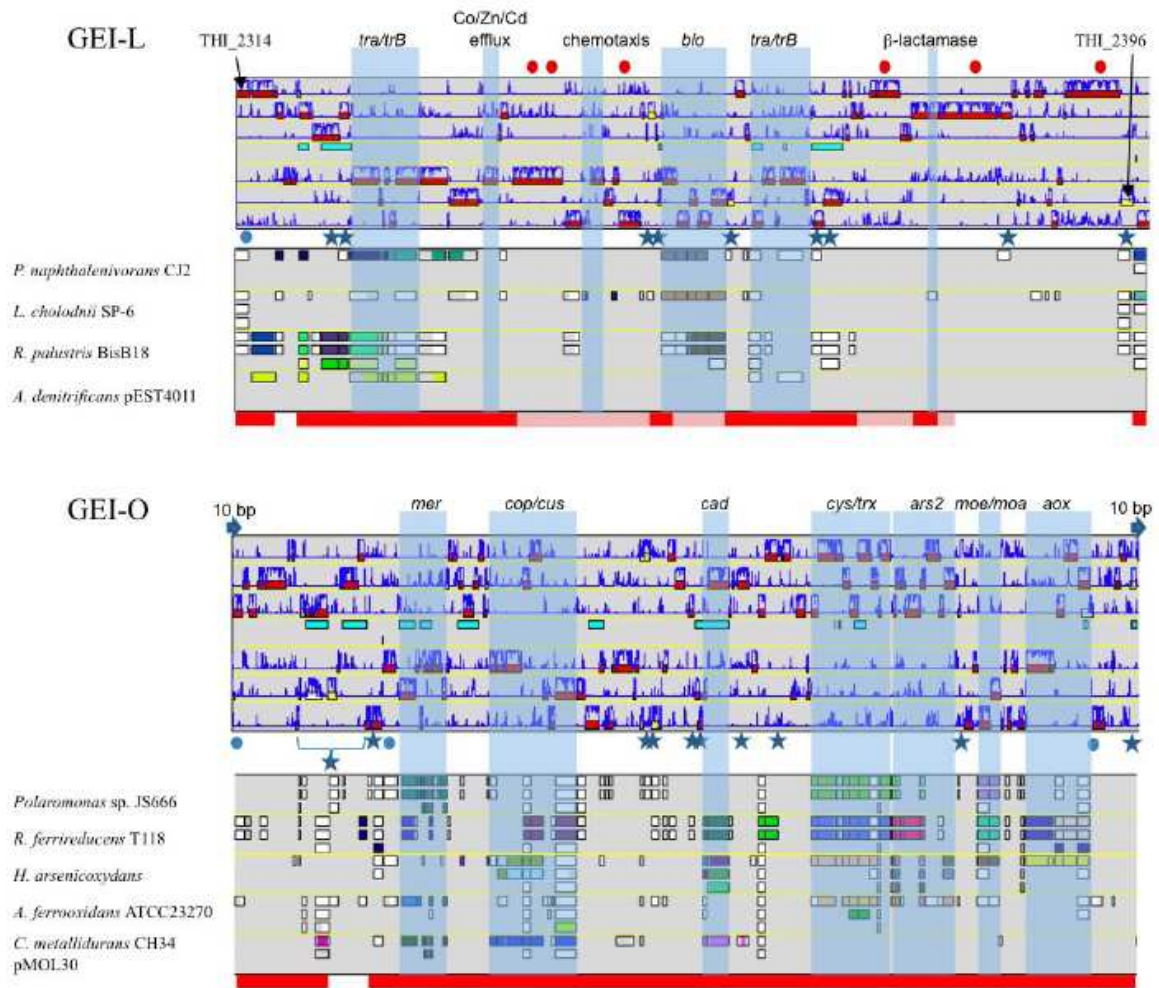


Figure 5

The supplementary data remain available upon request to the authors.

Chapitre II

Etude de la diversité et l'évolution de
gènes *aoxB* chez des bactéries
arsénite-oxydantes

I INTRODUCTION

Les gènes *aoxB* codant pour la sous-unité catalytique de l'enzyme sont largement distribués dans les environnements contaminés par l'arsenic (Inskeep *et al.*, 2007). Des amorces dégénérées ont été dessinées pour amplifier la première moitié du gène afin de mettre en évidence et analyser les gènes *aox* de diverses bactéries aérobies arsénite-oxydantes.

Les alignements protéiques de la première moitié de la protéine AoxB permettent de mettre en évidence une certaine variabilité des séquences. Néanmoins les zones catalytiques de l'enzyme ainsi que les zones de liaison à la sous-unité Rieske de chaque séquence présentent une forte homologie. Enfin ces analyses ont fait apparaître un rapport entre les études phylogéniques basées sur les séquences *aox* et sur les séquences du gène codant pour l'ARNr16S.

Ainsi le gène *aoxB* peut être considéré comme un marqueur moléculaire, afin d'investiguer le potentiel d'oxydation d'un environnement riche en arsenic.

Ce travail a été réalisé dans le cadre du GDR 2909 et en association avec le BRGM. Ma collaboration dans ce travail a porté sur les analyses moléculaires des souches cultivables, les tests de résistance aux métaux, les tests d'oxydation de l'As III ainsi que sur la réalisation et l'analyse des alignements de séquences protéiques. Enfin j'ai participé à l'écriture du manuscrit et à la réalisation des figures.

L'ensemble de ce travail a conduit à une publication dans *Applied Environmental Microbiology* présentée ci-dessous.

II MANUSCRIT

Diversity Surveys and Evolutionary Relationships of *aoxB* Genes in Aerobic Arsenite-Oxidizing Bacteria

Marianne Quéméneur, **Audrey Heinrich-Salmeron**, Daniel Muller, Didier Lièvremon, Michel Jauzein, Philippe N. Bertin, Francis Garrido, and Catherine Joulian

Applied and Environmental Microbiology, July 2008, p. 4567-4573

Diversity Surveys and Evolutionary Relationships of *aoxB* Genes in Aerobic Arsenite-Oxidizing Bacteria^{∇†}

Marianne Quéménéur,¹ Audrey Heinrich-Salmeron,² Daniel Muller,^{2,3} Didier Lièvreumont,² Michel Jauzein,⁴ Philippe N. Bertin,² Francis Garrido,¹ and Catherine Joulian^{1*}

BRGM, Ecotechnology Unit, Process and Environment Division, 3 Avenue Claude Guillemin, 45060 Orléans Cedex 02, France¹;
 Génétique Moléculaire, Génomique et Microbiologie, UMR 7156 CNRS and Université Louis Pasteur, 28 Rue Goethe,
 67000 Strasbourg, France²; Ecologie Microbienne, UMR CNRS 5557 USC INRA 1193 and Université de Lyon,
 69622 Villeurbanne Cedex, France³; and LIMOS, UMR 7137 CNRS and Université de Nancy, BP 239,
 54506 Vandœuvre-les-Nancy Cedex, France⁴

Received 18 December 2007/Accepted 12 May 2008

A new primer set was designed to specifically amplify ca. 1,100 bp of *aoxB* genes encoding the As(III) oxidase catalytic subunit from taxonomically diverse aerobic As(III)-oxidizing bacteria. Comparative analysis of AoxB protein sequences showed variable conservation levels and highlighted the conservation of essential amino acids and structural motifs. AoxB phylogeny of pure strains showed well-discriminated taxonomic groups and was similar to 16S rRNA phylogeny. Alphaproteobacteria-, Betaproteobacteria-, and Gammaproteobacteria-related sequences were retrieved from environmental surveys, demonstrating their prevalence in mesophilic As-contaminated soils. Our study underlines the usefulness of the *aoxB* gene as a functional marker of aerobic As(III) oxidizers.

Arsenic (As) exists mainly in two toxic soluble forms, arsenite, As(III), and arsenate, As(V), with the latter tending to associate with some oxyhydroxides and clay minerals. The bacterial oxidation of As(III) can thus contribute to a natural attenuation of As contamination by decreasing As bioavailability. These properties have recently been used to develop a bioprocess for removing As from a mining effluent by using the activity of As-metabolizing bacteria indigenous to the contaminated site (4). The feasibility of such a process depends on a good knowledge of the ability of the indigenous microflora to oxidize As(III) and requires reliable methods for detecting, identifying, and monitoring As(III) oxidizers in the environment.

More than 50 phylogenetically diverse As(III)-oxidizing strains distributed among 25 genera have been isolated from various environments so far. Bacterial aerobic As(III) oxidation is performed by a dedicated enzyme, the As(III) oxidase (1, 36, 40), which belongs to the dimethyl sulfoxide (DMSO) reductase of the molybdenum family (9). In *Alcaligenes faecalis*, it is an $\alpha_1\beta_1$ heterodimer comprising a large subunit incorporating a molybdenum center and a [3Fe-4S] cluster and a small subunit incorporating a Rieske-type [2Fe-2S] cluster (9). Genes encoding these subunits are cotranscribed as an operon and have been successively characterized in *Herminiimonas arsenicoxydans* (26), *Rhizobium* sp. strain NT-26 (36), and *Agrobacterium tumefaciens* (21). They have also been found in the genome of *Chloroflexus aurantiacus*, on a plasmid in *Ther-*

mus thermophilus, in two aerobic thermophilic As(III) oxidizers, and in the genome of strains for which the ability to oxidize As(III) has not been experimentally proven (27).

Due to the polyphyly of As(III)-oxidizing bacteria, the *aoxB* gene encoding the catalytic subunit of the enzyme seems to be a valuable molecular marker for investigating its ecology and the potential of As(III) oxidation in the environment. To this end, a recent study described primers targeting the first quarter of the *aoxB* gene to detect its presence and expression in the environment and suggested that the gene is widely distributed among the *Bacteria* and also is widespread in soil-water systems containing As (16).

In our present study, we designed new primers to extend the genetic information to the first half of the *aoxB* gene. We then explored the genetic diversity of this gene in order (i) to identify conserved structural and functional domains, (ii) to compare AoxB and 16S rRNA phylogenies to evaluate whether phylogenetic information about As(III)-oxidizing bacteria can be inferred from the *aoxB* gene, and (iii) to assess the composition of As(III)-oxidizing communities in environmental diversity surveys.

Primer evaluation and validation. Primers *aoxBM1-2F* (5'-CCACTTCTGCATCGTGGGNTGYGGNTA-3', positions 66 to 92 in the *H. arsenicoxydans aoxB* open reading frame) and *aoxBM3-2R* (5'-TGTCGTTGCCCCAGATGADNCCYTTYT C-3', positions 1150 to 1177) were designed on the most distantly located conserved regions with the CODEHOP program (32) from nine AoxB protein sequences from the *Bacteria* (*Alcaligenes faecalis* [GenBank accession no. AY297781], *Herminiimonas arsenicoxydans* [accession no. AF509588], "*Thiomonas arsenivorans*" [accession no. EU304260], *Agrobacterium tumefaciens* [accession no. DQ151549], *Rhizobium* sp. strain NT26 [accession no. AY345225], *Chloroflexus aurantiacus* [accession no. NZ_AAAH01000321],

* Corresponding author. Mailing address: BRGM, Ecotechnology Unit, Process and Environment Division, 3 Avenue Claude Guillemin, 45060 Orléans Cedex 02, France. Phone: 33(0)2 3864 3089. Fax: 33(0)2 3864 3680. E-mail: c.joulian@brgm.fr.

† Supplemental material for this article may be found at <http://aem.asm.org/>.

∇ Published ahead of print on 23 May 2008.

and *Thermus thermophilus* [accession no. NC_000854]), and the *Archaea* (*Aeropyrum pernix* [accession no. NC_000854] and *Sulfolobus tokodaii* [accession no. NC_003106]). The primers target the first part (amino acids 22 to 32, *H. arsenicoxydans* AoxB numbering) of the CX₂CX₃CX₇₀S motif (CHF CIVGCGYH) required for binding a [3Fe-4S] cluster and the consensus motif YEKGIWGN (amino acids 383 to 391).

Specific amplifications (35 cycles, annealing temperature of 52°C, 1 min 10 s elongation time) of ca. 1,100 bp were obtained for the following 21 chemoautotrophic and chemoheterotrophic As(III)-oxidizing strains belonging to *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, and *Chloroflexi*. *T. arsenivorans* b6^T (3), *Leptothrix* sp. strain S1.1 (4), *Variovorax* sp. strain 4.2 (4), *H. arsenicoxydans* ULPA1^T (43), *Acidovorax* sp. strain 75, *Acinetobacter* sp. strain 33, *Alcaligenes* sp. strain YI013H, *Alcaligenes* sp. strain T12RB, *Aminobacter* sp. strain 86, *Burkholderia* sp. strain YI019A, *Limnobacter* sp. strain 83, *Pseudomonas* sp. strains 1, 46, 72, 73, 89, and D2OHCJ, and *Ralstonia* sp. strain 22 were isolated in our laboratories from As-contaminated environments (see Table S1 in the supplemental material). *Thiomonas* sp. strain WJ68 and *Thiomonas* sp. strain NO115 were isolated from mining sites (7, 15). *Chloroflexus aurantiacus* DSM635 was isolated from hot springs (30). A less-specific amplification was obtained for the archaeon *Aeropyrum pernix* DSM11879. The detection of distinct genera thus indicated the ability to detect diverse taxa of this metabolic group with the newly designed primers. No amplification was obtained from the 10 non-As(III)-oxidizing bacteria used as negative controls, of which 6 harbor a molybdenum enzyme of the DMSO reductase family (i.e., *Rhodobacter sphaeroides*, *Desulfitobacterium hafniense*, *Bacillus selenitireducens*, *Geobacter metallireducens*, *Escherichia coli*, and *Halomonas denitrificans*).

Sensitivity tests showed that as little as 8.3 pg of template DNA of *H. arsenicoxydans* was needed for *aoxB* gene amplification. Knowing that the 3.4-Mb genome of *H. arsenicoxydans* contains only one *aoxB* gene (27), this would represent approximately 2.2×10^3 copies of the *aoxB* gene. The sensitivity was lower than that of *H. arsenicoxydans* for the phylogenetically distant *C. aurantiacus* strain, with the detection limit being 1.7×10^4 *aoxB* gene copies.

Protein and nucleotide sequence analyses. Analyses were conducted on 29 sequences of As(III)-oxidizing isolates, excluding recently published short sequences (16) and sequences from whole-genome data for which there is no experimental evidence that the corresponding strains oxidize As(III).

Examination of nucleotide sequence alignments using DnaSP version 4 software (33) revealed a high proportion of variable sites (82%), of which 74% were parsimony informative. Conservation degrees ranged from 48.6% to 92.3%. The second codon positions were clearly better conserved. The third codon positions were the most variable (99%), and saturation analysis using DAMBE program version 4.0.50 (44) showed that saturation occurred at this position. The number of substitution sites was higher for phylogenetically distant strains than for closely related ones. Nucleotide substitution causing amino acid substitution was more frequent than synonymous substitution, which accounted for only 23.7% of all substitutions.

Comparative analysis of the deduced AoxB protein sequences showed variable conservation levels (38.4 to 96.8% identity), comparable to data obtained for other functional genes, e.g., those involved in sulfur oxidation, nitrous oxide reduction, and sulfite reduction (28, 37, 42). The 15 partial AoxB sequences presented in Fig. 1 showed 22% identity on 372 amino acids. Three of the four domains of the AoxB subunit described for *A. faecalis* (9) were retrieved in the aligned region (the fourth domain extends outside the sequences presented here). Domain I extends from residue 21 to residue 119 (according to *A. faecalis* protein sequence numbering) and binds the Rieske subunit and the [3Fe-4S] cluster coordinated by the motif C₂₁-X₂-C₂₄-X₃-C₂₈-X₇₀-S₉₉. Except for amino acids present in the X₇₀ region, a high conservation was observed, particularly between residues 21 and 31 and between residues 86 and 103. A highly conserved region between residues 166 and 261 encompassed domains II (residues 120 to 195) and III (residues 196 to 392) and contains several structural motifs, such as alpha-helices and residues like H195 and E203 implicated in the substrate-binding site of the enzyme. Interestingly, the HNRPAYNSE motif is exactly conserved in all bacterial As(III) oxidases. The conserved S99 and A199 residues are important in the demarcation of the catalytic subunit of As(III) oxidase with the other members of the DMSO reductase family of molybdenum enzymes (9). A lower amino acid conservation was observed in the other parts of the protein, although the presence of multiple conservative substitutions suggests that the physicochemical properties of these amino acids play a role in the structure and/or the functioning of the As(III) oxidase enzyme. Finally, regions that contain no structural element showed a low amino acid conservation. For example, the *C. aurantiacus* sequence shows six additional amino acids between residues 330 and 331. In addition, the three *Alphaproteobacteria*, namely *A. tumefaciens*, *Rhizobium* sp. strain NT26, and an *Aminobacter* sp., harbor 24 additional amino acids in this region. This further supports the low amino acid conservation observed throughout the As(III) oxidase protein sequence in these microorganisms compared to that of the *Betaproteobacteria* protein sequences. Taken together, our observations suggest that the As(III) oxidases identified here show moderate amino acid conservation but share several structural and functional domains similar to those found in *A. faecalis* AoxB (Fig. 1), supporting the important role of these domains in As(III) oxidase activity.

Deduced protein sequences of closely related species were generally better conserved than nucleotide sequences. For example, *Alcaligenes* species exhibited sequence identity values ranging from 80.2 to 96.8% for proteins and from 76.6 to 92.2% for nucleotides. AoxB sequences within the same subclass of *Proteobacteria* showed generally higher identity values (e.g., between 72.6% and 96% for *Alphaproteobacteria*). The highest divergences in protein sequences were observed between distantly related bacterial lineages, i.e., *Proteobacteria* versus the nonproteobacterial *Thermus* and *Chloroflexus* strains (identities between 38.4% and 50.7%), and *Alphaproteobacteria* (except *Hydrogenophaga* sp. strain CL3; see below) versus *Betaproteobacteria* and *Gammaproteobacteria* (identities between 46.7 and 53.5%). In contrast, nucleotide sequences were better conserved than protein sequences for distantly related species. Similar events have been reported for other

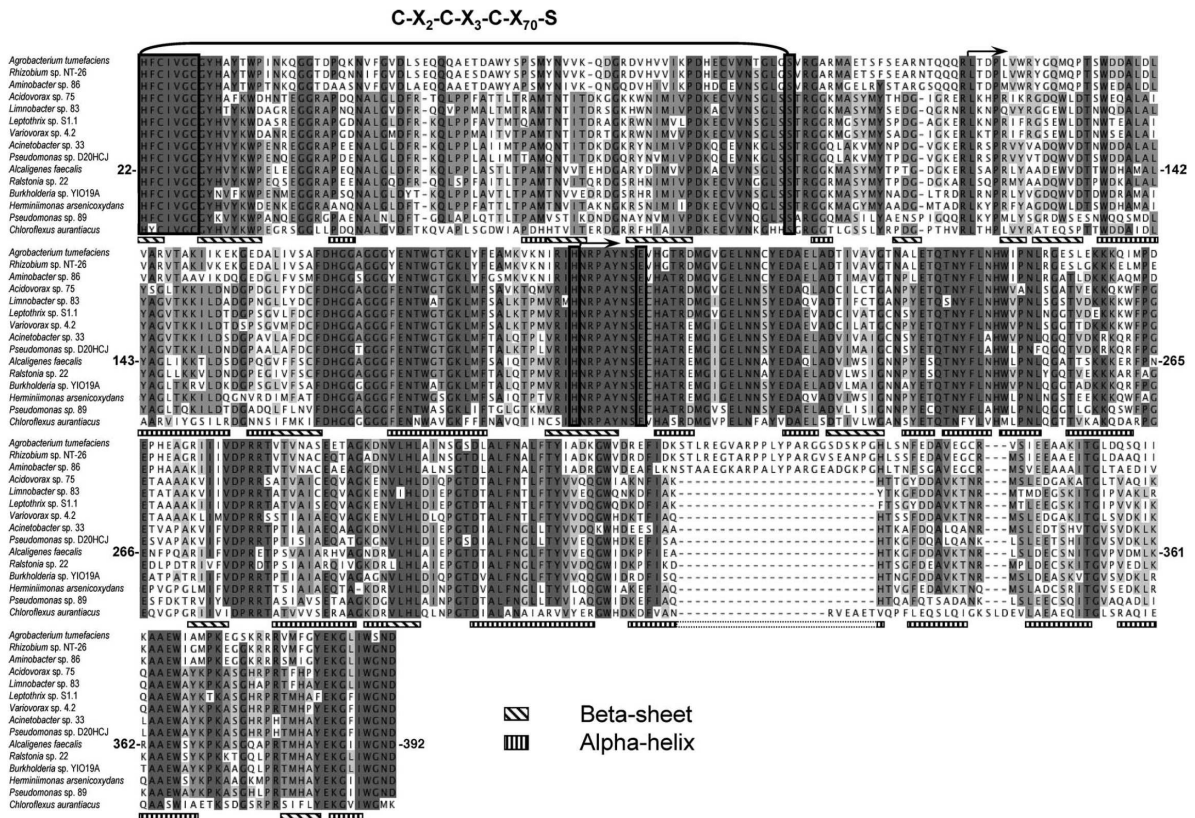


FIG. 1. Protein sequence alignment and putative secondary structure of As(III) oxidases. The *Alcaligenes faecalis* sequence and its secondary structure were retrieved from the PDB database (<http://www.rcsb.org/pdb/welcome.do>). Other As(III) oxidase protein sequences were deduced from nucleotide *aoxB* sequences. Sequence alignment was carried out by CLUSTAL W (39). Amino acid residue numbers are deduced according to the *A. faecalis* protein sequence, which starts at a histidine residue at position 22. Residues identical to those of the *A. faecalis* sequence are shown in gray-scale boxes. Residues known to play a role in As(III) oxidase activity are framed in black. Black arrows at positions 120 and 196 indicate the beginning of domains 2 and 3, respectively. The beta-sheet and alpha-helix are shown according to the three-dimensional structure of *A. faecalis* protein (9).

functional genes, e.g., those involved in denitrification and sulfur oxidation (5, 28). In the case of *aoxB* genes, these observations can be explained by saturation at the third codon position for comparison of distant strains. The observed divergences suggest that the *aoxB* gene may have a long evolutionary history and support the hypothesis of an early common ancestor (24). Rhine et al. (31) recently reported that autotrophic *aoxB* genes formed a phylogenetic group distinct from heterotrophic *aoxB* genes, with the exception of *A. tumefaciens aoxB* genes, and suggested that the two groups evolved separately from a common ancestor. However, our study demonstrated that the *aoxB* gene of the autotrophic *T. arsenivorans* is phylogenetically affiliated with heterotrophic *aoxB* genes and not with genes of autotrophic As(III) oxidizers. Further data on *aoxB* gene sequences from pure heterotrophic and autotrophic strains will thus be needed to elucidate the evolution of this gene.

Protein sequences are usually used for reconstructing the phylogeny of protein-encoding genes (6, 37). In addition, protein sequences appeared less noisy and more resolute than

nucleotide sequences for large-scale (phylum) analysis. Further reconstruction of As(III) oxidase phylogeny has thus been based on protein sequences.

Comparison of AoxB and 16S rRNA gene phylogenies. AoxB and 16S rRNA neighbor-joining (34) trees were constructed from unambiguous residues on distances estimated by the Kimura (22) and Jukes and Cantor (20) methods, respectively. The AoxB tree reconstructed the major taxonomic levels and was similar to the 16S rRNA tree (Fig. 2). Indeed, the *C. aurantiacus* AoxB sequence is clearly separated from the AoxB proteobacterial sequences, with a strong bootstrap support (100%). The *Alphaproteobacteria* and *Betaproteobacteria*/*Gamma*-*proteobacteria* AoxB sequences form distinct phylogenetic branches, supported by 100% bootstrap values. In addition, sequences from species of the same genus, such as *Thiomonas* or *Alcaligenes*, cluster together.

However, a detailed comparison of the two trees revealed some discrepancies. For example, the pseudomonads are divided into two groups in the AoxB-based tree. Group 1 forms a distinct phylogenetic branch supported by a high bootstrap

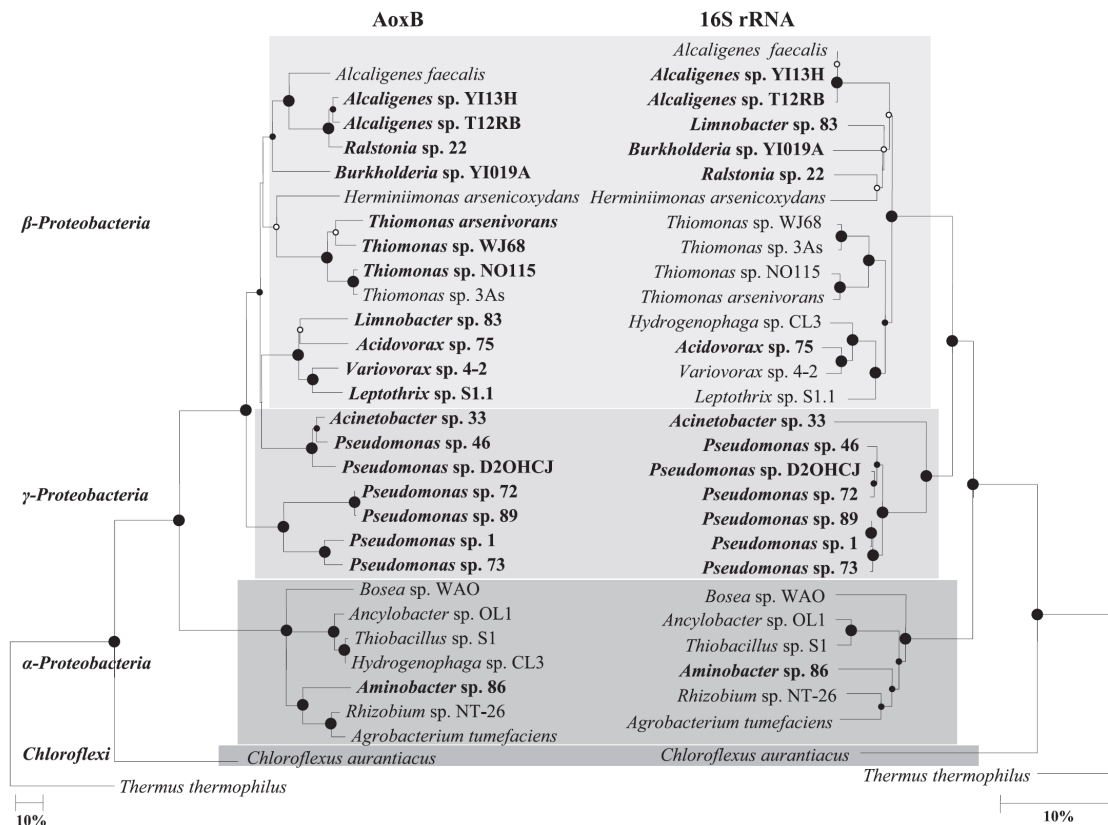


FIG. 2. Neighbor-joining phylogenetic trees showing the relationships between partial *AoxB* protein (351-residue) and partial 16S rRNA (1,139-nucleotide) sequences of bacterial As(III) oxidizers. Bacteria whose sequences were determined in this study are in bold. The *Thermus thermophilus* *AoxB* sequence was used as the outgroup. Circles at the branch nodes represent bootstrap percentages for 100 replicates (10): large filled circles, 95 to 100%; small filled circles, 75 to 95%; open circles, 50 to 75%. Scale bars correspond to 10 mutations per 100 residues. Phylogenetic programs were implemented in the TREECON package (41). Parsimony analysis (11) gave similar topologies.

value (100%), while group 2 forms a separate branch of sequences more closely related to those from the *Betaproteobacteria*. The major inconsistency is the position of the betaproteobacterium *Hydrogenophaga* sp. strain CL3 within *Alphaproteobacteria*, which suggests the possibility of lateral gene transfer. However, two *Hydrogenophaga* spp. were properly positioned among the *Betaproteobacteria* on an *AoxB* tree constructed from shorter sequences (16). However, strain CL3 is also branching within *Alphaproteobacteria* when a tree is built from short sequences (data not shown). Collection of *aoxB* sequences from other *Hydrogenophaga* isolates would help to determine whether or not this conspicuous event is limited to the *Hydrogenophaga* species. Interestingly, the *aoxB* gene is plasmid carried in *Thermus thermophilus* HB8 and is carried by a genomic island in *A. faecalis* (38) and *H. arsenicoxydans* (27), indicating that lateral transfer of the gene can occur. These events have also been demonstrated for other functional genes (12, 23), particularly for *arsC*, a gene involved in the As cycle and in encoding the cytoplasmic As(V) reductase (18). Consequently, *AoxB* sequences must be used with caution to infer phylogenetic information about environmental As(III)-oxidiz-

ing bacteria. A greater collection of *aoxB* genes will undoubtedly increase confidence in the *AoxB* phylogeny.

***AoxB* diversity in environmental samples.** Diversity surveys of *aoxB* genes were conducted for evaluating the pertinence of the newly developed assay to assess the composition of As(III) oxidizers in the environment. Gene libraries were constructed (TA cloning kit; Invitrogen, Carlsbad, CA) on (i) an As-contaminated industrial soil (T12R) obtained from a former cokery (19), (ii) an enrichment of this soil under As(III)-oxidizing chemoheterotrophic conditions, and (iii) an enrichment of this soil under As(III)-oxidizing chemoautotrophic conditions. Enrichments were made from a 100-fold-diluted (wt/vol) T12R soil suspension in CasO1 selective medium (2) containing 200 mg liter⁻¹ of As(III) and supplemented with 5 mM Na lactate, 5 mM acetate, and 0.2% yeast extract under heterotrophic conditions; As(III) oxidation was measured as described by Battaglia-Brunet et al. (2). DNA was extracted from 1 g of soil and 50 ml of enrichment cultures using the MoBio UltraClean soil extraction kit (MoBio, Solana Beach, CA).

Of the 133 clones screened by *RsaI*/*HaeIII* restriction analyses, 62 inserts were sequenced and finally grouped into 28

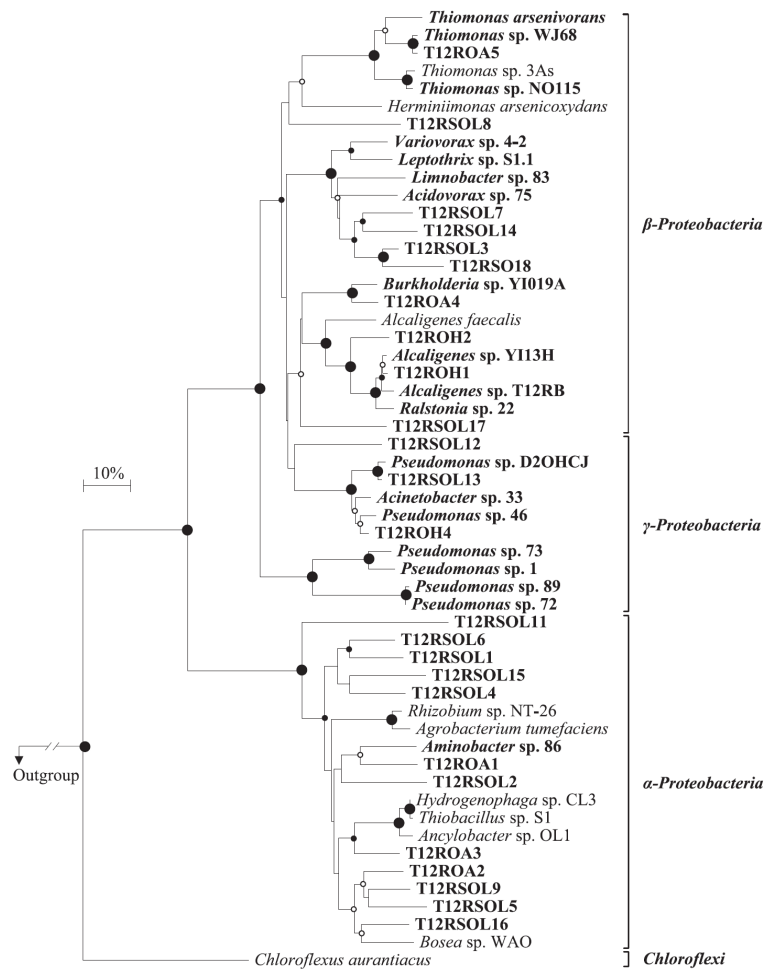


FIG. 3. Neighbor-joining phylogenetic tree of bacterial *AoxB* sequences retrieved from environmental surveys. T12RSOL, As-contaminated soil T12R; T12ROA, an As(III)-oxidizing chemoautotrophic enrichment of soil T12R; T12ROH, an As(III)-oxidizing chemoheterotrophic enrichment of soil T12R. Bacteria whose sequences were determined in this study are in bold. The *Thermus thermophilus* *AoxB* sequence was used as the outgroup. Circles at the branch nodes represent bootstrap percentages for 100 replicates (10): large filled circles, 95 to 100%; small filled circles, 75 to 95%; open circles, 50 to 75%. The scale bar corresponds to 10 mutations per 100 residues. All phylogenetic programs were implemented in the TREECON package (41).

operational taxonomic units (OTU) based on a 90% sequence identity. Only *aoxB* gene-related sequences were retrieved, and such exclusive specificity for the target gene has the advantage of reducing the stringency of annealing conditions, thus favoring greater sensitivity. All the environmental *AoxB* sequences contained the conserved residues predicted by Ellis et al. (9) as being essential for As(III) oxidase activity in *A. faecalis*, suggesting that these sequences produced functional enzymes. The environmental *AoxB* sequences were affiliated with sequences from strains belonging to the *Alphaproteobacteria*, *Betaproteobacteria*, and *Gammaproteobacteria*, with high bootstrap values (Fig. 3).

A large diversity (18 OTU) was detected from the soil T12R. The library was dominated (58%) by *Alphaproteobacteria*-related sequences, among which 19% were related to a *Bosea* sp. Five OTU formed separate branches of known *Alphapro-*

teobacteria sequences. *Betaproteobacteria*-related OTU (17%) were mainly affiliated with an *Acidovorax* sp. *Gammaproteobacteria*-related OTU (25%) were mainly represented by group 2 pseudomonads. Interestingly, incubation of this soil under chemoheterotrophic As(III)-oxidizing conditions led to a decrease in diversity with the selection of betaproteobacterial *Alcaligenes*-related sequences (88%). Only 1 OTU showed a close relationship to the *Gammaproteobacteria* group 2 pseudomonads. Remarkably, no alphaproteobacterium was detected, although *Alphaproteobacteria* were dominant in the soil. When this soil was incubated under chemoautotrophic As(III)-oxidizing conditions, OTU related to sequences of the autotrophic As(III)-oxidizing *Thiomonas* (55%) and *Burkholderia* (17%), genera of the *Betaproteobacteria*, dominated the library. The *Alphaproteobacteria*-related OTU represented 32% of the library (mainly an *Ancylobacter* sp.). To date, the

majority of known As(III)-oxidizing strains belonging to *Alphaproteobacteria* are autotrophs, suggesting a predominance of autotrophic metabolism among them (31). In contrast to the As(III)-oxidizing chemoheterotrophic enrichment, neither *Gammaproteobacteria*- nor *Alcaligenes* (*Betaproteobacteria*)-affiliated sequences were found. Our results underline the usefulness of functional molecular markers in diversity surveys for directly detecting the functionality of an environment without the need of cultivation.

The exclusive detection of environmental AoxB sequences belonging to the phylum *Proteobacteria* is probably related to the fact that we studied a mesophilic environment, which is likely to contain a large majority of As(III) oxidizers from this phylum. To date, all but two of the As(III)-oxidizing bacteria isolated from mesophilic sites belong to the phylum *Proteobacteria*, and a recent study (16) has shown that 98% of *aoxB*-like sequences retrieved from mesophilic As-contaminated soils and sediments belong to the *Proteobacteria*. The only reports on the ability of mesophilic gram-positive bacteria to perform As(III) oxidation concerned *Microbacterium lacticum* (25) and *Bacillus arsenoxydans* (14), with the general physiology of the latter organism being similar to that of the betaproteobacterium *A. faecalis* (29). As(III)-oxidizing bacteria belonging to other phyla, such as *Deinococcus-Thermus* (13), *Chloroflexi* (24), and *Aquificae* (8), have been isolated only from geothermic sites. It is also noteworthy that As(III)-oxidizing *Proteobacteria* are able to colonize geothermic sites (17, 35).

Our study has demonstrated that the *aoxB* gene has the major features of a molecular marker. (i) The *aoxB* gene has been found in all the aerobic As(III)-oxidizing bacteria tested so far. (ii) Conserved regions across this gene have enabled the design of valuable primers. (iii) The studied region is sufficiently large to obtain genetic variation, allowing the discrimination of phylogenetic groups. These data are encouraging for the further use of the *aoxB* gene as a functional marker specific to aerobic As(III) oxidizers in environmental diversity surveys. Future work will focus on evaluating the link between As speciation and concentration, and *aoxB* gene diversity and abundance, in the environment.

Nucleotide sequence accession numbers. The *aoxB* sequences have been deposited in the GenBank database under accession no. EU304260 to EU304278, EU304293 to EU304310, and EU304313 to EU304321, and the 16S rRNA gene sequences have been deposited under accession no. EU304279 to EU304292.

This is BRGM contribution no. 05252.

We thank the Agence de l'Environnement et de la Maîtrise de l'Energie (ADEME) and the French Ministry of Education and Research for financially supporting this research. The work was partly done within the framework of the Groupement de Recherche-Métabolisme de l'Arsenic chez les Prokaryotes (GDR2909-CNRS) and of the European Union FP6 Integrated Project AquaTerra (project no. GOCE 505428).

We are grateful to K. H. Hallberg for providing *Thiomonas* sp. strain WJ68 and *Thiomonas* sp. strain NO115. We also thank Marie-Claire Lett and Laurence Casalat for useful discussions.

REFERENCES

- Anderson, G. L., J. Williams, and R. Hille. 1992. The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum-containing hydroxylase. *J. Biol. Chem.* **267**:23674–23682.
- Battaglia-Brunet, F., M.-C. Dictor, F. Garrido, C. Crouzet, D. Morin, K. Dekeyser, M. Clarens, and P. Baranger. 2002. An arsenic(III)-oxidizing bacterial population: selection, characterization, and performance in reactors. *J. Appl. Microbiol.* **93**:656–667.
- Battaglia-Brunet, F., C. Joulain, F. Garrido, M.-C. Dictor, D. Morin, K. Coupland, D. Barrie Johnson, K. B. Hallberg, and P. Baranger. 2006. Oxidation of arsenite by *Thiomonas* strains and characterization of *Thiomonas arsenivorans* sp. nov. *Antonie van Leeuwenhoek* **89**:99–108.
- Battaglia-Brunet, F., Y. Itard, F. Garrido, F. Delorme, C. Crouzet, C. Greffier, and C. Joulain. 2006. A simple biogeochemical process removing arsenic from a mine drainage water. *Geomicrobiol. J.* **23**:1–11.
- Braker, G., J. Zhou, L. Wu, A. H. Devol, and J. M. Tiedje. 2000. Nitrite reductase genes (*nirK* and *nirS*) as functional markers to investigate diversity of denitrifying bacteria in Pacific northwest marine sediment communities. *Appl. Environ. Microbiol.* **66**:2096–2104.
- Case, R. J., Y. Boucher, I. Dahllöf, C. Holmström, W. F. Doolittle, and S. Kjelleberg. 2007. Use of 16S rRNA and *ppoB* genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* **73**:278–288.
- Coupland, K., F. Battaglia-Brunet, K. B. Hallberg, M.-C. Dictor, F. Garrido, and D. B. Johnson. 2004. Oxidation of iron, sulfur and arsenic in mine waters and mine wastes: an important role for novel *Thiomonas* spp., p. 639–646. *In* M. Tsezos, A. Hatzikioseyan, and E. Remoudaki (ed.), *Biohydrometallurgy: a sustainable technology in evolution*. National Technical University of Athens, Zografou, Greece.
- Donahoe-Christiansen, J., S. D'Imperio, C. R. Jackson, W. P. Inskeep, and T. R. McDermott. 2004. Arsenite-oxidizing *Hydrogenobaculum* strain isolated from an acid-sulfate-chloride geothermal spring in Yellowstone National Park. *Appl. Environ. Microbiol.* **70**:1865–1868.
- Ellis, P. J., T. Conrads, R. Hille, and P. Kuhn. 2001. Crystal structure of the 100 kDa arsenite oxidase from *Alcaligenes faecalis* in two crystal forms at 1.64 Å and 2.03 Å. *Structure* **9**:125–132.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
- Felsenstein, J. 1989. PHYLIP—phylogeny inference package (version 3.2). *Cladistics* **5**:164–166.
- Friedrich, M. W. 2002. Phylogenetic analysis reveals multiple lateral transfers of adenosine-5'-phosphosulfate reductase genes among sulfate-reducing microorganisms. *J. Bacteriol.* **184**:278–289.
- Gühring, T. M., G. K. Druschel, R. B. McCleskey, R. J. Hamers, and J. F. Banfield. 2001. Rapid arsenite oxidation by *Thermus aquaticus* and *Thermus thermophilus*: field and laboratory investigations. *Environ. Sci. Technol.* **35**:3857–3862.
- Green, H. H. 1918. Isolation and description of a bacterium causing oxidation of arsenite to arsenate in cattle-dipping baths. *Rep. Dir. Vet. S. Afr.* **6**:593–599.
- Hallberg, K. B., and D. B. Johnson. 2003. Novel acidophiles isolated from moderately acidic mine drainage waters. *Hydrometallurgy* **71**:139–148.
- Inskeep, W. P., R. E. Macur, N. Hamamura, T. P. Warelow, S. A. Ward, and J. M. Santini. 2007. Detection, diversity and expression of aerobic bacterial arsenite oxidase genes. *Environ. Microbiol.* **9**:934–943.
- Jackson, C. R., H. W. Langner, J. Donahoe-Christiansen, W. P. Inskeep, and T. R. McDermott. 2001. Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. *Environ. Microbiol.* **3**:532–542.
- Jackson, C. R., and S. L. Dugas. 2003. Phylogenetic analysis of bacterial and archaeal *arsC* gene sequences suggests an ancient, common origin for arsenate reductase. *BMC Evol. Biol.* **23**:3–18.
- Joubert, A. V. P., L. Lucas, F. Garrido, C. Joulain, and M. Jauzein. 2007. Effect of temperature, gas phase composition, pH and microbial activity on As, Zn, Pb and Cd mobility in selected soils in the Ebro and Meuse Basins in the context of global change. *Environ. Pollut.* **148**:749–758.
- Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules, p. 21–132. *In* H. N. Munro (ed.), *Mammalian protein metabolism*. Academic Press, New York, NY.
- Kashyap, D. R., L. M. Botero, W. L. Franck, D. J. Hassett, and T. R. McDermott. 2006. Complex regulation of arsenite oxidation in *Agrobacterium tumefaciens*. *J. Bacteriol.* **188**:1081–1088.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequence. *J. Mol. Biol.* **16**:111–120.
- Klein, M., M. Friedrich, A. J. Roger, P. Hugenholz, S. Fishbain, H. Abicht, L. L. Blackall, D. A. Stahl, and M. Wagner. 2001. Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes. *J. Bacteriol.* **183**:6028–6035.
- Lebrun, E., M. Brugna, F. Baymann, D. Muller, D. Lièvreumont, M.-C. Lett, and W. Nitschke. 2003. Arsenite oxidase, an ancient bioenergetic enzyme. *Mol. Biol. Evol.* **20**:686–693.
- Mokashi, S. A., and K. M. Paknikar. 2002. Arsenic (III) oxidizing *Microbacterium lacticum* and its use in the treatment of arsenic contaminated groundwater. *Lett. Appl. Microbiol.* **34**:258–262.
- Muller, D., D. Lièvreumont, D. D. Simeonova, J.-C. Hubert, and M.-C. Lett. 2003. Arsenite oxidase *aox* genes from a metal-resistant β -proteobacterium. *J. Bacteriol.* **185**:135–141.
- Muller, D., C. Médigue, S. Koehler, V. Barbe, M. Barakat, E. Talla, V.

- Bonnefoy, E. Krin, F. Arsène-Plotze, C. Carapito, M. Chandler, B. Cournoyer, S. Cruveiller, C. Dossat, S. Duval, M. Heymann, E. Leize, A. Lieutaud, D. Lièvreumont, Y. Makita, S. Mangenot, W. Nitschke, P. Ortet, N. Perdrial, B. Schoepp, P. Siguier, D. D. Simeonova, Z. Rouy, B. Segurens, E. Turlin, D. Vallenet, A. Van Dorsselaer, S. Weiss, J. Weissenbach, M.-C. Lett, A. Danchin, and P. N. Bertin. 2007. A tale of two oxidation states: bacterial colonization of arsenic-rich environments. *PLoS Genet.* **3**:e53.
28. Petri, R., L. Podgorsek, and J. F. Imhoff. 2001. Phylogeny and distribution of the *soxB* gene among thiosulfate-oxidizing bacteria. *FEMS Microbiol. Lett.* **197**:171–178.
29. Phillips, S. E., and M. L. Taylor. 1976. Oxidation of arsenite to arsenate by *Alcaligenes faecalis*. *Appl. Environ. Microbiol.* **32**:392–399.
30. Pierson, B. K., and R. W. Castenholz. 1974. A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Arch. Microbiol.* **100**:5–24.
31. Rhine, E. D., S. M. Ni Chadhain, G. J. Zylstra, and L. Y. Young. 2007. The arsenite oxidase genes (*aroAB*) in novel chemoautotrophic arsenite oxidizers. *Biochem. Biophys. Res. Commun.* **354**:662–667.
32. Rose, T. M., E. R. Schultz, J. G. Henikoff, S. Pietrokovski, C. M. McCallum, and S. Henikoff. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res.* **26**:1628–1635.
33. Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**:2496–2497.
34. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
35. Salmassi, T. M., J. J. Walker, D. K. Newman, J. R. Leadbetter, N. R. Pace, and J. G. Hering. 2006. Community and cultivation analysis of arsenite oxidizing biofilms at Hot Creek. *Environ. Microbiol.* **8**:50–59.
36. Santini, J. M., and R. N. vanden Hoven. 2004. Molybdenum-containing arsenite oxidase of the chemolithoautotrophic arsenite oxidizer NT-26. *J. Bacteriol.* **186**:1614–1619.
37. Scala, D. J., and L. J. Kerkhof. 1999. Diversity of nitrous oxide reductase (*nosZ*) genes in continental shelf sediments. *Appl. Environ. Microbiol.* **65**:1681–1687.
38. Silver, S., and L. T. Phung. 2005. Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. *Appl. Environ. Microbiol.* **71**:599–608.
39. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
40. vanden Hoven, R. N., and J. M. Santini. 2004. Arsenite oxidation by the heterotroph *Hydrogenophaga* sp. str. NT-14: the arsenite oxidase and its physiological electron acceptor. *Biochim. Biophys. Acta* **1656**:148–155.
41. Van de Peer, Y., and R. De Wachter. 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput. Appl. Biosci.* **10**:569–570.
42. Wagner, M., A. J. Roger, J. L. Flax, G. A. Brusseau, and D. A. Stahl. 1998. Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *J. Bacteriol.* **180**:2975–2982.
43. Weeger, W., D. Lièvreumont, M. Perret, F. Lagarde, J.-C. Hubert, M. Leroy, and M.-C. Lett. 1999. Oxidation of arsenite to arsenate by a bacterium isolated from an aquatic environment. *Biometals* **12**:141–149.
44. Xia, X., and Z. Xie. 2001. DAMBE: software package for data analysis in molecular biology and evolution. *J. Hered.* **92**:371–373.

COMPLEMENTARY INFORMATION ON AS(III)-OXIDIZING ISOLATES

Fourteen phylogenetically diverse As(III)-oxidizing strains have been isolated from various As-polluted environments on selective media (1, 3), amended (chemoheterotrophs) or not (chemoautotrophs) with carbon sources. The sources of inoculum were (i) soils of two industrial sites in the Meuse Basin (Belgium), (ii) As-polluted surface waters collected along the Isle River Basin (Limousin, France) impacted by gold mining activities, and (iii) soils affected by mining contamination (Salsigne gold mine, Aude, France).

The capacity of isolated colonies to oxidize As(III) was confirmed as described by Battaglia-Brunet *et al.* (1) or by Weeger *et al.* (3). Isolates D2OHCJ and T12RB from Meuse Basin soils and YI013H and YI019A from Isle River waters were able to totally oxidize 200 mg l⁻¹ of As(III) (measured after 7 days); among them, only isolate YI019A was a chemoautotroph. Bacterial strains isolated from the Salsigne mining site were able to oxidize 100 mg l⁻¹ of As(III) (measured after 2 days) and exhibited chemoheterotrophic growth.

Table 1 presents the phylogenetic affiliation of each isolate determined by comparison of 1214 aligned and unambiguous nucleotides (RDP alignment refined manually) of 16S rRNA gene sequence with sequences of known bacteria, using distance calculations implemented in the BioEdit program (2). The fourteen As(III)-oxidizing isolates were related to the *Alpha-*, *Beta-*, and *Gamma-proteobacteria*. The majority of isolated strains belonged to genera comprising species already known to oxidize As(III), e.g. *Pseudomonas* spp. and *Alcaligenes* spp. To our knowledge, it is the first time that members of the genera *Aminobacter*, *Burkholderia*, and *Limnobacter* are proven to oxidize As(III). They were all used as positive controls for testing the primers developed to specifically target *aoxB* genes.

Table 1. Source and phylogenetic affiliation of As(III)-oxidizing isolates

Isolates	Sources	Most closed related species [Accession number]	16S rRNA gene identities (%)	Classes
1	Soils of Salsigne gold mine (France)	<i>Pseudomonas stutzeri</i> [AJ312175]	99.6	<i>Gammaproteobacteria</i>
22		<i>Ralstonia solanacearum</i> [EF016365]	97.8	<i>Betaproteobacteria</i>
33		<i>Acinetobacter calcoaceticus</i> [AJ888984]	99.3	<i>Gammaproteobacteria</i>
46		<i>Pseudomonas reactans</i> [AY747594]	99.9	<i>Gammaproteobacteria</i>
72		<i>Pseudomonas putida</i> [AY958233]	99.0	<i>Gammaproteobacteria</i>
73		<i>Pseudomonas stutzeri</i> [DQ537505]	99.4	<i>Gammaproteobacteria</i>
75		<i>Acidovorax facilis</i> [AF078765]	99.0	<i>Betaproteobacteria</i>
83		<i>Limnobacter thiooxidans</i> [AJ289885]	99.6	<i>Betaproteobacteria</i>
86		<i>Aminobacter lissarensis</i> [AF107722]	99.9	<i>Alphaproteobacteria</i>
89		<i>Pseudomonas stutzeri</i> [AJ312175]	99.7	<i>Gammaproteobacteria</i>
D2OHCJ		Soils of Meuse Basin (Belgium)	<i>Pseudomonas putida</i> [AY918067]	99.8
T12RB	<i>Alcaligenes faecalis</i> [AJ509012]		99.7	<i>Betaproteobacteria</i>
YI013H	Waters of Isle River Basin (France)	<i>Alcaligenes faecalis</i> [AJ509012]	99.7	<i>Betaproteobacteria</i>
YIO19A		<i>Burkholderia phenazinium</i> [AY154375]	99.7	<i>Betaproteobacteria</i>

REFERENCES

1. Battaglia-Brunet, F., M.-C. Dictor, F. Garrido, C. Crouzet, D. Morin, K. Dekeyser, M. Clarens, and P. Baranger. 2002. An As(III)-oxidizing bacterial population: selection, characterization, and performance in reactors. *J. Appl. Microbiol.* **93**:656–667.
2. Hall, T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**:95-98.
3. Weeger, W., D. Lievreumont, M. Perret, F. Lagarde, J. C. Hubert, M. Leroy, and M.-C. Lett. 1999. Oxidation of arsenite to arsenate by a bacterium isolated from an aquatic environment. *Biometals* **12**:141-149.

Chapitre III

Etude de la diversité des gènes *AOXB* dans
la communauté bactérienne de l'ancienne
mine de Sainte-Marie-aux-Mines

I INTRODUCTION

Le chapitre précédent a permis de montrer que les gènes codant les arsénite-oxydases peuvent être considérés comme un marqueur moléculaire de souches aérobies arsénite-oxydantes (Quéméneur *et al.*, 2008). Il était alors intéressant d'élargir les connaissances en analysant les gènes *aoxB* d'un environnement arsénié en se focalisant sur la communauté. Deux approches avaient été entreprises pour étudier des gènes *aox* issus d'une communauté complexe d'environnements contaminés et présentant des conditions plus ou moins acides (Inskeep *et al.*, 2007) (Hamamura *et al.*, 2009). Cependant ces approches avaient été menées sur des sites riches en arsenic et présentant donc une diversité microbienne faible. Il était intéressant d'utiliser des approches similaires mais sur des communautés présentant une plus grande diversité.

Au laboratoire, un projet a été initié en collaboration avec le Laboratoire des Interactions Ecotoxicologie Biodiversité Ecosystèmes de Metz et le Laboratoire de Spectrométrie de Masse Bio-organique de Strasbourg. L'objectif du projet était d'identifier les enzymes synthétisées par une communauté microbienne et jouant un rôle dans le métabolisme bactérien lié à l'arsenic en se basant sur une approche intégrée impliquant à la fois l'utilisation de méthodes classiques de microbiologie, de biologie moléculaire, de biochimie et de chimie analytique. Le projet s'est articulé autour de trois points : la caractérisation taxonomique et fonctionnelle de la communauté microbienne du site par l'analyse des protéines exprimées par la communauté par une approche de métaprotéomique et l'analyse des gènes impliqués dans le métabolisme de l'arsenic.

L'ensemble de ces approches a été mené pour la première fois sur l'ancienne mine d'argent et d'arsenic de Gabe-Gottes près de Sainte-Marie-aux-Mines en Alsace. Il s'agit d'un site présentant une contamination en arsenic relativement faible (320 mg/kg) et qui présente aussi une diversité bactérienne importante.

Ma participation à ce projet a consisté en l'exploration de la diversité des gènes *aox* de la communauté de ce site. De façon plus précise, j'ai participé au choix du site, à l'échantillonnage, à l'analyse des gènes *aox* de la communauté, à l'interprétation des données protéomiques et des données de séquences *aox* (alignement protéique et arbre phylogénétique), à l'écriture du manuscrit et à la réalisation de figures.

Par ces approches intégrées, nous avons mis en évidence une grande diversité des gènes *aox*. En particulier, de nouvelles séquences AoxB ont été trouvées et la divergence entre les séquences des alpha-proteobacteries et des beta-, gamma-proteobacteries a été décrite en détail. Une

séquence signature pour les alpha-protéobactéries a été proposée. L'ensemble des analyses fait l'objet d'une publication en cours de finalisation et est présenté ci-après.

II MANUSCRIT

Characterization of an arsenic contaminated environment, Sainte-Marie-aux-Mines, France, revealed a large prokaryotic diversity with arsenic-specific adaptation capacities

Audrey Heinrich-Salmeron, Audrey Cordi, Sébastien Gallien, David Halter, Christophe Pagnout, Florence Goulhen-Chollet, Alain Van Dorsselaer, Christine Schaeffer, Philippe Bertin, Pascale Bauda, Florence Arsène-Ploetze.

En préparation.

Characterization of an arsenic contaminated environment, Sainte-Marie-aux-mines, France, revealed a large prokaryotic diversity with arsenic-specific adaptation capacities.

Audrey Heinrich-Salmeron^{1*}, Audrey Cordi^{2*}, Sébastien Gallien³, David Halter¹, Christophe Pagnout², Florence Goulhen-Chollet¹, Elham Abbaszadehfard², Alain Van Dorsselaer³, Christine Schaeffer³, Philippe Bertin¹, Pascale Bauda², Florence Arsène-Ploetze^{1,&}.

¹ Laboratoire de Génétique Moléculaire, Génomique, Microbiologie, Département Microorganismes, Génomes, Environnement, UMR7156 Université de Strasbourg/CNRS, 28 rue Goethe, 67083 Strasbourg Cedex, France.

² Laboratoire des Interactions Ecotoxicologie Biodiversité Ecosystèmes (LIEBE), UMR7146, CNRS, Université Paul Verlaine, Campus Bridoux, rue du Général Delestraint, 57070 Metz, France.

³ Laboratoire de Spectrométrie de Masse Bio-organique, Institut Pluridisciplinaire Hubert Curien, UMR7178 Université de Strasbourg/CNRS, 25 rue Becquerel, 67087 Strasbourg, France.

* Equal contributors

& Corresponding author : Florence.ploetze@gem.u-strabg.fr

ABSTRACT

The Sainte-Marie mine site (France) is characterized by a mildly arsenic contamination and a neutral pH. An integrated study coupling chemistry, phylogeny of cultured and uncultured prokaryotes, metaproteomics and *aox* gene diversity analyses was performed to explore arsenic-dependent adaptation capacities of the microbial community. A large diversity was observed but only some of these prokaryotes were arsenic resistant and most of them belonged to Gamma-proteobacteria. Such bacteria are particularly adapted to persist in this environment as illustrated by our metaproteomic analyses. In addition, 77 *aoxB* genes that we classified in 4 distinct groups were characterized from both cultured and uncultured strains. In depth analyses of these genes allowed the definition of specific signature sequences for Alpha-proteobacteria. Altogether, this study revealed that despite a moderate arsenic contamination, some of the prokaryotes found in this environment have developed arsenic-specific adaptation processes.

INTRODUCTION

Arsenic is one of the most toxic inorganic elements in the environment and constitutes an important public health problem in many countries (Smith *et al.*, 2002). This metalloid is distributed ubiquitously around the world, with a large panel of concentrations. Indeed, uncontaminated soil or sediment, contain 0.1 to 40 mg Kg⁻¹, with a mean value of 6 mg Kg⁻¹ (Mandal & Suzuki, 2002). However, this value varies considerably among geographical regions depending on soil geochemical characteristics (natural contamination) or industrial activities (anthropogenic contaminations). Arsenic contents in unpolluted waters ranges from 1 to 10 µg L⁻¹, however 5 mg L⁻¹ arsenic can be observed in contaminated waters in USA, South America, India or Bangladesh (Smedley & Kinniburgh, 2002). Arsenic-rich waters are mainly located in geothermal regions or in mining activity areas due to Acid Mine Drainage (AMD). AMD are generally characterized by elevated sulfate, iron and other metal concentrations (Coupland *et al.*, 2004)(Hallberg & Johnson, 2005). In these environments, arsenic occurs mainly as inorganic species, arsenate (As(V)) or arsenite (As(III)) (Mandal & Suzuki, 2002). As(III) is more soluble, more mobile and therefore considered as more toxic than As(V).

Some arsenic-resistant archaea and bacteria, especially Proteobacteria, are able to oxidize As(III) to As(V) in order to detoxify their immediate environment. It has been proposed that the bacterial As(III) oxidation contributes to a natural attenuation of arsenic contamination by decreasing arsenic bioavailability (Lievremont *et al.*, 2009). This bacterial oxidation involves an arsenite oxidase encoded by *aoxAB*, (*aroAB* or *asoAB* genes). In *Alcaligenes faecalis*, it has been shown that the

arsenite oxidase belongs to the dimethyl sulfoxide (DMSO) reductase of the molybdenum family (Anderson *et al.*, 1992) (Ellis *et al.*, 2001). This protein is an heterodimer comprising a large subunit with a molybdenum center and a [3Fe-4S] cluster and a small subunit with a Rieske-type [2Fe-2S] cluster. The *aoxAB* genes encoding arsenite oxidase, were also described, in *Herminiimonas arsenicoxydans*, and are co-transcribed as an operon (Muller *et al.*, 2003) (Muller *et al.*, 2007). Other As-oxidizing bacteria recently studied include *Thermus* HR13 (Gihring & Banfield, 2001) (Gihring *et al.*, 2001), *Hydrogenophaga* sp. Str. NT-14 (Vanden Hoven & Santini, 2004), *Agrobacterium tumefaciens* (Kashyap *et al.*, 2006), *Ochrobactrum tritrici* SCII24 (Branco *et al.*, 2009), *Rhizobium* NT-26 (Santini *et al.*, 2004), strain MLHE-1 (Oremland *et al.*, 2002) and *Thiomonas* sp. (Duquesne *et al.*, 2008). Besides these well characterized arsenic-oxidizing bacteria, *aox* genes have been found in many phylogenetically diverse bacteria distributed among 25 genera, isolated from various arsenic-rich environments (Quéméneur *et al.*, 2008) (Inskeep *et al.*, 2007) (Chang *et al.*, 2009). For these reasons those genes have been considered as molecular markers of arsenic contamination (Quéméneur *et al.*, 2008). Therefore, one recent challenge was to investigate the diversity of *aox* genes in bacteria, even in uncultured bacteria. More generally, new approaches have been recently developed to access to the gene pool of the uncultured microorganisms (Bertin *et al.*, 2008). These approaches include phylogenetic and genomic content analysis (metagenomic) as well as global analysis of proteins synthesized by the communities (metaproteomic). Such approaches were recently used to analyze arsenic-rich AMD (Tyson *et al.*, 2004)(Tyson *et al.*, 2005)(Goltsman *et al.*, 2009) and also to identify new *aox* genes in environmental DNA extracted from arsenic-contaminated sites (Inskeep *et al.*, 2007). Besides these studies performed on arsenic-rich sites, where the *aox* genes analyses presented an evident interest, their presence and diversity in the prokaryotic community of slightly contaminated samples was poorly documented, such analyses are yet of particular interest in order to explore arsenite oxidation capacities in a community presenting high biodiversity.

In this study we explored the bacterial community and their arsenic-specific adaptation genes in a newly characterized environment located in the Sainte-Marie-aux-Mines valley, eastern France. Using environmental genomic approaches coupled with classical phylogenetical and molecular analyses, novel arsenic-resistant bacteria and *aox* genes were characterized.

MATERIALS AND METHODS

Site sampling procedure and physicochemical determinations *in situ*

Sampling site corresponds to a creek localized downstream an ancient mining exploitation at Sainte-Marie-aux-Mines, France. Samples were collected at different points along the creek (30 m long area) and homogenized. One litter water was filtered through a 0.22 μm sterile nucleopore filter and stored at 4°C. For further chemical analysis, water and sediments were separated. Sediments were stored at -80°C for further DNA extraction or in glycerol (25%, w/v) at -80°C for further protein extraction.

Chemical analysis

Water samples were collected into polyethylene bottles, kept cold (4°C) and analyzed within 48 h. Stream pH was measured in the laboratory using a microprocessor pH meter (pH 3000, WTW), and acid-neutralizing capacity (ANC) was determined by Gran's titration. Conductivity was measured with a Metrohm Herisau Conductometer E518 (Herisau, Switzerland) at 25°C. Concentrations of metals-metalloids (after acidification with HNO_3) were determined by atomic absorption spectrophotometry (Aanalyst 100; Perkin Elmer and Varian SpectrAA-300) and concentrations of Br^- , Cl^- , SO_4^{2-} , NO_3^- and NO_2^- by ion chromatography (Dionex 1500i with a AS 4 A SC column; Sunnyvale, USA). Water concentrations of suspended solids, total phosphorus (TP), total Kjeldahl nitrogen (TKN), and COD (Chemical Oxygen Demand), BOD5 (Biological Oxygen Demand after 5 days), and TSS (Total Suspended Solids) were determined according to the French norms for water analyses (AFNOR, 1994). Total phosphorus was determined according to the French norm AFNOR NF T90-023. Chlorophyllian pigments were quantified by spectrophotometry (Specord 205) at 750 and 755 nm.

Mineralization of sediments (0.5 g dry weight) was performed according to the French norm AFNOR NFX 31-151, before metal analysis by atomic absorption spectrophotometry. Organic matter content was determined by mass loss after treatment at 550°C during 4 hours.

Bacterial strains and plasmids used, culture conditions and DNA analyses

Characteristics of bacteria, plasmids and primers used in this study are presented in Table 1. Bacteria isolation was performed by adding 20 mL of sterile physiological water (NaCl 8.5 g L^{-1}) to 10 g of collected sediment. Several dilutions (1/5, 1/25, 1/125, 1/625, 1/3125, 1/15625 and 1/78125) were spread on DR2A medium (Tamaki *et al.*, 2005), gelled with agar or gellane gum, in absence of arsenic or in media containing 7 mM As(III) or 20 mM As(V). Strains were grown 18 days at 20°C, in aerobic or anaerobic conditions (GazPack system).

DNA was extracted directly from the sediment community using PowerMax Soil DNA isolation Kit according to the manufacturer recommendation (MoBio Laboratories, Inc.), concentrated by precipitation in 2.5 v/v 100% ethanol and 0.1 v/v sodium acetate (3M, pH 5.6) and

stored at -20°C until further processing. DNA extractions from cultivable strains were performed using UltraClean™ Microbial DNA Isolation Kit according to the manufacturer recommendations (MoBio Laboratories, Inc.).

16S rRNA specific PCR amplifications using the cultivable strain DNA as matrix were performed with primers W01 and W02 (Table 2). PCR reactions were carried out with a Mastercycler gradient (Eppendorf), with an initial denaturation step at 95°C for 5 min, followed by 35 cycles of denaturation (40 s at 95°C), an annealing step of 30 s at 52°C, 45 s elongation step at 68°C and a final denaturation step of 10 min at 68°C. Products were separated by agarose gel electrophoresis, purified by PCR purification Kit Qiagen and sequenced by GATC Biotech.

1.5 kb bacterial and archaeal 16SrRNA encoding genes were amplified using the metagenomic DNA as a matrix, and the universal bacterial and archaeal primers W01/W02 and ARC344F/ARC915R, respectively. The obtained fragments were cloned in plasmids pCR®2.1 (Table 1) using TOPO® TA cloning Kit according to the manufacturer recommendation (INVITROGENE Laboratories). 500 clones containing bacterial 16S rRNA encoding gene and 260 clones containing archaeal 16s RNA were obtained, among them 248 bacterial and 96 archaeal fragments were sequenced using the primers W01 or ARC344F, respectively (Table1). Sequences harboring more than 400 pb were used for phylogenetic analysis. Indeed, these sequences were compared to those in the GenBank database or RDP database by using BLAST searches (Altschul *et al.*, 1990).

Arsenite oxidase (*aoxB*) gene amplifications were performed by PCR with primers *aoxBM1-2F/aoxBM3-2R* carrying or not *XhoI* and *EcoRI* cassettes, using either the DNA extracted from cultivable strains or from the sediment community. The PCR reactions were carried out with an initial denaturation step at 95°C for 1 min, followed by 35 cycles of denaturation (40 s at 95°C), an annealing step of 40 s at 50°C, and a 1.5 min elongation step at 72°C. The products were separated by electrophoresis in 0.7% low-melting agarose and 5-10 kb fragments were purified either by QIAquick Gel Extraction Kit (Qiagen laboratories) or NucleoSpin extract II Kit. When only one fragment was obtained (cultivable strains), this fragment was directly sequenced. Otherwise, the fragments were ligated directly into pcR2.1, or digested with *XhoI* and *EcoRI* for cloning in pcDNA2.1 vector. The ligated product was electroporated into *E. coli* XL1 blue competent cells at 1700V. White transformants were selected on ampicillin-containing LB agar plates (100 mg L⁻¹) supplemented with X-gal (40 µg m L⁻¹) and IPTG (20 mM). Each clone were suspended in a lysis solution (0.05 M NaOH, 0.25% SDS) and boiled 10 min. Volume was adjusted at 200 µl with ddH₂O and PCR were performed as described previously with primers M13-20 primer and Reverse primer (Table 2), with an annealing step at 54°C, and a 6 min elongation step. These fragments were sequenced.

Metaproteomic analyses

To recover microbial cells from sediments, 8 x 12.5 g of frozen sediment were washed with 8 x 12.5 mL of salts solution (0.15 g L⁻¹: Na₂SO₄, 0.45 g L⁻¹: (NH₄)₂SO₄, 10H₂O, 0.05 g L⁻¹: KCL, 0.5 g L⁻¹: MgSO₄, 7H₂O, 0.05 g L⁻¹: KH₂PO₄, 0.014 g L⁻¹: Ca(NO₃)₂, 4H₂O) and gently shaken one night at 4°C. After 30 min decantation, the supernatant (8 x 7.5 mL) was added without mixing on Nycodenz solution (8 x 17.5 mL, 65%, w/v) (Axis-Shield, Dundee, Scotland), and then centrifuged 3 h at 10,000 g. The cellular fraction was removed and washed by adding 3 volumes of NaCl 0.9% and centrifuged 1 h at 10,000 g and 4°C.

Cells were resuspended in NaCl 0.9% and stained using the LIVE/DEAD BacLight Bacterial Viability kit (L13152, Molecular Probes, Invitrogen) according to the manufacturer recommendation. Preparation was incubated in darkness for 15 min at room temperature and viewed using an epifluorescent microscope Nikon Eclipse TE2000 S. Images were captured with an integrated Hamamatsu digital camera (C4742-95; Nikon UK).

Proteins were extracted from cells and separated by 1D-SDS-PAGE, according the method of Laemmli (Laemmli, 1970) using a 12% gradient slab gels (PROTEAN II, Bio-Rad laboratories). Electrophoresis was carried out at 200 V. Proteins were stained with silver nitrate, bands were cut every 1.5 mm and the 29 gel pieces were stored at -20°C before mass spectrometry analysis.

In gel digestion of gel bands was performed as previously described (Weiss *et al.*, 2009). The resulting peptide were analyzed by nanoLCMS/MS on nanoACQUITY Ultra-Performance-LC (UPLC, Waters, Milford, MA) coupled to SYNAPT hybrid quadrupole orthogonal acceleration time-of-flight tandem mass spectrometer (Waters, Milford, MA). The capillary voltage was set at 3,500 V and the cone voltage at 35 V. For tandem MS experiments, the system was operated with automatic switching between MS and MS/MS modes. The 3 most abundant peptides, preferably doubly and triply charged ions, were selected on each MS spectrum for further isolation and CID fragmentation with 2 energies set using collision energy profile. The complete system was fully controlled by MassLynx 4.1 (SCN 566, Waters, Milford, MA). Raw data collected during nanoLC-MS/MS analyses were processed and converted with ProteinLynx Browser 2.3 (23, Waters, Milford, MA) into .pkl peak list format.

The MS/MS data were analyzed using the MASCOT 2.2.0. algorithm (Matrix Science, London, UK). Spectra were searched with a mass tolerance of 30 ppm for MS and 0.1 Da for MS/MS data, allowing a maximum of one missed cleavage with trypsin and with carbamidomethylation of cysteins and oxidation of methionines specified as variable modifications. Spectra were firstly searched against a target-decoy version of NCBI nr database restricted to bacterial organisms. Protein identifications were validated when at least two peptides with high quality MS/MS spectra (less than

20 points below Mascot's threshold score of identity at 95% confidence level) were detected. All spectra that did not satisfy these criteria were exported using Scaffold software 2.2.0 (Proteome Software, Portland, USA) and searched against in the same database. Protein identifications were validated when one peptide with a very high quality MS/MS spectrum (more than 19 points above Mascot's threshold score of identity at 95 % confidence level) was detected. All spectra that did not pass this threshold were exported again and searched against a target-decoy version of NCBI nr database restricted to eukaryotic organisms. Protein identifications were validated when at least two peptides with high quality MS/MS spectra (less than 16 points below Mascot's threshold score of identity at 95% confidence level) were detected. In the case of one-peptide hits, the score of the unique peptide had to be greater (minimal "difference score" of 4) than the 95% significance Mascot threshold. After removal of common contaminants, such as keratins and trypsin, these thresholds allowed identifying 27 unique proteins with an estimated false discovery rate less than 3%.

RESULTS

Sampling site description and chemical analysis

According to geochemical data (<http://infoterre.brgm.fr>), sediments or soils of several sites in eastern France (Vosges) are contaminated with geochemical origin arsenic (above 100 mg kg⁻¹). A preliminary sampling was performed at those sites and arsenic-specific markers (*aox* genes) were amplified from extracted DNA samples. The site with the highest arsenic contamination presented a positive *aox* specific PCR amplification whereas the others did not (data not shown). This site should contain specifically adapted bacteria to arsenic-contaminated environment and was therefore chosen.

This sampling point is localized downstream of the Gabe-Gottes mine site and located in the Sainte-Marie-aux-Mines valley (France) (Fig. 1). In this mining exploitation, extractions processed until 1940, and were composed of silver ores, and more recently of arsenic ores. Sediments of a creek emerging from the old mine and located 40 m of the mine entrance were collected in December 2007 (N 48°12'863, E 007°09'569). The physicochemical composition of sediment and water samples was analyzed (Table 2), revealing a moderate arsenic concentration in sediments (320 mg kg⁻¹), but this contamination was higher than the average geochemical content in this region (48 mg kg⁻¹ in the Vosges, <http://infoterre.brgm.fr>). These samples were also slightly contaminated by other metals such as Al, Mn and Fe, Cr, Cu, Ni, Pb and Zn.

Composition of the bacterial community in sediments

An inventory of the bacterial and archaeal population present at this sampling point was performed by combining a culture-based approach and phylogenetic analyses, using 16S rRNA-based community analysis. The 16S rRNA gene library analyses revealed a majority of bacteria from the phyla Bacteroidetes (12%), Betaproteobacteria (15%), Acidobacteria (17%) and Deltaproteobacteria (12%) and a minority (5% to 10%) of Alphaproteobacteria, Actinobacteria, Gammaproteobacteria, Nitrospira and Planctomycetes (Figure 2). Regarding the archaeal population, sediments contained Euryarchaeota (40%) (Methanomicrobia (12%, 1% *Methanothrix*), Methanobacteria (3.5%, 2.3% *Methanobacterium*) and 1% Thermococcus (*Pyrococcus*)) and Crenarchaealota (60%) mainly Thermoprotei (55%).

As a complement to this culture independent approach, the cultivable cells were numbered and phylogenetically analyzed. Among the 1.35×10^7 CFU g⁻¹ cells cultured under aerobic conditions, 20 to 57% (depending on the gelling agent) corresponded to bacteria resistant to high As(V) concentration (20 mM), whereas less than 3% of bacteria grew in the presence of 7 mM of As(III). The total cell number was lower under anaerobic conditions (3.75×10^5 CFU g⁻¹), as compared to aerobic conditions with 60% of As(V) resistant and less than 3% of As(III) resistant bacteria. Among the As(III) and As(V) resistant strains, 260 bacteria were phylogenetically characterized using a 16S rRNA-based analysis (Figure 2). These strains are phylogenetically related to 4 distinct phyla: 61% belonged to Proteobacteria (4% alpha-, 4% beta- and 53% gamma-proteobacteria), 27% to Actinobacteria, 11% to Bacteroidetes and 1% to Firmicutes (Figure 3). Interestingly, *Pseudomonas* sp., *Flavobacterium* sp. and *Rhodococcus* sp. were over-represented, whereas no *Acidobacteria* were isolated even if this phylum is highly represented in the culture independent approach. A same proportion of Firmicutes and Actinobacteria was found among the As(V) and As(III) resistant cultivable cells, whereas a higher proportion of bacteria belonging to the Bacteroidetes phylum was observed among the As(V) resistant bacteria as compared to the As(III) resistant strains, and a higher proportion of Gamma-Proteobacteria were isolated in As(III) containing media as compared to As(V) containing media. Altogether, these phylogenetic analyses revealed a huge diversity among both the bacterial and the archaeal communities found in this arsenic contaminated environment.

Metaproteomic analysis of the sediment community

To get insight into the major community functionalities and to highlight the active community *in situ*, a metaproteomic analysis was performed. Microorganisms were separated from sediments and interstitial water by a Nycodenz density gradient. An aliquot of these cells were stained with a live/dead BacLight kit (Invitrogen), revealing that the majority of the cells extracted from the sediments had intact membranes (data not shown). Moreover, observations of these cells

revealed that bacteria were embedded in extracellular matrix. A total of 27 proteins expressed from these cells were identified, 58% originated from bacteria (19 proteins), 36% from plant and algae (6 proteins each) and 6% from a eukaryotic microorganism *Euglena* (Table 3).

The nineteen identified bacterial proteins were similar to proteins from Firmicutes, Cyanobacteria, Alpha-, Beta-, Gamma and Delta-Proteobacteria, with a large proportion of proteins originated from Gamma-Proteobacteria. These proteins are involved in replication and transcription (ribosomal proteins, elongation factors, helicases), energy metabolism (D-ribose binding proteins, ATP synthase), carbon metabolism (amino acids synthase, formate acetyltransferase, tryptophanase), cell wall synthesis (gentamicin acetyltransferase), in regulation (histone family, transcriptional regulator) and stress response (Hsp90, helicase, ethanolamine utilization, thiol peroxidase). These findings indicated that these bacteria had an active energetic and carbon metabolism enabling them to persist in this environment.

No proteins involved in arsenic response were identified among the major proteins expressed by this community, as for example the arsenite oxidase or the *ars* encoded proteins. Such observation was expected considering that these proteins are not strongly expressed in arsenite-oxidizing bacteria cultivated in laboratory conditions (Weiss *et al.*, 2009) (Bryan *et al.*, 2009). In order to test whether bacteria are specifically adapted to arsenic, the *aox* genes specific arsenic markers were amplified from DNA extracted from both cultured and uncultured strains.

Arsenite oxidase encoding operon analysis

First, arsenite oxidase encoding genes were successfully amplified from 22 strains among the 260 isolated strains. According to phylogenetic analyses, these strains belonged to *Flavobacterium*, *Bosea*, *Pseudomonas*, *Achromobacter*, *Agromyces*, *Acinetobacter* genera. Moreover, we identified an *aoxB* gene in two Gram-positive bacteria belonging to *Bacillus* and *Rhodococcus* genus, whereas no *aoxB* sequences were amplified from anaerobe isolates. In parallel, 55 *aoxB* genes were amplified and sequenced from metagenomic DNA. Altogether, 77 *aoxB* genes were further analyzed. Among these sequences, the corresponding amino acid sequences of two of them shared only limited similarities with sequences found in databases. Indeed, the AoxB C3 sequence harbored 52% identity (170/322 amino acids) with *Chloroflexus aurentiacus* J-10-f1 and the C15A sequence presented 33% identity (133/401 amino acids) with *C. aurentiacus* J-10-f1. The comparison of the 77 AoxB sequences (supplementary Figure 1) revealed that residues present in the catalytic site or arsenic binding site were conserved in almost all sequences including the more divergent sequences C3 and C15A, related to *Chloroflexus* (data not shown). These sequences were therefore considered as AoxB homologues.

In depth analyses of the AoxB sequences revealed that the Rieske binding motif C₂₁-X₂-C₂₄-X₃-C₂₈-X₇₀-S₉₉ (according to *A. faecalis* protein sequence numbering) is conserved in almost all the 77 sequences but the 70 amino acids localized between C28 and S99 (according to *Alcaligenes* sequence) were less conserved, and S99 is only conserved in Beta- and Gamma-proteobacteria. The amino acids corresponding to *Alcaligenes* H195 and E203 implicated in the arsenic-binding site were conserved in all sequence except C15A. As observed previously among 14 *aoxB* of Proteobacteria (Quéméneur *et al.*, 2008), the motif HNRPAYNSE involved in arsenic binding was conserved in all sequences.

Interestingly, these phylogenetic analyses revealed the existence of four groups: two groups contained sequences originated from Alpha-proteobacteria and the two other groups contained sequences from *Thermus/Chloroflexus* and Beta- and Gamma-proteobacteria, respectively. Indeed, beyond the well conserved functional regions, the AoxB sequences from Beta- and Gamma-proteobacteria shared a significant conservation (Supplementary Figure 1). Because of such similarities, the metagenomic sequences C22, C40, C11, C15, C1A, C8, C17 and C18 could not be assigned either to Beta or Gamma-proteobacteria affiliated sequences. Nevertheless, AoxB comparison suggests that sequences C1A and C40 were more closely related to Beta-proteobacteria than to Gamma-proteobacteria.

Surprisingly, no strong divergence was observed among AoxB sequences from Gram-positive bacteria and the Beta- or Gamma-proteobacteria. On the contrary, a stronger divergence was observed between these sequences and AoxB from Alpha-proteobacteria or from *Chloroflexus/Thermus* affiliated sequences (Figure 4). Indeed, all Alpha-proteobacteria sequences present two divergent regions as compared to other proteobacteria, at position 59 to 78 and 187 to 191 (according to *A. faecalis* protein sequence numbering) (supplementary fig. 1). This last region located between an α -helix and a β -sheet upstream from the arsenic-binding site present a perfect conservation among Alpha-proteobacteria. These residues (M-K-V-K-N/H-I/C) may be thus considered as an Alpha-proteobacteria signature sequence. Moreover, among sequences originated from Alpha-proteobacteria, two subgroups were distinguishable, because residues from 257 to 275 (according to *A. faecalis* protein sequence numbering) were absent in the C19A, C29, C2 and *Roseavarius* AoxB sequences, and a strong conservation was observed with amino acid surrounding these sequences. This region was different as compared to the other Alpha-proteobacteria AoxB (i.e. *Bosea*, 43. *Bosea*, *Nitrobacter*, *Thiobacillus*, C5A and C17A). Altogether, thanks to the AoxB sequences identified in this study from cultured strains and community analyses, new *aoxB* genes that were divergent from known sequences were identified. Moreover, these analyses revealed that AoxB

sequences could be classified in 4 distinct groups, with particular signature sequences for Alpha-proteobacteria.

DISCUSSION

In the present work we studied a mildly arsenic-contaminated site. The physico-chemical analyses of the samples revealed a moderate but significant arsenic contamination (320 mg Kg^{-1} and $6 \mu\text{g L}^{-1}$ in sediments and in water, respectively) and a neutral pH. Despite this arsenic contamination, a high prokaryotic diversity was observed that make the study of this site of particular interest to analyze the diversity of arsenic adaptation processes. Indeed, analyses of cultured and uncultured strains and metaproteomic approaches allowed us to draw up an inventory of the micro-organisms found in this moderately contaminated environment. Interestingly, several observations suggested that bacteria have develop several mechanisms of adaptation to arsenic in this environment, as illustrated by the presence of at least 77 different *aox* genes, the *in situ* synthesis of proteins involved in stress response, and the observation that bacteria were embedded in extracellular matrix, a mechanism previously proposed to be involved in arsenic sequestration (Muller *et al.*, 2007).

In the present study, 16S rRNA analyses revealed that at least 17 bacterial and two archaeal phyla were represented. Such a diversity is large as compared to previous analyses on more contaminated sites (Baker & Banfield, 2003) (Johnson & Hallberg, 2003), as for example in Carnoules AMD (less than 10 phyla, (Bruneel *et al.*, 2005) (Bruneel *et al.*, 2006) (Bruneel *et al.*, 2008), in metagenomic analyses on AMD biofilm (4 to 6 organisms type) (Tyson *et al.*, 2004). Similarly, such a diversity was not observed in geothermal springs (6 phyla, (Hamamura *et al.*, 2009)).

As usually observed, no *Acidobacteria* were isolated even if this phylum is highly represented in the total phylogenetic analysis performed with the 16S rRNA-based approach, suggesting that the community is more complex than supposed on the basis of the cultured population. This reinforces the interest of using culture independent approaches to explore the diversity of environmental communities. Nevertheless, among the proteins identified in our metaproteomic analyses, no protein originated from *Acidobacteria* was found, maybe because sequences are not enough represented in sequence databases or because such bacteria were less abundant than other bacteria, or were easily lost during extraction.

The metaproteomic data revealed that a large diversity of organisms expressed proteins in the environment under study, suggesting that such prokaryotes persist in this environment. 19 proteins originating from bacteria were identified and the metabolic function of some of them revealed that several bacteria found have metabolic activities and may be particularly adapted to grow in this environment. Interestingly, among these bacteria, the largest proportion of arsenic

resistant bacteria belonged to Gamma-proteobacteria and a majority of bacterial proteins identified probably originated from organisms related to this phylum. These observations suggest that a large proportion of the Gamma-proteobacteria are particularly well adapted to this environment. In addition, proteins originated from *Euglena* were identified in our metaproteomic analyses. Eukaryotic photosynthetic *Euglena* sp. is a common inhabitant of arsenic contaminated environments such as AMD, and it is frequently considered as an 'indicator species' of these environments (Amaral Zettler *et al.*, 2003)(Brake *et al.*, 2001a)(Olaveson & Nalewajko, 2000)(Brake *et al.*, 2001b).

A large proportion of bacteria found at the study site were not arsenic resistant. Indeed, according to our quantification of cultured arsenic resistant strains, 20 to 57% were As(V) resistant, whereas only less than 3% were As(III) resistant. Nevertheless, because of the huge amount of cells present at this site (at least 2×10^7 CFU g⁻¹), we isolated 260 arsenic resistant different strains. Among those bacteria, strains belonged to phyla that are rarely found in arsenic-rich environments, e.g. as Bacteroides. These observations suggest that the diversity of arsenic resistant bacteria and archaea is larger than expected so far.

Bacteria resist to arsenic by expressing *ars* genes (Mukhopadhyay & Rosen, 2002). Recently, metagenomic explorations revealed that *ars* genes are widely spread in micro-organisms, even in bacteria found in environment that are not contaminated by arsenic (Chauhan *et al.*, 2009). These *ars* genes cannot therefore be considered as markers of arsenic contamination. In contrast, *aox* genes may be considered as arsenic contamination specific markers. In the present work, several *aox* genes were characterized from isolated clones or uncultured bacteria. Indeed, culture-based and environmental genomic approaches were combined, allowing identifying 77 sequences. These analyses thus extend previous analyses of AoxB phylogenetic clusters suggesting that Alpha-, Beta-, Gamma-proteobacteria and *Chloroflexus* or *Thermus* are well separated (Quéméneur *et al.*, 2008). Our study goes nevertheless further into detail since three particular regions specific to Alpha-proteobacteria were defined in AoxB sequences. Moreover, the present study allowed us to include *aoxB* genes from Gram-positive bacteria in phylogenetic analyses (*Bacillus* and *Rhodococcus*), showing that these sequences are more related to Gamma- and Beta-proteobacteria than the Alpha-proteobacteria sequences. These observations suggest that *aox* gene evolution does not correlated strictly to bacterial evolution. Finally, novel *aox* sequences close to *Chloroflexus* but phylogenetically far from the previously known sequences were characterized. Thus these observations show that the *aoxB* gene diversity is larger than expected by analyzing arsenic-rich environment. Taken together, our results demonstrated that (meta)genomic approaches are of particular interest to explore the

diversity of arsenic adapted prokaryotic communities, including specific genes such as *aoxB*, in mildly arsenic contaminated environment.

ACKNOWLEDGMENTS

Audrey Cordi was supported by grants of the French Ministry of Research and the “Zone Atelier Moselle” (ZAM). Audrey Heinrich-Salmeron, David Halter and Sebastien Gallien were supported by grants of the French Ministry of Research and the Université de Strasbourg. Florence Goulhen-Chollet was supported by the Agence Nationale de la Recherche (ANR), RARE project (Reactivity of an Arsenic-Rich Ecosystem).

The study was financed by the EC2CO program (Institut National des Sciences de l'Univers, CNRS). This work was performed in the framework of the Groupement de Recherche “Métabolisme de l'Arsenic chez les micro-organismes: de la résistance à la détoxification” (GDR2909-CNRS).

REFERENCES

- AFNOR (1994).** Qualité de l'Eau 2Ed (Paris 1994), 826.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990).** Basic local alignment search tool. *J Mol Biol* **215**, 403-410.
- Amaral Zettler, L. A., Messerli, M. A., Laatsch, A. D., Smith, P. J. S. & Sogin, M. L. (2003).** From genes to genomes: Beyond biodiversity in Spain's Rio Tinto. *Biological Bulletin* **204**, 205-209.
- Anderson, G. L., Williams, J. & Hille, R. (1992).** The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum-containing hydroxylase. *Journal of Biological Chemistry* **267**, 23674-23682.
- Baker, B. J. & Banfield, J. F. (2003).** Microbial communities in acid mine drainage. *FEMS Microbiol Ecol* **44**, 139-152.
- Bertin, P. N., Médigue, C. & Normand, P. (2008).** Advances in environmental genomics: Towards an integrated view of micro-organisms and ecosystems. *Microbiology* **154**, 347-359.
- Brake, S. S., Dannelly, H. K. & Connors, K. A. (2001a).** Controls on the nature and distribution of an alga in coal mine-waste environments and its potential impact on water quality. *Environmental Geology* **40**, 458-469.
- Brake, S. S., Dannelly, H. K., Connors, K. A. & Hasiotis, S. T. (2001b).** Influence of water chemistry on the distribution of an acidophilic protozoan in an acid mine drainage system at the abandoned Green Valley coal mine, Indiana, USA. *Applied Geochemistry* **16**, 1641-1652.

- Branco, R., Francisco, R., Chung, A. P. & Morais, P. V. (2009).** Identification of an *aox* system that requires cytochrome c in the highly arsenic resistant strain *Ochrobactrum tritici* SC1124. *Appl Environ Microbiol.*
- Bruneel, O., Duran, R., Koffi, K., Casiot, C., Fourçans, A., Elbaz-Poulichet, F. & Personné, J. C. (2005).** Microbial diversity in a pyrite-rich tailings impoundment (Carnoulès, France). *Geomicrobiology Journal* **22**, 249-257.
- Bruneel, O., Duran, R., Casiot, C., Elbaz-Poulichet, F. & Personné, J. C. (2006).** Diversity of microorganisms in Fe-As-rich acid mine drainage waters of Carnoulès, France. *Appl Environ Microbiol* **72**, 551-556.
- Bruneel, O., Pascault, N., Egal, M., Bancon-Montigny, C., Goñi-Urriza, M. S., Elbaz-Poulichet, F., Personné, J. C. & Duran, R. (2008).** Archaeal diversity in a Fe-As rich acid mine drainage at Carnoulès (France). *Extremophiles* **12**, 563-571.
- Bryan, C. G., Marchal, M., Battaglia-Brunet, F., Kugler, V., Lemaitre-Guillier, C., Lievremont, D., Bertin, P. N. & Arsene-Ploetze, F. (2009).** Carbon and arsenic metabolism in *Thiomonas* strains: differences revealed diverse adaptation processes. *BMC Microbiol* **9**, 127.
- Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987).** XL1-Blue: A high efficiency plasmid transforming *recA E. coli* strain with beta-galactosidase selection. *Biotechniques* **5**, 376-378.
- Chang, J. S., Yoon, I. H., Lee, J. H., Kim, K. R., An, J. & Kim, K. W. (2009).** Arsenic detoxification potential of *aox* genes in arsenite-oxidizing bacteria isolated from natural and constructed wetlands in the Republic of Korea. *Environ Geochem Health.*
- Chauhan, N. S., Ranjan, R., Purohit, H. J., Kalia, V. C. & Sharma, R. (2009).** Identification of genes conferring arsenic resistance to *Escherichia coli* from an effluent treatment plant sludge metagenomic library. *FEMS Microbiol Ecol* **67**, 130-139.
- Coupland, K., Battaglia-Brunet, F., Hallberg, K. B., Dictor, M.-C., Garrido, F. & Johnson, D. B. (2004).** Oxidation of iron, sulfur and arsenic in mine waters and mine wastes: an important role of novel *Thiomonas* spp. *Biohydrometallurgy: a sustainable technology in evolution* Edited by Tsezos AHM, Remondaki E Zografou, Greece:National Technical University of Athens, 639-646.
- Duquesne, K., Lieutaud, A., Ratouchniak, J., Muller, D., Lett, M. C. & Bonnefoy, V. (2008).** Arsenite oxidation by a chemoautotrophic moderately acidophilic *Thiomonas* sp.: From the strain isolation to the gene study. *Environmental Microbiology* **10**, 228-237.

Ellis, P. J., Conrads, T., Hille, R. & Kuhn, P. (2001). Crystal structure of the 100 kDa arsenite oxidase from *Alcaligenes faecalis* in two crystal forms at 1.64 Å and 2.03 Å. *Structure* **9**, 125-132.

Gihring, T. M. & Banfield, J. F. (2001). Arsenite oxidation and arsenate respiration by a new *Thermus* isolate. *FEMS Microbiology Letters* **204**, 335-340.

Gihring, T. M., Druschel, G. K., McCleskey, R. B., Hamers, R. J. & Banfield, J. F. (2001). Rapid arsenite oxidation by *Thermus aquaticus* and *Thermus thermophilus*: Field and laboratory investigations. *Environmental Science and Technology* **35**, 3857-3862.

Goltsman, D. S., Deneff, V. J., Singer, S. W. & other authors (2009). Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing "*Leptospirillum rubarum*" (Group II) and "*Leptospirillum ferrodiazotrophum*" (Group III) bacteria in acid mine drainage biofilms. *Appl Environ Microbiol* **75**, 4599-4615.

Hallberg, K. B. & Johnson, D. B. (2005). Microbiology of a wetland ecosystem constructed to remediate mine drainage from a heavy metal mine. *Sci Total Environ* **338**, 53-66.

Hamamura, N., Macur, R. E., Korf, S., Ackerman, G., Taylor, W. P., Kozubal, M., Reysenbach, A. L. & Inskeep, W. P. (2009). Linking microbial oxidation of arsenic with detection and phylogenetic analysis of arsenite oxidase genes in diverse geothermal environments. *Environmental Microbiology* **11**, 421-431.

Hanahan, D. (1985). DNA Cloning. *DNA cloning, ed Glover, DM (IRL Press, Oxford)* **1**, 109-135.

Inskeep, W. P., Macur, R. E., Hamamura, N., Warelow, T. P., Ward, S. A. & Santini, J. M. (2007). Detection, diversity and expression of aerobic bacterial arsenite oxidase genes. *Environmental Microbiology* **9**, 934-943.

Johnson, D. B. & Hallberg, K. B. (2003). The microbiology of acidic mine waters. *Research in Microbiology* **154**, 466-473.

Kashyap, D. R., Botero, L. M., Franck, W. L., Hassett, D. J. & McDermott, T. R. (2006). Complex regulation of arsenite oxidation in *Agrobacterium tumefaciens*. *Journal of Bacteriology* **188**, 1081-1088.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lievremont, D., Bertin, P. N. & Lett, M. C. (2009). Arsenic in contaminated waters: Biogeochemical cycle, microbial metabolism and biotreatment processes. *Biochimie*.

- Mandal, B. K. & Suzuki, K. T. (2002).** Arsenic round the world: a review. *Talanta* **58**, 201-235.
- Mukhopadhyay, R. & Rosen, B. P. (2002).** Arsenate reductases in prokaryotes and eukaryotes. *Environmental Health Perspectives* **110**, 745-748.
- Muller, D., Lièvreumont, D., Simeonova, D. D., Hubert, J. C. & Lett, M. C. (2003).** Arsenite oxidase *aox* genes from a metal-resistant β -proteobacterium. *Journal of Bacteriology* **185**, 135-141.
- Muller, D., Médigue, C., Koechler, S. & other authors (2007).** A tale of two oxidation states: Bacterial colonization of arsenic-rich environments. *PLoS Genetics* **3**, 0518-0530.
- Olaveson, M. M. & Nalewajko, C. (2000).** Effects of acidity on the growth of two *Euglena* species. *Hydrobiologia* **433**, 39-56.
- Oremland, R. S., Hoefl, S. E., Santini, J. M., Bano, N., Hollibaugh, R. A. & Hollibaugh, J. T. (2002).** Anaerobic oxidation of arsenite in Mono Lake water and by a facultative, arsenite-oxidizing chemoautotroph, strain MLHE-1. *Appl Environ Microbiol* **68**, 4795-4802.
- Quéméneur, M., Heinrich-Salmeron, A., Muller, D., Lièvreumont, D., Jauzein, M., Bertin, P. N., Garrido, F. & Joulain, C. (2008).** Diversity surveys and evolutionary relationships of *aoxB* genes in aerobic arsenite-oxidizing bacteria. *Appl Environ Microbiol* **74**, 4567-4573.
- Raskin, L., Stromley, J. M., Rittmann, B. E. & Stahl, D. A. (1994).** Groupspecific 16s rRNA hybridization probes to describe natural communities of methanogens. *Appl Environ Microbiol* **60**, 1232-1240.
- Santini, J. M., Streimann, I. C. A. & vanden Hoven, R. N. (2004).** *Bacillus macyae* sp. nov., an arsenate-respiring bacterium isolated from an Australian gold mine. *International Journal of Systematic and Evolutionary Microbiology* **54**, 2241-2244.
- Smedley, P. L. & Kinniburgh, D. G. (2002).** A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry* **17**, 517-568.
- Smith, A. H., Lopipero, P. A., Bates, M. N. & Steinmaus, C. M. (2002).** Arsenic epidemiology and drinking water standards. *Science* **296**, 2145-2146.
- Stahl, D. A. & Amann, R. I. (1991).** Development and application of nucleic acid probes in bacterial systematics. In Stackebrandt E, Goodfellow M (Eds). *Sequencing and hybridization techniques in bacterial systematics*, Wiley, Chichester, 205-248.

Tamaki, H., Sekiguchi, Y., Hanada, S., Nakamura, K., Nomura, N., Matsumura, M. & Kamagata, Y. (2005). Comparative analysis of bacterial diversity in freshwater sediment of a shallow eutrophic lake by molecular and improved cultivation-based techniques. *Appl Environ Microbiol* **71**, 2162-2169.

Tyson, G. W., Chapman, J., Hugenholtz, P. & other authors (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**, 37-43.

Tyson, G. W., Lo, I., Baker, B. J., Allen, E. E., Hugenholtz, P. & Banfield, J. F. (2005). Genome-directed isolation of the key nitrogen fixer *Leptospirillum ferrodiazotrophum* sp. nov. from an acidophilic microbial community. *Appl Environ Microbiol* **71**, 6319-6324.

Vanden Hoven, R. N. & Santini, J. M. (2004). Arsenite oxidation by the heterotroph *Hydrogenophaga* sp. str. NT-14: The arsenite oxidase and its physiological electron acceptor. *Biochimica et Biophysica Acta - Bioenergetics* **1656**, 148-155.

Weiss, S., Carapito, C., Cleiss, J. & other authors (2009). Enhanced structural and functional genome elucidation of the arsenite-oxidizing strain *Hermiimonas arsenicoxydans* by proteomics data. *Biochimie* **91**, 192-203.

FIGURES and TABLESTable 1: Strains, plasmids and primers used in this work. For *AoxXhoI* and *AoxEcoRI*, nucleotides red nucleotides corresponded to *XhoI* and *EcoRI* restriction site respectively.

	DESCRIPTION	REFERENCES
STRAINS		
<i>E. coli</i> DH5- α ATCC53868	Δ lac_(<i>lacZ</i>)	(Hanahan, 1985)
<i>E. coli</i> XL1	<i>recA</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi</i> , <i>hsdR17</i> (<i>rk-</i> , <i>mk+</i>), <i>supE44</i> , <i>relA1</i> , <i>lambda-</i> , <i>lac-</i>	(Bullock <i>et al.</i> , 1987)
PLASMIDS		
pcDNA2.1	<i>Ori ColE1</i> , Amp ^R , <i>flori</i> , <i>lacZ</i> , Promotor T7	Invitrogen
pCR@2.1	Amp ^R , Kn ^R ,	Invitrogen
AMORCES		
W01	5'-AGAGTTTGATCMTGGCTC-3'	Universal primer
W02	5'-GNTACCTTGTTACGACTT-3'	Universal primer
ARC344F	5'-ACGGGGYGCAGCAGGCGCA-3'	(Raskin <i>et al.</i> , 1994)
ARC915R	5'-GTGCTCCCCGCAATTCCT-3'	(Stahl & Amann, 1991)
BM1-2F*	5'-CCACTTCTGCATCGTGGGNTGYGGNTA-3'	(Quéméneur <i>et al.</i> , 2008)
BM3-1R*	5'-TGCCCCAGATGATGCCYTTYTCRWA-3'	(Quéméneur <i>et al.</i> , 2008)
M13-20	5'-GTAAAACGACGGCCAGT-3'	amerces universelles
Reverse primer	5'-GGAAACAGCTATGACCATG-3'	amerces universelles

*: Restriction site of *XhoI* and *EcoRI* were added to these primers for cloning into pcDNA2.1

Table 2: Physicochemical characteristics of water and sediment samples

	Water	Sediments
pH	6.9	-
Density (Kg/l)	-	2.07
Moisture 20° (%)	-	18
Moisture 105° (%)	-	19.3
Fraction <2mn (%)	-	57.8
Fraction >2mn (%)	-	42.2
OM on <2mn (%)	-	2
P total	-	0.08
Al	98 µg/l	15150 µg/g
As	6 µg/l	320 µg/g
Cd	ND	0.3 µg/g
Cr	ND	40 µg/g
Cu	1 µg/l	103 µg/g
Fe	ND	23745 µg/g
Mn	ND	510 µg/g
Ni	ND	23 µg/g
Pb	ND	49 µg/g
Zn	2 µg/l	88 µg/g
Conductivity (µS/cm at 25°C)	50.4	-
COD (mg O/l)	5	-
BOD5 (mg O/l)	0.4	-
NH₄ (mg N/l)	0.02	ND
NO₂ (mg N/l)	<0.01	ND
NO₃ (mg N/l)	1.2	ND
N Kj (mg N/l)	<0.1	ND
PO₄ (mg P/l)	0.02	ND
P total (mg/l)	0.03	ND
Cl (mg/l)	2.1	ND
SO₄ (mg/l)	6.2	ND
Br (mg/l)	0.004	ND
Alacalinity (meq/l)	0.18	-
Chlo a (µg/l)	0.3	-
Phéo (µg/l)	0.2	-
TSS (mg/l)	0.7	-
VSS (%)	100	-

Legend: ND: non detected NT: non tested. OM: Organic Matter, COD: Chemical Oxygen Demand, BOD5: Biological Oxygen Demand after 5 days, Chlo a: chlorophyll a, TSS: Total Suspended Solids, VSS: Volatile Suspended Solids.

Table 3: Proteins expressed by the community identified by MS/MS.

Protein name	Closest organisms	Protein accession numbers	Protein molecular weight (Da)	Number of unique peptides	Percentage sequence coverage
50S ribosomal protein L9	<i>Escherichia coli</i> O157:H7 EDL933	gi 15804792 ref NP_290833.1	15 750,7	4	34,90%
Chain C, Structure Of The 30s Subunit Of A Pre-Translocational E. Coli Ribosome	<i>Escherichia coli</i>	gi 116667412 pdb 2GY9 C,gi 1173247 sp P46172 RS3_BUCAK,gi 117620103 ref YP_854847.1	20 052,6	1	6,67%
Elongation factor Tu	<i>Shigella flexneri</i> 5 str. 8401	gi 110807172 ref YP_690692.1 ,gi 11514297 pdb 1D8T A,gi 157158302 ref YP_001465471.1 ,gi 157830940 pdb 1EFM A	31 508,0	5	24,90%
5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	<i>Alkalilimnicola ehrlichei</i> MLHE-1	gi 114321254 ref YP_742937.1 ,gi 149377318 ref ZP_01895064.1	85 687,0	1	1,32%
Ethanolamine utilization protein eutM	<i>Heliobacillus mobilis</i>	gi 111075049 gb ABH04894.1 ,gi 116619462 ref YP_821618.1 ,gi 116619522 ref YP_821678.1	8 785,9	1	12,60%
Gentamicin 3'-acetyltransferase (Gentamicin acetyltransferase I)(Aminoglycoside N(3')-acetyltransferase I) (AAC(3)-I)	<i>Serratia marcescens</i>	gi 112944 sp P23181 AAC1_PSEAE,gi 239721 gb AAB20441.1 ,gi 280053 pir A60451,gi 3745773 pdb 1BO4 A	19 396,7	2	11,30%
Probable formate acetyltransferase 3	<i>Shigella flexneri</i> 5 str. 8401	gi 110806996 ref YP_690516.1 ,gi 146313196 ref YP_00178270.1 ,gi 157155072 ref YP_001464583.1	85 939,7	1	1,44%
Tryptophanase	<i>Escherichia fergusonii</i>	gi 130833466 gb ABO32521.1 ,gi 130833475 gb ABO32527.1 ,gi 134105349 pdb 2OQX A	52 724,7	3	7,01%
ATP-dependent RNA helicase DeaD	<i>Shigella flexneri</i> 5 str. 8401	gi 110807031 ref YP_690551.1 ,gi 117625458 ref YP_858781.1 ,gi 12517770 gb AAG58298.1 AE005544_12	70 347,6	1	2,23%
Chain A, D-Ribose-Binding Protein From <i>Escherichia Coli</i>	<i>Escherichia Coli</i>	gi 3318867 pdb 1URP A	28 456,5	7	41,00%
FOF1 ATP synthase subunit beta	<i>Sphingopyxis alaskensis</i> RB2256	gi 103487769 ref YP_617330.1 ,gi 108759590 ref YP_635037.1 ,gi 109648044 ref ZP_01371947.1 ,	50 955,7	2	5,23%
Photosystem q(b) protein	<i>Trichodesmium erythraeum</i> IMS101	gi 113474085 ref YP_720146.1 ,gi 113478129 ref YP_724190.1 ,gi 148242539 ref YP_001227696.1	39 440,8	2	6,41%
Heat shock protein 90	<i>Pelobacter propionicus</i> DSM 2379	gi 118579799 ref YP_901049.1 ,gi 15617081 ref NP_240294.1	72 803,5	2	2,24%
Heat shock protein 90	<i>Rhodospseudomonas palustris</i> BisA53	gi 115526885 ref YP_783796.1 ,gi 146343100 ref YP_001208148.1 ,gi 147676445 ref YP_001210660.1	72 062,3	1	2,15%
Histone family protein DNA-binding protein	<i>Parvibaculum lavamentivorans</i> DS-1	gi 154250962 ref YP_001411786.1	12 014,1	1	8,85%
Transcriptional regulator HU subunit alpha	<i>Escherichia coli</i> O157:H7 EDL933	gi 15804591 ref NP_290632.1	9 517,2	3	63,30%
Major outer membrane lipoprotein precursor (Murein-lipoprotein)	<i>Serratia marcescens</i>	gi 127525 sp P02938 LPP_SE RMA,gi 146311423 ref YP_001176497.1 ,gi 152970683 ref YP_001335792.1	7 193,2	1	20,90%
Thiol peroxidase	<i>Escherichia Coli</i>	gi 1103833 gb AAC43517.1 ,gi 157155990 ref YP_001462632.1 ,gi 15801846 ref NP_287864.1 ,gi 42543455 pdb 1QXH A,gi 74312309 ref YP_310728.1	17 877,3	2	17,90%
Hypothetical protein BpseS_03000875	<i>Burkholderia pseudomallei</i> S13	gi 100125706 ref ZP_01331031.1 ,gi 113866733 ref YP_725222.1 ,gi 119897265 ref Y	56 942,1	2	7,88%

Chapitre III – Etude de la diversité des gènes *aoxB* dans la communauté bactérienne de l'ancienne mine de Sainte-Marie-aux-Mines

P_932478.1|,

DNA-directed RNA polymerase subunit beta'	<i>Hahella chejuensis</i> KCTC 2396	gi 83648858 ref YP_437293.1	154 503,6	2	1,22%
F0F1 ATP synthase subunit beta	<i>Planctomyces maris</i> DSM 8797	gi 149173275 ref ZP_01851906.1 ,gi 32476317 ref NP_869311.1 ,gi 457413 emb CAA40489.1 ,gi 87311750 ref ZP_01093865.1	52 729,3	2	4,97%
Chloroplast light harvesting protein isoform 15	<i>Isochrysis galbana</i>	gi 77024121 gb ABA55531.1	21 820,8	1	5,53%
Fucoxanthin chlorophyll a/c binding protein	<i>Cyclotella cryptica</i>	gi 2996369 emb CAA04400.1 ,gi 2996381 emb CAA04178.1 ,gi 2996383 emb CAA04226.1	21 184,1	2	10,60%
Fucoxanthin-chlorophyll a/c light-harvesting protein	<i>Skeletonema costatum</i>	gi 1778725 gb AAB70101.1 ,gi 1778741 gb AAB40912.1	20 383,2	1	6,91%
Photosystem II protein H	<i>Porphyra purpurea</i>	gi 11465791 ref NP_053935.1 ,gi 11467750 ref NP_050802.1 ,gi 149072040 ref YP_001293614.1 ,gi 51209909 ref YP_063573.1	7 408,3	1	20,90%
Photosystem II 44 kDa protein	<i>Euglena gracilis</i>	gi 11466985 ref NP_041892.1 ,gi 224089 prf 1010249B	14 958,0	1	24,20%
Chloroplast sedoheptulose-1,7-bisphosphatase	<i>Euglena gracilis</i>	gi 99903740 gb ABF68591.1	87 948,3	3	4,79%

LEGENDS

Figure 1: Gabe-Gottes mining site map in Sainte-Marie-aux-Mines (France) and location of sampling site.

Figure 2: Global phylogeny analyses from Gabe-Gottes sediments. Panel A: Bacterial phyla repartition obtained by cloning/sequencing of 260 clones (82 strains selected on As(V), 84 strains selected on As(III), both cultivated in aerobic conditions, 33 and 61 strains resistant to As(V) or As(III) isolated in anaerobic conditions, respectively). Panel B: Archaeal phyla repartition obtained by cloning/sequencing of 83 clones. In blue Crenarchaeota and in pink Euryarchaeota.

Figure 3: Panel A: Phylogenetic relationship of cultivable arsenic resistant bacteria isolated from Gabe-Gottes sediment related to taxa based on a combined sequence analysis of partial 16S rDNA genes (1500 nucleotids). Tree was generated using near full-length 16S rDNA gene sequences and the neighbor-joining method. The scale bar represents the numbers of substitution per 100 nucleotides within the 16S rDNA sequence. Panel B: Quantification of cultivable bacterial cell extracted from sediments

Figure 4: Neighbor-joining phylogenetic tree of bacterial AoxB sequences from isolated clones or metagenomic approach.

Supplementary Figure 1: Protein sequence alignment and putative secondary structure of AoxB. Sequence alignment were carried out by CLUSTALX. Amino acid numbers are defined according to *Alcaligenes faecalis* protein sequence. Residues known to play a role in As(III) oxidase activity are framed in black.

The supplementary table 1, because of its size, cannot be printed. It is available upon request to the authors.

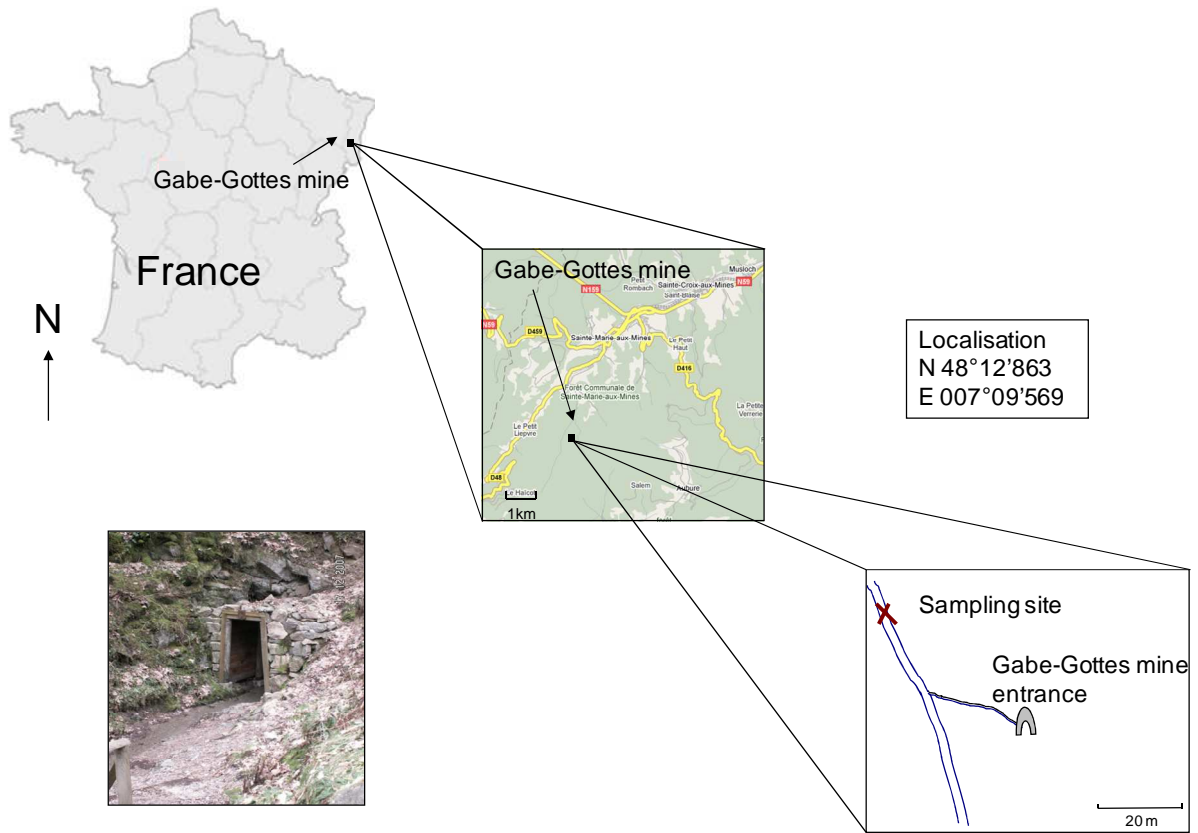
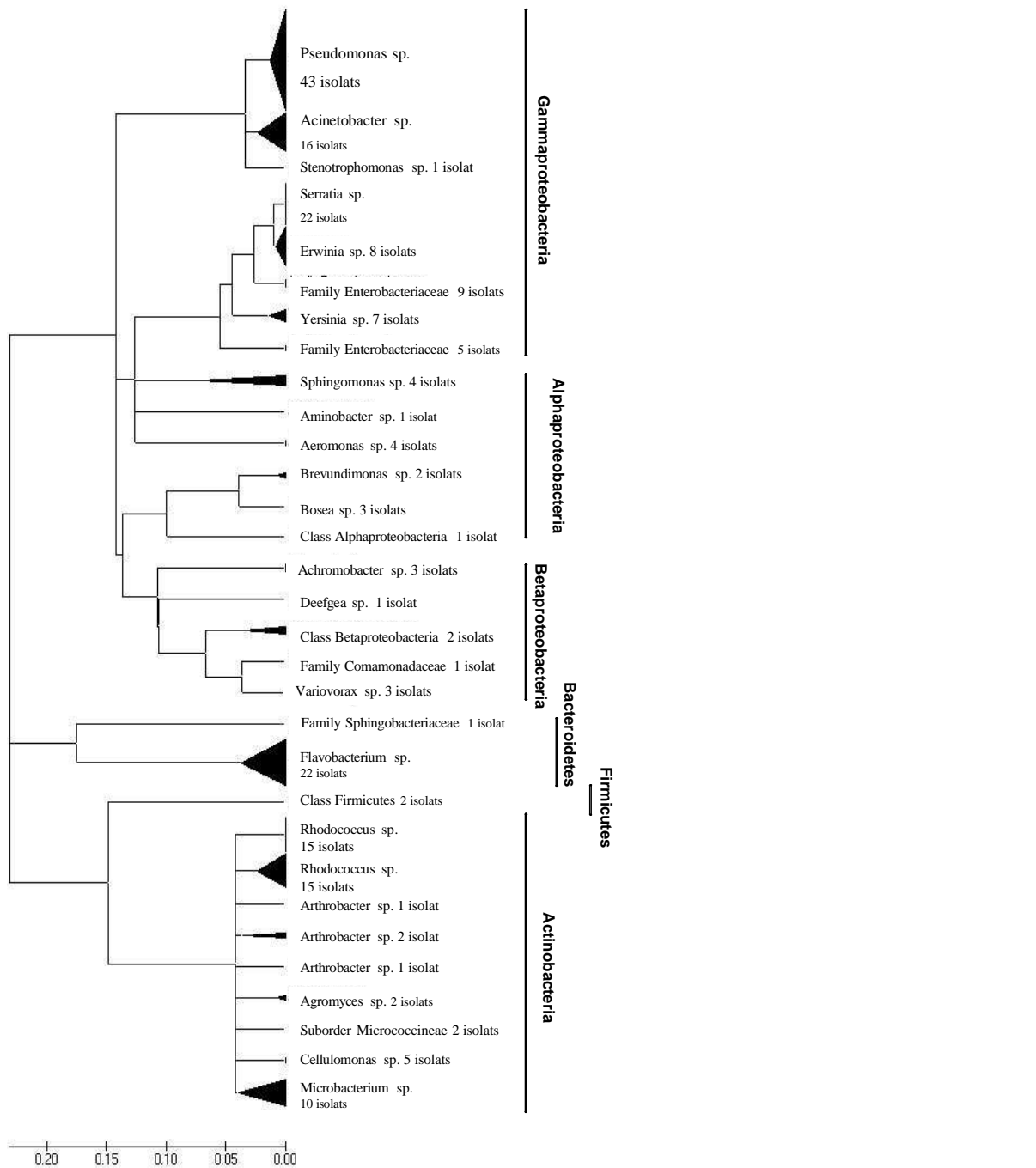


Figure 1



	10 ³ CFU/ g sediment ^a
Aerobic conditions	11479-13594
Aerobic conditions + AsIII (7mM)	271-61
Aerobic conditions + AsV (20mM)	6504-2683
Anaerobic conditions	ND 375
Anaerobic conditions + AsIII (7mM)	16-9
Anaerobic conditions + AsV (20mM)	168-232

^a depending culture conditions: Agar or Gellane gum

Figure 2

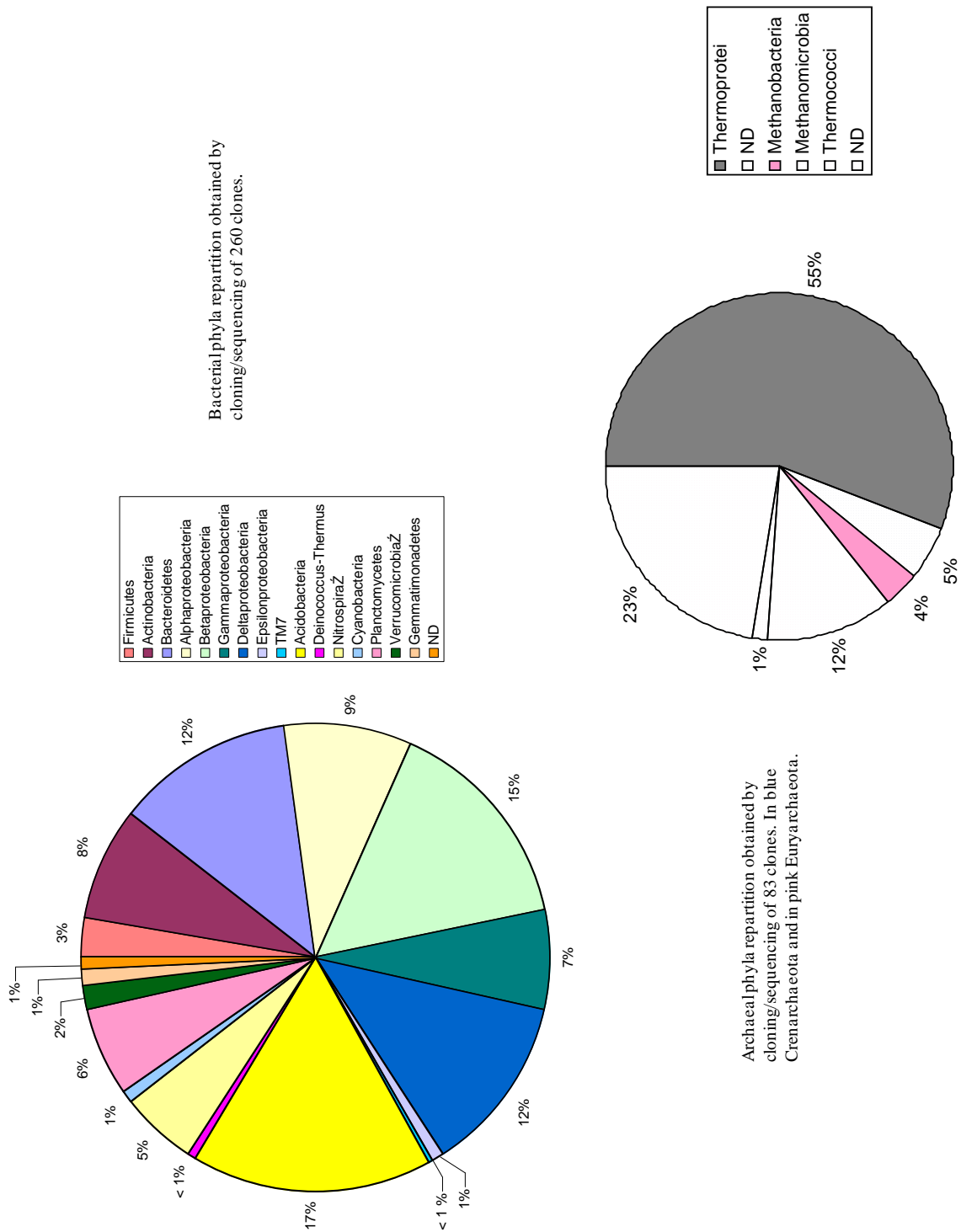


Figure 3

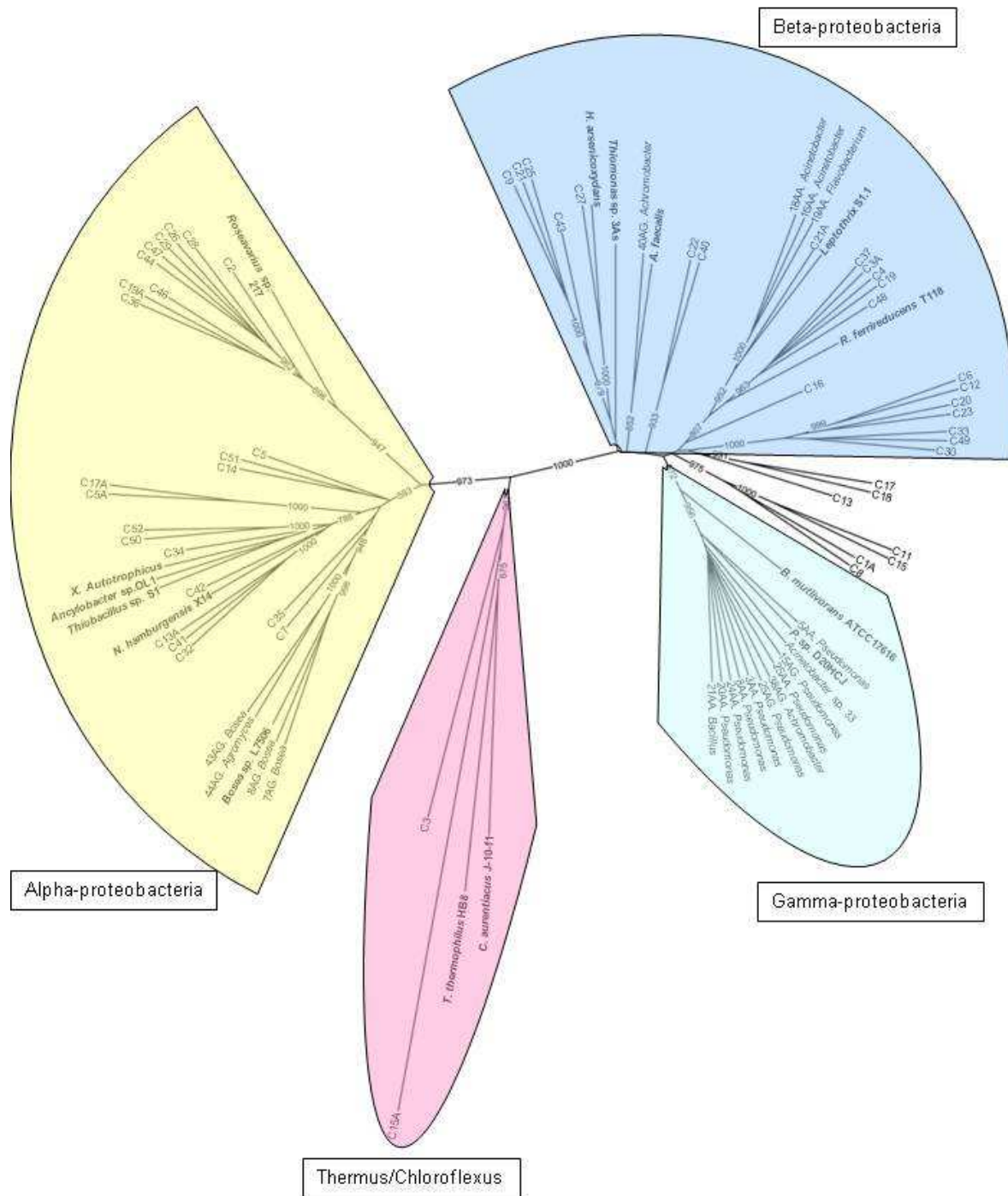


Figure 4

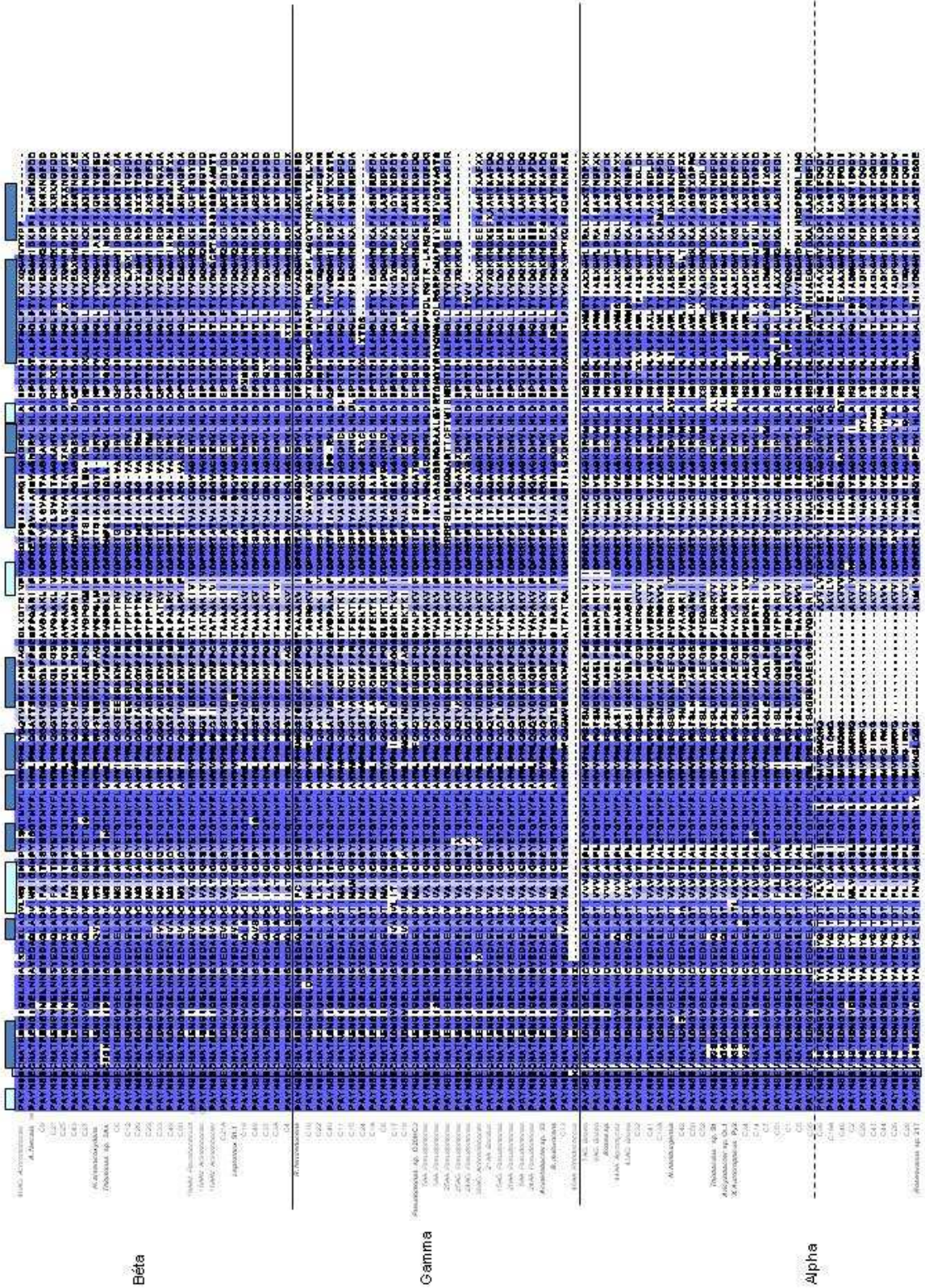


Figure S1

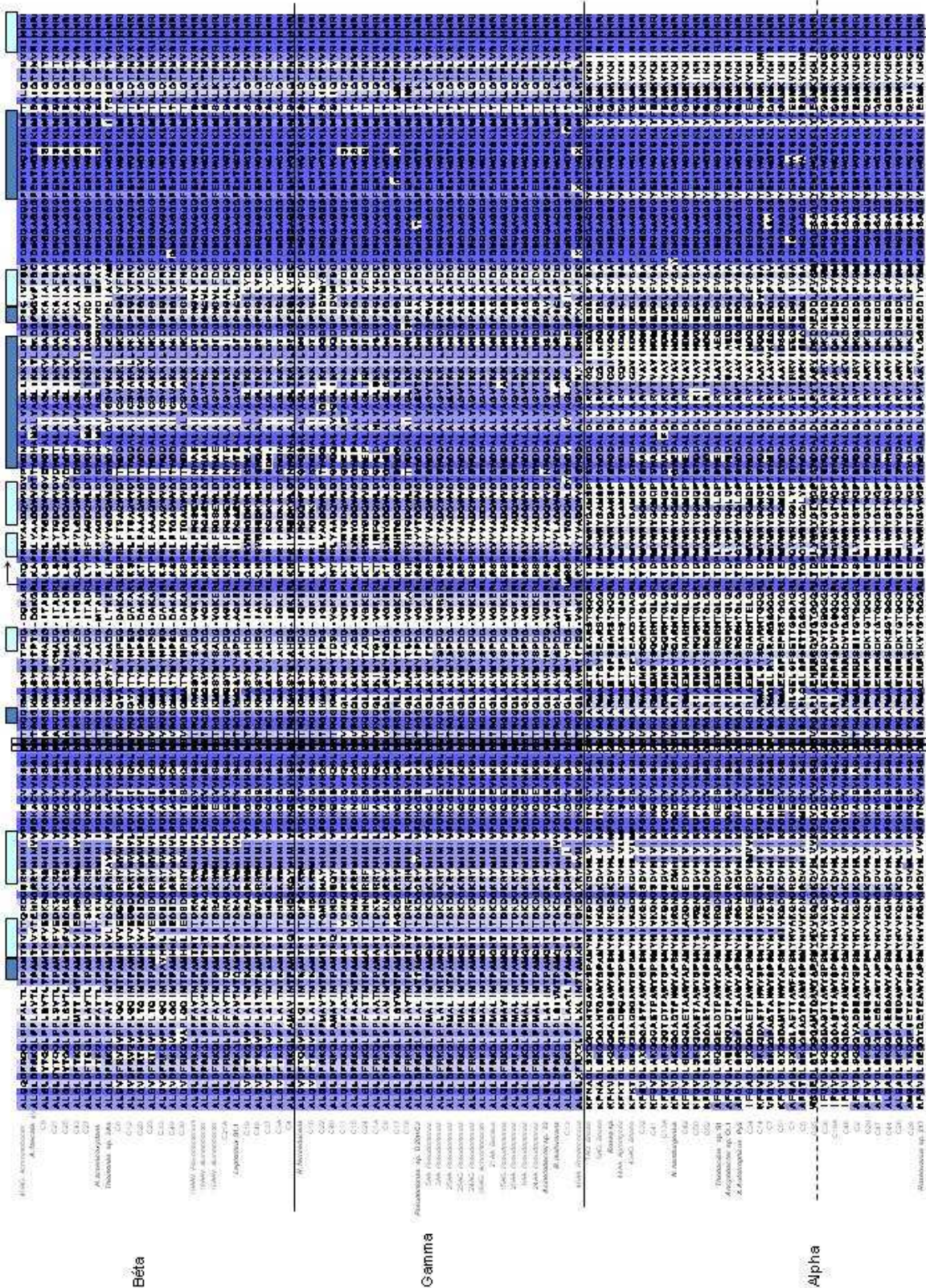


Figure S1 (suite)

Chapitre IV

Etude d'une communauté complexe
issue d'un AMD par des approches
protéo-génomiques

I INTRODUCTION

L’utilisation des approches de métagénomique permet aujourd’hui de s’affranchir des étapes de culture et d’élucider les questions concernant le rôle des communautés microbiennes dans leur globalité, y compris leurs interactions dans un écosystème complexe. La métagénomique couplée à une étude métaprotéomique permet de caractériser le potentiel génétique et métabolique des pseudo-génomés reconstruits issus des communautés microbiennes de cet environnement (Allen & Banfield, 2005) (Bertin *et al.*, 2008).

Ces approches, en s’affranchissant des biais et des *a priori* des études classiques, permettent de mettre en lumière des diversités bactériennes et génétiques inaccessibles jusqu’alors. Nous les avons, de ce fait, appliqués à l’étude de systèmes impliqués dans le métabolisme de l’arsenic.

Le site minier de Carnoulès fait l’objet de nombreuses études, en particulier concernant ses caractéristiques physico-chimiques (Morin *et al.*, 2003) (Casiot *et al.*, 2003a) (Leblanc *et al.*, 1996). Ces études permettent de le définir comme un remarquable exemple d’environnement hautement contaminé par l’arsenic. En effet, on retrouve dans le ruisseau de Reigous qui compose cet AMD, des concentrations d’arsenic parmi les plus importantes au monde, jusque 12 000 mg/l (Koffi *et al.*, 2002) (Bruneel *et al.*, 2006) et un pH équivalent à 3. Le Reigous est aussi un bon exemple d’atténuation naturelle, puisque la concentration en arsenic soluble chute de 95 % un kilomètre en aval de la source, cette diminution étant à attribuer à la précipitation de l’arsenic.

En collaboration avec le Génoscope, 550 000 séquences ont été obtenues, selon la méthode Sanger, sur l’ADN bactérien total des micro-organismes extraits de sédiments. Les séquences assemblées en sept pseudo-génomés correspondent aux génomes bactériens majoritaires de cet environnement, dont cinq à des souches non-cultivées. Nous avons d’autre part mis en évidence un nouveau phylum *Candidatus* Fodinabacter aminivorans particulièrement impliqué dans le recyclage de matière organique.

Une analyse globale des potentialités génétiques (voies métaboliques et énergétiques) des sept pseudo-génomés a été réalisée, ainsi que l’annotation experte des gènes appartenant aux différentes voies métaboliques d’intérêt impliquées dans l’adaptation et la colonisation de cet AMD par ces bactéries.

L’analyse descriptive et fonctionnelle de l’écosystème a mis en exergue l’organisation coopérative de ces sept souches, c’est-à-dire le rôle de chaque pseudo-génome dans le métabolisme arsénié, mais aussi dans la fixation du carbone inorganique et de l’azote, le recyclage du fer, du soufre et de matière organique. Enfin, nous avons identifié les fonctions permettant aux bactéries de

décontaminer ce site en précipitant l'As V et le FeIII en oxy-hydroxides (par exemple le tooeleite - précipité d'As V, de Fe III et de SO_4^{2-}) grâce à l'oxydation de l'arsénite en arséniate.

Cette étude a été réalisée dans le cadre d'un projet financé par l'Agence Nationale de la Recherche (ANR), intitulé « Reactivity of an Arsenic-Rich Ecosystem » (RARE), du Groupement de Recherche (GDR) 2909 « Métabolisme de l'arsenic chez les micro-organismes : de la résistance à la détoxification », et en collaboration avec 7 laboratoires du réseau. Ma collaboration dans ce projet a porté sur quatre points :

- L'analyse globale des potentialités des sept pseudo-génomomes et de leur interactions au sein de l'écosystème.
- L'annotation experte de certaines voies métaboliques.
- L'intégration des données métaprotéomiques aux analyses métagénomiques.
- L'écriture du manuscrit et la réalisation des figures.

L'ensemble de ces analyses a donné lieu à une publication Bertin *et al.*, 2009 présentée ci-après.

II MANUSCRIT

Diversity of trophic interactions inside an arsenic-rich microbial ecosystem

Philippe N. Bertin, **Audrey Heinrich-Salmeron**, Eric Pelletier, Florence Goulhen-Chollet, Florence Arsène-Ploetze, Sébastien Gallien, Alexandra Calteau, David Vallenet, Corinne Casiot, Béatrice Chane-Woon-Ming, Ludovic Giloteaux, Mohamed Barakat, Violaine Bonnefoy, Odile Bruneel, Michael Chandler, Jessica Cleiss-Arnold, Robert Duran, Françoise Elbaz-Poulichet, Nuria Fonknechten, David Halter, Béatrice Lauga, Marie Marchal, Damien Mornico, Philippe Ortet, Eduardo P. C. Rocha, Patricia Siguier, Emmanuel Talla, Adam Alexander Thil Smith, Sara Vieira-Silva, Alain Van Dorsselaer, Jean Weissenbach, Claudine Médigue, and Denis Le Paslier.

Soumis à Science

Diversity of trophic interactions inside an arsenic-rich microbial ecosystem

Philippe N. Bertin¹, Audrey Heinrich-Salmeron¹, Eric Pelletier^{2,3,4}, Florence Goulhen-Chollet¹, Florence Arsène-Ploetze¹, Sébastien Gallien⁶, Alexandra Calteau^{3,4,5}, David Vallenet^{3,4,5}, Corinne Casiot⁷, Béatrice Chane-Woon-Ming^{3,4,5}, Ludovic Giloteaux⁸, Mohamed Barakat⁹, Violaine Bonnefoy¹⁰, Odile Bruneel⁷, Michael Chandler¹¹, Jessica Cleiss-Arnold¹, Robert Duran⁸, Françoise Elbaz-Poulichet⁷, Nuria Fonknechten^{2,3,4}, David Halter¹, Béatrice Lauga⁸, Marie Marchal¹, Damien Mornico^{3,4,5}, Philippe Ortet⁹, Eduardo P. C. Rocha¹², Christine Schaeffer⁶, Patricia Siguier¹¹, Emmanuel Talla¹⁰, Adam Alexander Thil Smith^{3,4,5}, Sara Vieira-Silva¹², Alain Van Dorsselaer⁶, Jean Weissenbach^{2,3,4}, Claudine Médigue^{3,4,5}, Denis Le Paslier^{2,3,4}.

1 Génétique Moléculaire, Génomique et Microbiologie, UMR7156 CNRS & Uds, Strasbourg, France, **2** CEA, DSV, IG, Genoscope, Laboratoire de Métagénomique des Procaryotes, **3** CNRS UMR8030, Evry, France, **4** UEVE, Université d'Evry, **5** CEA, DSV, IG, Genoscope, Laboratoire de Génomique Comparative, **6** Laboratoire de Spectrométrie de Masse Bio-Organique, Institut Pluridisciplinaire Hubert Curien, UMR7178 CNRS & Uds, Strasbourg, France, **7** Laboratoire Hydrosociences Montpellier, UMR 5569, Montpellier, France, **8** Environnement et Microbiologie, UMR5254 CNRS & UPPA, Institut Pluridisciplinaire de Recherche sur l'Environnement et les Matériaux, Pau, France, **9** Laboratoire d'Ecologie Microbienne de la Rhizosphère et d'Environnements Extrêmes, UMR6191 CNRS, CEA & Université Aix-Marseille II, Saint-Paul-lez-Durance, France, **10** Laboratoire de Chimie Bactérienne, UPR9043 CNRS et Université de la Méditerranée, Marseille, France, **11** Laboratoire de Microbiologie et Génétique Moléculaires, UMR5100 CNRS, Toulouse, France, **12** Institut Pasteur, Microbial Evolutionary Genomics, CNRS, URA2171, Paris, France

Running title: Deciphering the Carnoulès community

*To whom correspondence should be addressed. E-mail: philippe.bertin@unistra.fr

The understanding of nutrient cycles requires the deciphering of both the structure and the function of microbial communities, which has been hampered until recently by a lack of genomic data on uncultivated microorganisms. We used here an environmental genomics approach to reconstruct from an environment highly contaminated with arsenic and toxic metals the genomes of seven dominant microbial strains. Most of them represent yet uncultivated bacteria, including microorganisms belonging to a novel bacterial phylum. In depth analysis of metagenomic data combined with a functional metaproteomic approach unravelled some unsuspected capabilities of those microorganisms, in particular regarding sulfur, iron or arsenic oxidation, and organic compound or cofactor metabolism. From these results, we built an integrated model of the microbial interactions inside the arsenic-rich ecosystem, leading to an efficient natural remediation.

Microorganisms are involved in biogeochemical nutrient cycles and play therefore a crucial role in the biosphere. Although microbial ecosystems constitute a major reservoir of our planet's genetic biodiversity, it is generally recognized that most of the microorganisms present in the environment are not accessible by culture-dependent techniques. However, the recent development of genomics has given an unprecedented opportunity to gain insight into the structure and the function of microbial communities. Indeed, environmental genomics has extended their analysis far beyond the sole taxonomic studies, giving rise to an integrated picture of ecosystems, including those considered as extreme (1-3).

Arsenic, in the As(III) form (arsenite), may have been an important mineral substrates providing chemolithotrophic organisms with energy in the environments harbouring the first forms of life on earth (4). Nowadays, high arsenic contents are encountered in mine drainage waters, where the toxic metalloid is usually associated with sulfur, iron and other metals (5). Few forms of life thrive in these acid waters, as compared with neutral waters (6).

The former mine of Carnoulès, Gard (France), provides an outstanding example of an extreme environment. The sulfurous wastes contain As-rich pyrite and the leached waters are the source of a small stream called the Reigous that usually contains between 50 and 350 mg.l⁻¹ of soluble arsenic, mainly in the form of As(III) (table S1). However, although the arsenic levels remain still high, this concentration decreases by 95% between the source of the Reigous and its confluence with the river Amous, 1.5 km downstream. This natural process of remediation is mainly due to microbial metabolism, leading to the oxidation of iron and arsenic into Fe(III) and As(V), and their co-precipitation of with sulfur (7-9). Despite many recent works, few of these microbes have been isolated (9-11). In the present study, we used a multidisciplinary approach to get an integrated view of the microorganisms present in Carnoulès and their role in the complex metabolic processes at work.

DNA recovered from the upper zone of sediments was sequenced and assembled in 7 major bins (12), using a combination of 16S rRNA gene sequences, GC%, mean coverage of the various contigs, mean polymorphism frequency and similarity to already sequenced genomes, including *Acidithiobacillus ferrooxidans* and *Thiomonas* sp. 3As (<https://www.genoscope.cns.fr/agc/mage/arsenoscope>). Even though the assembling of sequences from closely related organisms cannot be excluded, each supercontig group was considered to correspond to the genome of one single major organism (table S2, and figs. S1 and S2). We evaluated the origin of bins by a phylogenetic analysis of 16S rRNA sequences obtained from both a clone library and metagenome sequencing (Fig. 1). We also made a phylogenomic analysis of the 7 bins using 27 universal marker genes (individually or combined) supposed to be sufficiently conserved to

build a tree of life (13). Sequence analyses revealed a dominance of bacteria belonging to the Proteobacteria phylum (β - and γ -Proteobacteria with 38 and 23%, respectively). Remaining clones belong to Firmicutes (4%) and Acidobacteria (3%), and to a lesser extent to Spirochaetes (1%) and α -Proteobacteria (1%). This suggests that the microbial community at Carnoulès includes a few other bacterial phyla in addition to those represented by the 7 major genomes, and few Archaea (7). More precisely, and except for CARN4 whose missing rRNA sequence precluded its 16S-based classification, and CARN6 where no marker was found, the candidates CARN2, CARN3 and CARN5 were affiliated to the β -Proteobacterium *Thiomonas* genus, the Acidobacteria clade and to the γ -Proteobacterium *At. ferrooxidans*, respectively (table S3). In addition, the comparison of CARN7 16S rRNA gene with sequences in the RDP database suggests that this organism is phylogenetically related to the *Gallionella* β -Proteobacterium. All phylogenetic results for the bins were supported, at least in part, by an analysis of gene order conservation known to be correlated with evolutive distances (14). Remarkably, the candidates corresponding to the CARN1 and CARN4 bins showed an important phylogenetic relationship and may represent 2 subpopulations (figs. S1 et S2), as recently observed in other contexts (3). Those bins did not correspond to any known taxonomic class and, according to the metabolic properties identified in the present study (see below), this new genus was herein named *Candidatus Fodinabacter communicans*, in accordance with the recommendations for incompletely characterized microorganisms (15). Interestingly, similar 16S rRNA sequences have already been identified in acid mine environments, suggesting that these uncultivated bacteria are widespread in such ecosystems (data not shown).

The size of the 7 bins ranged from 1.5 Mbp to more than 4.0 Mbp and their GC content ranged approximately from 52 to 65% (table S4). DNA sequencing led to a low coverage for some contigs, in particular those of CARN3, which suggests that the size of the corresponding genome is underestimated. In addition, the comparison of the CARN5 data with those of the 2 already available *At. ferrooxidans* genomes revealed the existence of at least one large gap in CARN5 and ATCC 53993, as compared to ATCC 23270. Similar results were obtained by 454 pyrosequencing, which supports the lack of any major bias resulting from the cloning procedure for genomic library construction (fig. S3). In ATCC 23270, this region contains a duplication of most tRNA encoding genes, several genes involved in the synthesis of membrane sugars and multiple transposases suggesting a probable acquisition by horizontal gene transfer.

We investigated the fitness of the bacterial community, in particular regarding arsenic, by exploring *in silico* the gene content of each bin. In parallel, we analyzed the proteome synthesized *in situ*, which allowed the reliable identification of more than 500 unique proteins (tables S5a and S5b). Arsenic toxicity results from various biological effects, including oxidative stress and DNA damage. As

expected, both metagenomic and/or metaproteomic data showed that all Carnoulès strains can express functions protecting against oxidative stress, including those coding for catalase, superoxide dismutase, glutathion peroxidase or thioredoxin/thioredoxin reductase. Most bins also contain genes coding for the bacterioferritin comigratory protein, which protects cells against toxic hydroxyl radicals. Multiple genes coding for enzymes involved in DNA recombination and repair were also identified and were recently shown to be important for viability in the presence of arsenic (16). In addition, the genome of each bin carries at least one *ars* operon conferring resistance to arsenic by the means of an ArsB efflux pump (Fig. 2). Surprisingly, the 3 operons identified in *Thiomonas* sp. CARN2 were located in the vicinity of 3 *aoxAB* operons, suggesting the existence of 3 distinct arsenic islands in this strain (fig. S5). One of those islands was flanked by multiple transposases (table S6), suggesting a possible acquisition by horizontal gene transfer. Arsenic methylation may also be considered as a resistance mechanism, and *arsM* genes coding for an arsenite S-adenosylmethyltransferase were identified, in particular in the CARN6 bin. This observation is in agreement with the presence of methylated forms of arsenic, i.e. monomethylarsonate and dimethylarsinate, in Carnoulès sediments (C. Casiot, unpublished). Finally, because of the structural similarity between P and As(V), arsenic metabolizing strains may preferentially transport phosphate via the specific Pst phosphate transport system rather than the Pit general transport mechanism to reduce the entry of As(V) (16, 17). No Pit gene was identified in any bin while all strains contain a Pst system (table S5b), further supporting this hypothesis.

Due to their role in biofilm formation, motility and aggregation have been often demonstrated to be major determinants for survival in toxic environments, including arsenic (16). Except for *Acidithiobacillus* sp. CARN5, most bins contain multiple genes required for flagellum biosynthesis and motility (Fig. 2). Moreover, *Thiomonas* sp., *Acidithiobacillus* sp. and *Gallionella*-related strains also carry Type 4 pilus operons, which play an important role in twitching motility and adhesion. Finally, *Acidithiobacillus* sp. CARN5 contains several genes involved in capsular biosynthesis, e.g. *kdtA*, *crd* and *gmd*, suggesting the possible formation of a biofilm. Such adaptive traits typically depend on multiple regulatory mechanisms that allow bacteria to respond to a wide panel of stimuli, including high concentrations of iron and arsenic. In this respect, all bins encode nucleoid-associated proteins such as HU, IHF, FIS or H-NS (18), which play a major role in both the structure and the function of chromosomal DNA. Multiple regulators, e.g. various alternative sigma factors and regulatory proteins belonging to the two-component systems or those involved in cell-to-cell signaling, were also identified in all strains (tables S7a and S7b).

A wide diversity of bioenergetic electron chains may be needed to accommodate the fluctuations in the various physico-chemical parameters. Accordingly, several terminal oxidases might

be operative in all strains of the Carnoulès microorganisms, e.g. the cytochrome oxidase *cox* and *cta* operons and the *cyo* and *cyd* operons encoding quinol oxidases. In addition, several operons involved in anaerobic respiration were present in the bin of *Thiomonas* sp. (CARN2), in particular those involved in the metabolism of nitrate and nitrite, such as *ntr*, *nar*, and *nas*, or the use of fumarate as an electron acceptor under anaerobic conditions (Fig. 2). More importantly, possible inorganic electron donors may comprise sulfur compounds and arsenite. This was attested by the synthesis by *Thiomonas* sp. of proteins such as *sor*, *sox*, *tetH* and *sqr* as well as *aox* involved in the oxidation of reduced inorganic sulfur compounds, e.g. sulfide, sulfur, sulfite, thiosulfate and tetrathionate, and As(III), respectively. Similarly, the CARN7 bin harbours a *dsr*ABEFHCMKLJOPN operon. The DsrA protein showed 72% similarity with ortholog in the sulfur-oxidising *Thiobacillus denitrificans*. This suggests that the *Gallionella*-related strain is also able to use sulfur compounds as electron donor in its energy metabolism, as suggested in *Gallionella ferruginea* (19). In addition, the *Acidithiobacillus* sp. CARN5 bin expressed the sole RusA protein (table S5b). This rusticyanin is involved in electron transport with iron used as an energy source, which suggests a major role of CARN5 in the iron oxidation observed on the study site (9, 20). All these metabolisms may play a key role in the natural bioremediation observed at Carnoulès. Finally, the oxidation of H₂ to protons may also be a source of energy due to the presence in CARN2, CARN5 and CARN6 of several hydrogenase-encoding genes. In contrast, no respiratory arsenate reductase, which allows anaerobic respiration of As(V), was identified in any of the 7 bins.

The existence of metabolic exchanges regarding carbon and nitrogen may be of prime importance inside a microbial community. As expected from a partly oligotrophic environment and suggested by the predicted slow growth rate of each bin (table S8), we identified several enzymes responsible for autotrophic metabolism (Fig. 2). Carbon fixation may depend on genes coding for ribulose 1,5-biphosphate carboxylase/oxygenase (involved in Calvin cycle), carboxysome structural proteins and carbon monoxide dehydrogenase (involved in acetyl-coenzyme A synthesis) that were identified in and expressed by *Thiomonas* sp. (CARN2), *Acidithiobacillus* sp. (CARN5) or *Gallionella*-related (CARN7) strains. On the other hand, CARN5 and CARN7 were also able to fix nitrogen by the means of *nif* genes, which encode a nitrogenase. Nevertheless, mixotrophic or heterotrophic metabolism may also exist in Carnoulès, suggesting that microorganisms recycle some organic compounds released by others. For example, CARN2, CARN5 and CARN7 encode lactate, formate or succinate dehydrogenases, while *fucP* and *exuT* genes coding for L-fucose and hexuronate transporters, respectively are present in CARN3. In addition, the CARN6 bin carries the cellulase encoding gene *bczS* and the alpha-amylase *amyM* gene, suggesting that this strain is able to metabolize more complex carbohydrates. Remarkably, the CARN1 and CARN4 bins of *Candidatus* F.

communicans carry the arginase-encoding gene *rocF* involved in urea biosynthesis, while a complete urease-encoding *ure* operon was present in *Thiomonas* sp. CARN2 strain. This observation is in agreement with the identification of urease activity in various *Thiomonas* strains (F. Arsène-Ploetze, unpublished) and further supports a metabolic interaction between the Carnoulès microorganisms. Similarly, several enzymes required for amino acid transport and metabolism were identified in the bins lacking the carbon and nitrogen fixation genetic determinants. In particular, the CARN1 and CARN4 bins of *Candidatus* F. communicans, which were shown to express most of the identified proteins (table S5b), encode multiple peptidases and also contain *liv* and *opp* genes involved in branched amino-acid and oligopeptide transport, respectively. Genes involved in fermentation were also identified in both bins, e.g. those of the lysine pathway which converts lysine into butyrate, acetate and ammonia, principally in obligate anaerobic bacteria (21). Finally, the CARN1 and CARN4 bins carry genes involved in cobalamin (vitamin B12) synthesis (22) while cobalamin transporters are present in the genome of CARN2 and CARN5, as observed in already sequenced related bacteria (<https://www.genoscope.cns.fr/agc/mage/arsenoscope>). This biosynthesis may also be useful to eucaryotes, in particular *Euglena* sp. present on the study site (23) whose cell cycle is known to require cobalamin (24, 25). Remarkably, iron oxidation by *At. ferrooxidans* was shown to be strongly increased in the presence of cobalamin (fig. S6), which suggests that such a metabolic cooperation may be of importance in the natural remediation observed at Carnoulès.

These observations were further supported by a Multiple Factor Analysis (MFA) of both metagenomic and metaproteomic data on a reaction frequency matrix, suggesting an important and specific role for each organism inside the Carnoulès community. Indeed, factorial planes segregated CARN2 from the other bins (Fig. 3a) and separated CARN1-CARN4 from CARN5-CARN7. From the variable classification results, several clusters of reactions were then associated to bin groups. These include, for example, Cluster 4 linked to CARN2 (energy metabolism, inorganic nutrient metabolism and arsenic detoxification (table S9a); Cluster 1 groups reactions common to CARN2 and CARN5 (Calvin-Beson-Bassham cycle and urea degradation pathways); Cluster 5 includes reactions related to lysine fermentation and other amino acid degradation pathways in CARN1 and CARN4, and also gathers reactions related to cobalamin biosynthesis. In addition, CARN3 was opposed to CARN6 (Fig. 3b), revealing CARN6-specific reactions (Clusters 5 and 6) that include those involved in purine degradation and cellulose metabolism (table S9b). Interestingly, no such correlation was observed between our data and those from the acidic AMD biofilm (26), which suggests that the diversity of these two ecosystems differ markedly and highlights the metabolic complementarity inside the Carnoulès community (fig. S7).

Taken together, these environmental genomics approaches allowed us to inventory the multiple metabolic pathways in the ecosystem under study and to visualize several activities expressed *in situ* (Fig. 2 and table S5b). More importantly, our observations revealed the diversity of trophic interactions that may exist inside a microbial community and allow microorganisms, including yet uncultivated bacteria, to efficiently colonize an arsenic-rich environment. These processes include not only the fixation of inorganic carbon and nitrogen but also the use or the recycling of other mineral and organic resources such as arsenic, iron, sulfur, urea, vitamins and amino acids. Such an equilibrium between organic and inorganic compound metabolism may provide all partners with nutrients. In this respect, the role of *Thiomonas* sp. in arsenic oxidation associated with other metabolic activities, including sulfur oxidation by the *Gallionella*-related strain and iron oxidation by *Acidithiobacillus* sp., seems to be of prime importance in the partial but efficient natural remediation of the Carnoulès contaminated ecosystem. Other microorganisms, in particular *Candidatus* F. communificans, may recycle organic compounds released by others or provide them with essential cofactors. Our observations therefore support an indirect but crucial role of this first representative of a novel bacterial phylum in the bioremediation process of the arsenic-rich ecosystem. In the future, similar strategies will lead to important advances in microbial ecology, revealing what was recently regarded as impossible to explore.

References and Notes

1. E. E. Allen, J. F. Banfield, *Nat. Rev. Microbiol.* **3**, 489 (2005).
2. P. N. Bertin, C. Médigue, P. Normand, *Microbiology* **154**, 347 (2008).
3. P. Wilmes, S. L. Simmons, V. J. Denef, J. F. Banfield, *FEMS Microbiol. Rev.* **33**, 109 (2009).
4. E. Lebrun *et al.*, *Mol. Biol. Evol.* **20**, 686 (2003).
5. D. J. Vaughan, *Elements* **2**, 71 (2006).
6. D. B. Johnson, K. B. Hallberg, *Res. Microbiol.* **154**, 466 (2003).
7. O. Bruneel *et al.*, *Extremophiles* **12**, 563 (2008).
8. C. Casiot *et al.*, *Water Res.* **37**, 2929 (2003).
9. K. Duquesne *et al.*, *Appl. Env. Microbiol.* **69**, 6165 (2003).
10. K. Duquesne *et al.*, *Environ. Microbiol.* **10**, 228 (2008).
11. O. Bruneel *et al.*, *J. Appl. Microbiol.* **95**, 492 (2003).
12. Materials and methods are available as supporting material on Science Online
13. F. D. Ciccarelli *et al.*, *Science* **311**, 1283 (2006).
14. M. A. Huynen, P. Bork, *Proc. Natl. Acad. Sci. U S A.* **95**, 5849 (1998).
15. R. G. E. Murray, E. Stackebrandt, *Int. J. Syst. Bacteriol.* **45**, 186 (1995).
16. D. Muller *et al.*, *PLoS Genet.* **3**, e53: 518 (2007).
17. S. Weiss *et al.*, *Biochimie* **91**, 192 (2009).
18. C. Tendeng, P. N. Bertin, *Trends Microbiol.* **11**, 511 (2003).

19. S. Lutters-Czekalla, *Arch. Microbiol.* **154**, 417 (1990).
20. G. Morin *et al.*, *Environ. Sci. Technol.* **37**, 1705 (2003).
21. A. Kreimeyer *et al.*, *J. Biol. Chem.* **282**, 7191 (2007).
22. J. C. Escalante-Semerena, *J. Bacteriol.* **189**, 4555 (2007).
23. C. Casiot, O. Bruneel, J. C. Personné, M. Leblanc, F. Elbaz-Poulichet, *Sci. Total Environ.* **320**, 259 (2004).
24. M. M. Olaveson, P. M. Stokes, *J. Phycol.* **25**, 529 (1989).
25. M. H. Bré, E. El Ferjani, M. Lefort-Tran, *Protoplasma* **108**, 301 (1981).
26. G. W. Tyson *et al.*, *Nature* **428**, 37 (2004).
27. A. H.-S. was supported by a grant from the French Ministry of Education and Research and F. G.-C. by a grant from the Agence Nationale de la Recherche (ANR). S. V.-S. is part of the Gulbenkian PhD Program in Computational Biology. Financial support came from the Université de Strasbourg (UdS), the Consortium National de Recherche en Génomique (CNRG), the Centre National de la Recherche Scientifique (CNRS) and the Agence Nationale de la Recherche (ANR) in the frame of the RARE and MicroScope projects. ISfinder is supported by the CNRS and has received some support from the Association pour la Recherche sur le Cancer (ARC). This work was done in the frame of the « Groupement de Recherche - Métabolisme de l'Arsenic chez les Microorganismes (GDR2909-CNRS) » (<http://gdr2909.u-strasbg.fr>).

Figure legends

Fig. 1. Phylogenetic tree representing the taxonomic affiliation of the Carnoulès community microorganisms. The 16S rRNA gene sequences were obtained from DNA sediments after PCR amplification (clones CG-X) or metagenomic sequencing (CARN bins, except for CARN4, see text). A total of 759 positions in the final dataset were obtained after all positions containing gaps and missing data were eliminated. The distances were computed using the Maximum Composite Likelihood method and the tree was inferred using the Neighbor-Joining method ; a Maximum-likelihood method gave rise to a similar phylogenetic reconstruction (data not shown). The scale bar corresponds to 0.05 substitutions per site. Percentages of 1000 bootstrap resamplings that supported the branching orders in each analysis are shown above or near the relevant nodes. Bootstrap values are shown for branches with more than 50% bootstrap support.

Fig. 2. Model of the Carnoulès bacterial community highlighting the major functions identified by metagenome sequencing or metaproteome characterization. These activities include carbon and nitrogen fixation, energy metabolism, flagellum and capsule biosynthesis, amino acid transport and degradation, detoxification and stress response, arsenic and iron metabolism. The possible interactions between these microorganisms or with other chemical or biological components present on the study site are indicated by arrows. CARN bins are numbered from 1 to 7.

Fig. 3. Multiple Factorial Analysis of the 7 Carnoulès bins, performed on a two dimensional-matrix combining bins and enzymatic reactions, respectively. To highlight possible metabolic distinctions between bins, 3 axes (F1 to F3) capturing the highest dispersion in the resulting dot cloud were selected; they represent more than half the total dispersion. Colored lines represent the vectors corresponding to the enzymatic reaction frequencies, the concentric disks differentiating those identified in the metaproteomic data. The reaction vectors were then hierarchically clustered, which led to 7 and 9 classes (indicated by the colors) for the first and second factorial planes, respectively. The corresponding functions are listed in table S9ab.

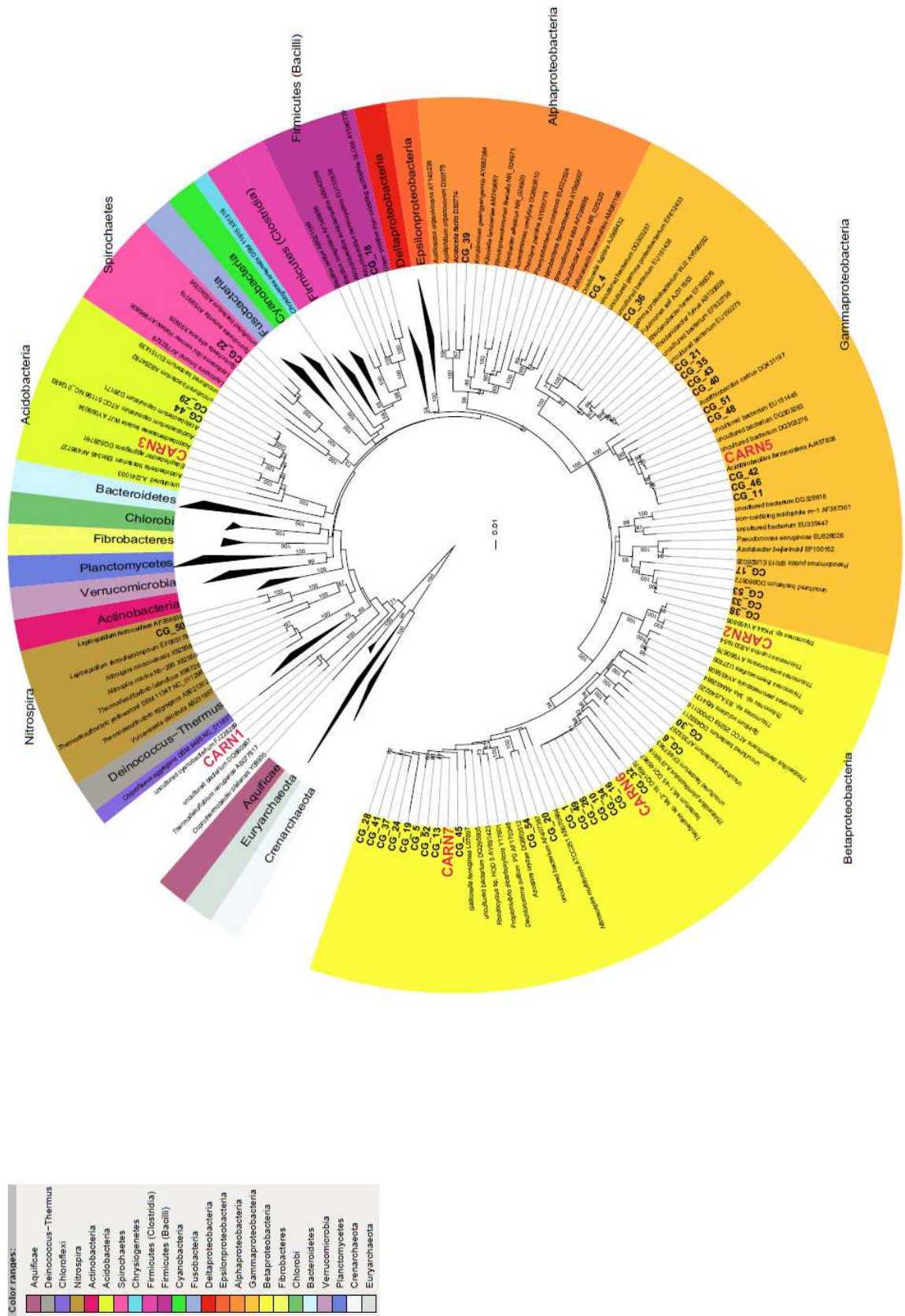


Figure 1

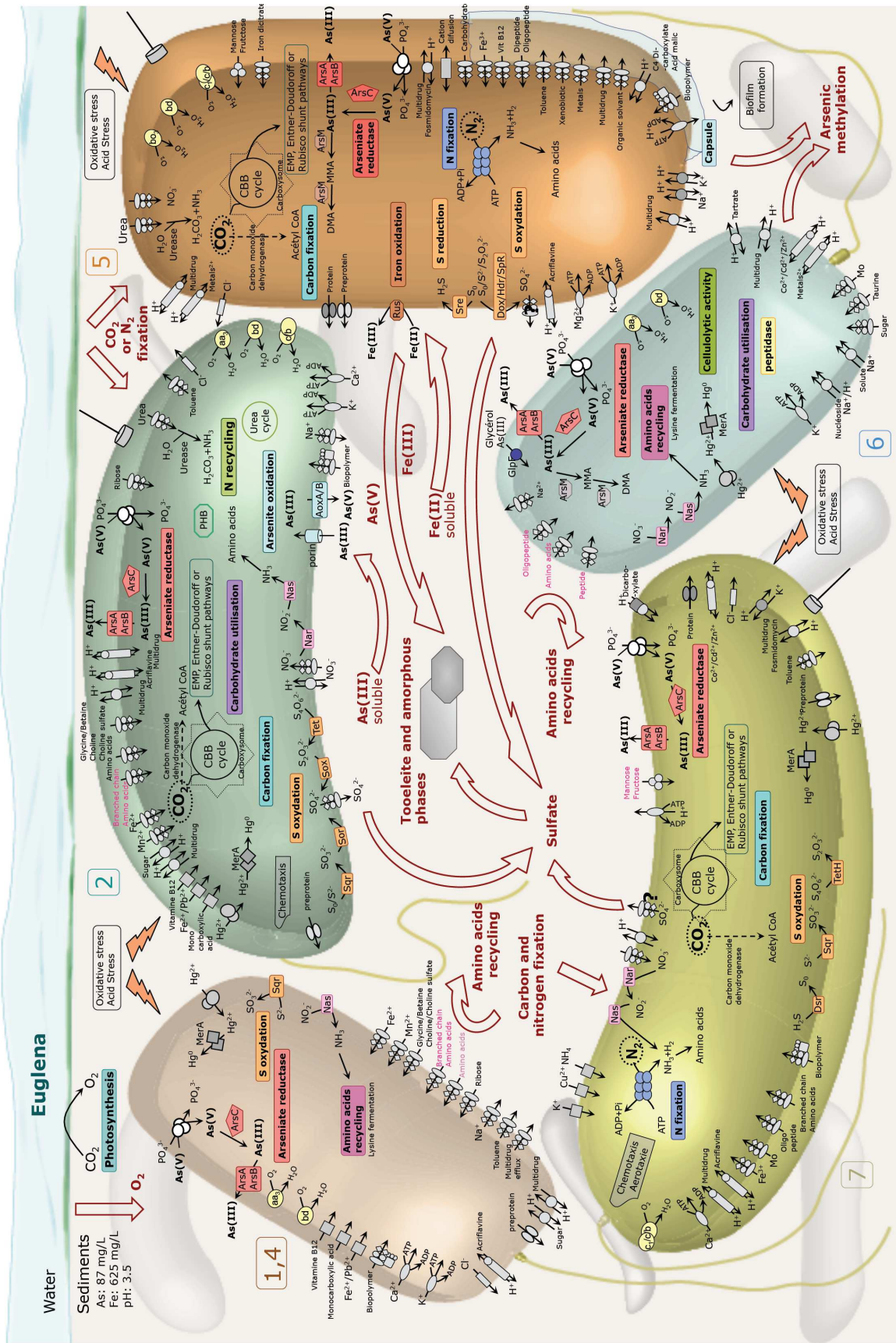


Figure 2

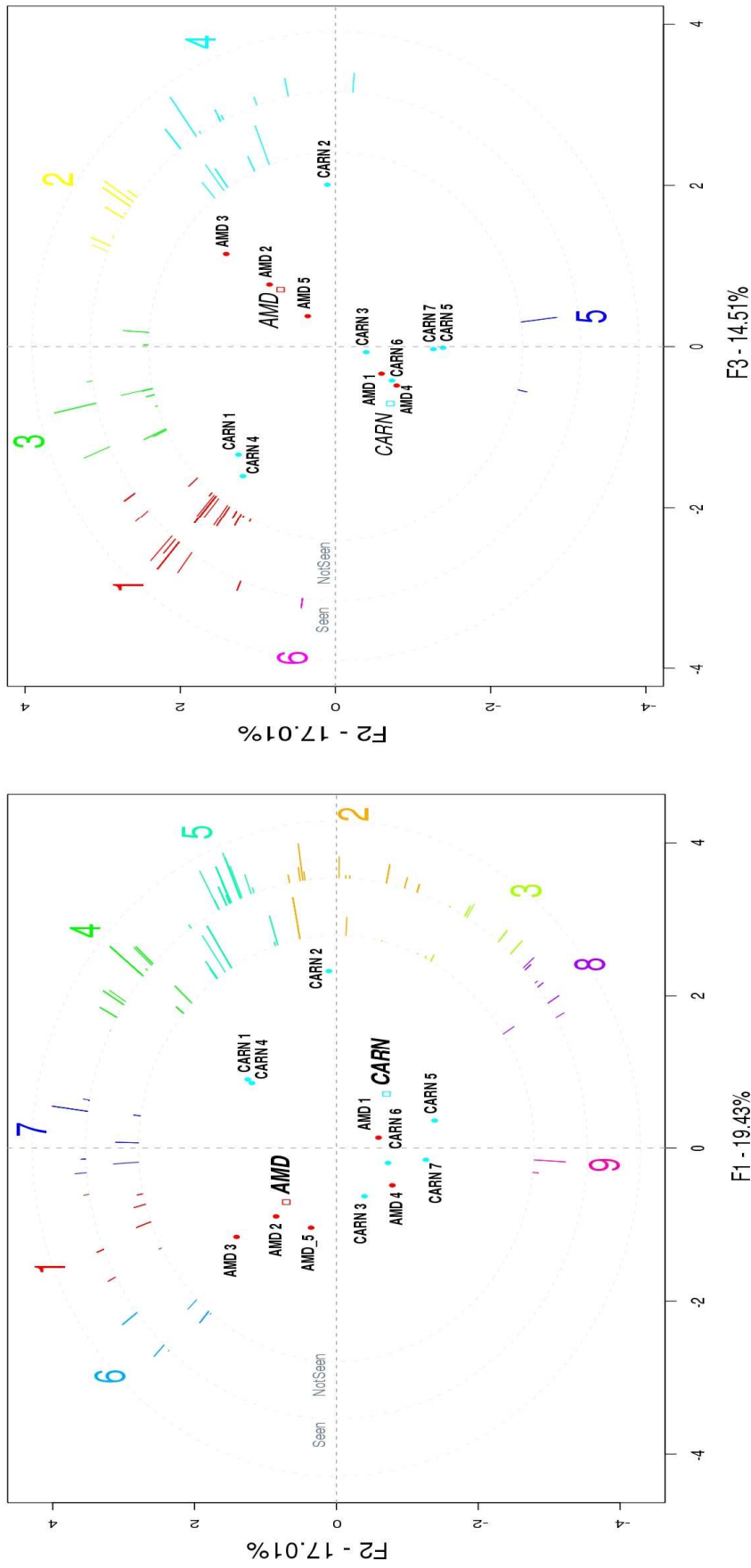


Figure 3

SUPPLEMENTARY DATA

Author contributions

P.B. coordinated the project, contributed to the analysis of genome information and wrote the manuscript, A.H.-S. and F.G.-C. contributed to metagenomic and proteomic data analysis and to the writing of the manuscript, E.P. performed genome binning and polymorphism analysis, F.A.-P. initiated the experimental work and contributed to proteomic experiments, B.C.W.M. and A.C. performed phylogenetic analyses, L.G., B.L. and R.D. constructed the 16S rDNA environmental library and contributed to data analysis, D.V., D.M. and A.A.T.S performed the statistical analysis of metabolic data, S.G. and A.V.D. contributed to proteomic data analysis, C.C. and F.E.-P. performed physico-chemical analyses, O.B., J.C., E.T., N.F., V.B., P.O., M.B., P.S., and M.C. contributed to genome annotation and data analysis, C.M. coordinated software development for genome analysis (MaGe), J.W. and D.L.P. contributed to metagenome assembling and to the writing of the manuscript.

Supplementary figure legends

Figure S1. Distribution plot showing the repartition of the seven major contigs based on their respective GC% and mean coverage. Colors indicate the different bin assignment of supercontigs (CARN1, white, CARN2, red; CARN3, purple; CARN4, green, CARN5, blue; CARN6, yellow and CARN7, light red).

Figure S2. Amount of polymorphism in the sequenced population regarding the CARN bin reference sequences. **A:** SNPs+indel events frequencies relative to each bin sequence length; **B:** distribution of the polymorphism along the respective sequences of CARN1 and CARN4 bins.

Figure S3. Comparison between CARN5 genomic data and *Acidithiobacillus ferrooxidans* ATCC 23270. Both genomes were compared using MUMmer 3.0 (Kurtz *et al.*, 2004), using the nucleic mode, with *Acidithiobacillus ferrooxidans* as a reference, and individual contigs of the CARN5 bin as a query. X-axis represents the reference genome, Y-axis the % of similarity between the two sequences, the bottom red line representing the projection of the different segments mapped onto the reference genome. Similar results were obtained with the *Acidithiobacillus ferrooxidans* reference strain ATCC 53993.

Figure S4. Comparison between the experimental protein pattern obtained by MS/MS identification of the proteins expressed *in situ* (CARN1, green, CARN2, yellow; CARN3, black; CARN4,

pink, CARN5, light blue; CARN6, brown and CARN7, blue) and the theoretical distribution predicted from metagenomic data in grey.

Figure S5. Genetic environment of the 3 arsenic islands identified in the genome of the *Thiomonas* sp. CARN2 bin. The various CDS are indicated in the 6 phases according to their position along the chromosome on the MaGe interface (Vallenet *et al.*, 2006). A red color indicates that an expert annotation has been performed. A vertical black box indicates the limit of a contig (<http://www.genoscope.cns.fr/agc/mage/carnoulescope>).

The supplementary tables 5 to 8 are oversized and thus cannot be printed. They remain available upon request to the authors.

pH	T (°C)	DO ¹ (mg.l ⁻¹)	Eh ² (mV)	As(III) (mg.l ⁻¹)	As(V) (mg.l ⁻¹)	Fe(II) (mg.l ⁻¹)	Fe(III) (mg.l ⁻¹)	SO ₄ ²⁻ (mg.l ⁻¹)
3.83	15.1	8.3	459	71 (±4)	16 (±1)	625(±30)	<DL	3209 (±300)

¹ Dissolved oxygen

² Redox potential

Supplementary Table 1: Physico-chemical characteristics of the Carnoules water during the sampling at COWG on May 2007

Carn	Size	Events (SNPs + indels)	ratio
Carn-1	2558671	53768	2,10E-02
Carn-2	4069759	3856	9,47E-04
Carn-3	1567831	786	5,01E-04
Carn-4	2760983	76894	2,79E-02
Carn-5	2984465	12651	4,24E-03

Supplementary Table 2: Amount of relative variability of each Carn bin, as determined by mapping individual reads on the bin sequences using ssaha2 (Ning *et al.*, 2001).

Group	Domain	Phylum	Class	Order	Closest organism
CARN1	Bacteria	-	-	-	CARN4
CARN2	Bacteria	Proteobacteria	β-Proteobacteria	Burkholderiales	<i>Thiomonas</i> sp. 3As
CARN3	Bacteria	Acidobacteria	Unclassified Acidobacteria	Unclassified Acidobacteria	<i>A. bacterium</i> Ellin345
CARN4	Bacteria	-	-	-	CARN1
CARN5	Bacteria	Proteobacteria	γ-Proteobacteria	Acidithiobacillales	<i>At. ferrooxidans</i> ATCC 23270
CARN6	Bacteria	-	-	-	-
CARN7	Bacteria	Proteobacteria	β-Proteobacteria	-	-

Supplementary Table 3: Phylogenomic analysis of the Carnoules seven bin-genomes using 27 universal marker genes (Ciccarelli *et al.*, 2006)

Chapitre 4 – Etude d’une communauté complexe issue d’une AMD par des approches protéogénomiques

	Carnoulès 1	Carnoulès 2	Carnoulès 3	Carnoulès 4	Carnoulès 5	Carnoulès 6	Carnoulès 7
Number of base pairs	2558671	4069759,00	1567831,00	2760983,00	2984465,00	2824537,00	2708310,00
G+C content (%)	62,13	64,66	58,55	61,89	58,40	58,51	51,70
No. of predicted protein-coding genes	2664	4302	1498	2806	3292	2851	2949
No. of predicted proteins of unknown function (%)	1267 (47,56 %)	1561 (36,29 %)	781 (52,14 ù)	1359 (48,43 ù)	1500 (45,57 %)	1531 (53,70 %)	1322 (44,83 %)

	Carnoulès 1	Carnoulès 2	Carnoulès 3	Carnoulès 4	Carnoulès 5	Carnoulès 6	Carnoulès 7
	Nb (%)	Nb (%)	Nb (%)	Nb (%)	Nb (%)	Nb (%)	Nb (%)
Cellular processes and signaling							
Cell cycle control, cell division, chromosome partitioning	33 (1,24)	50 (1,16)	9 (0,06)	49 (1,75)	46 (1,4)	38 (1,33)	51 (1,73)
Cell wall/membrane/envelope biogenesis	139 (5,22)	202 (4,7)	70 (4,67)	177 (6,31)	193 (5,86)	153 (5,37)	137 (4,65)
Cell motility	60 (2,25)	112 (2,6)	29 (1,94)	61 (2,17)	53 (1,61)	63 (2,21)	97 (3,29)
Posttranslational modification, protein turnover, chaperones	101 (3,79)	177 (4,1)	46 (3,07)	119 (4,24)	136 (4,13)	88 (3,09)	118 (4)
Signal transduction mechanisms	121 (4,54)	230 (5,35)	55 (3,67)	121 (4,31)	100 (3,04)	105 (3,68)	156 (5,29)
Intracellular trafficking, secretion, and vesicular transport	62 (2,33)	120 (2,79)	52 (3,47)	79 (2,82)	131 (3,98)	89 (3,12)	104 (3,56)
Defense mechanisms	56 (2,1)	107 (2,49)	36 (2,4)	66 (2,35)	74 (2,25)	59 (2,07)	62 (2,1)
Extracellular structures	1 (0,04)	1 (0,02)	1 (0,07)	3 (0,11)	2 (0,06)	1 (0,04)	1 (0,03)
RNA processing and modification		1 (0,02)			1 (0,03)		1 (0,03)
Information storage and processing							
Chromatin structure and dynamics	1 (0,04)	1 (0,02)		1 (0,04)	1 (0,03)		1 (0,03)
Translation, ribosomal structure and biogenesis	161 (6,04)	177 (4,1)	73 (4,87)	166 (5,92)	161 (4,89)	75 (2,63)	103 (3,49)
Transcription	120 (4,5)	230 (5,35)	66 (4,41)	148 (5,27)	141 (4,28)	109 (3,82)	143 (4,85)
Replication, recombination and repair	202 (7,58)	372 (8,65)	99 (6,61)	182 (6,49)	272 (8,26)	264 (9,26)	307 (10,41)
Metabolism							
Energy production and conversion	161 (6,04)	335 (7,79)	64 (4,27)	177 (6,31)	199 (6,05)	149 (5,23)	172 (5,83)
Amino acid transport and metabolism	313 (11,75)	488 (11,34)	120 (8,01)	309 (11,01)	277 (8,41)	214 (7,51)	207 (7,02)
Nucleotide transport and metabolism	69 (2,59)	94 (2,19)	25 (1,67)	61 (2,17)	73 (2,22)	63 (2,21)	63 (2,14)
Carbohydrate transport and metabolism	139 (5,22)	232 (5,39)	101 (6,74)	173 (6,17)	171 (5,19)	107 (3,75)	125 (4,24)
Coenzyme transport and metabolism	108 (4,05)	152 (3,53)	42 (2,8)	110 (3,92)	113 (3,43)	72 (2,53)	102 (3,46)
Lipid transport and metabolism	112 (4,20)	164 (3,81)	36 (2,4)	112 (3,99)	75 (2,28)	63 (2,21)	65 (2,2)
Inorganic ion transport and metabolism	117 (6,64)	322 (7,48)	78 (5,21)	183 (6,52)	231 (7,02)	136 (4,77)	158 (5,36)
Secondary metabolites biosynthesis, transport and catabolism	64 (2,40)	132 (3,07)	29 (1,94)	76 (2,71)	68 (2,07)	36 (1,26)	44 (1,49)

Supplementary Table 4: General features of the 7 Carnoulès bins

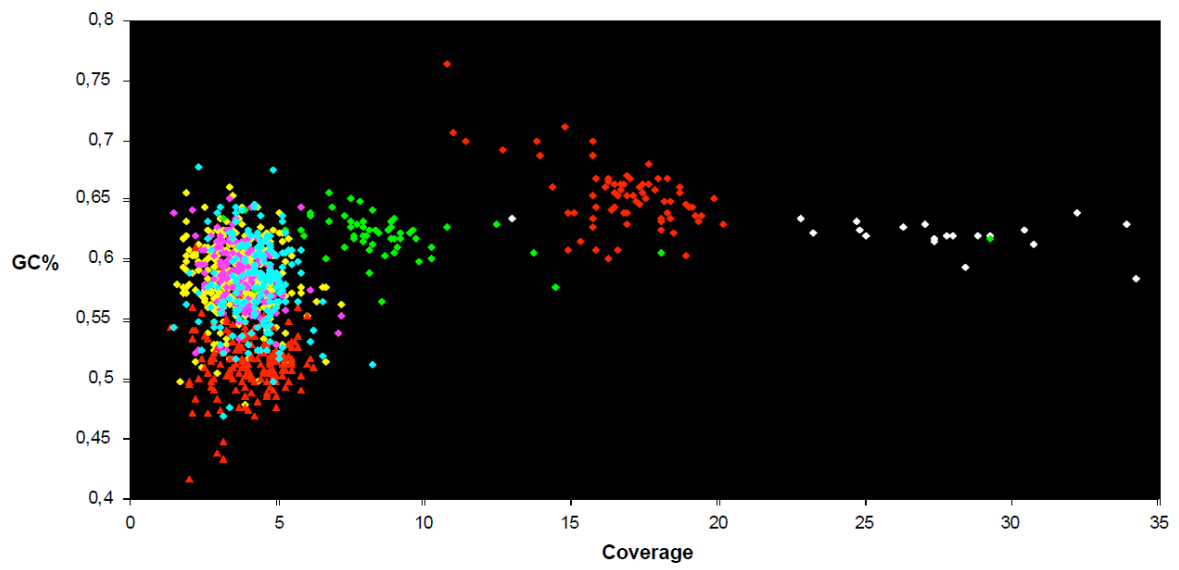


Figure S1

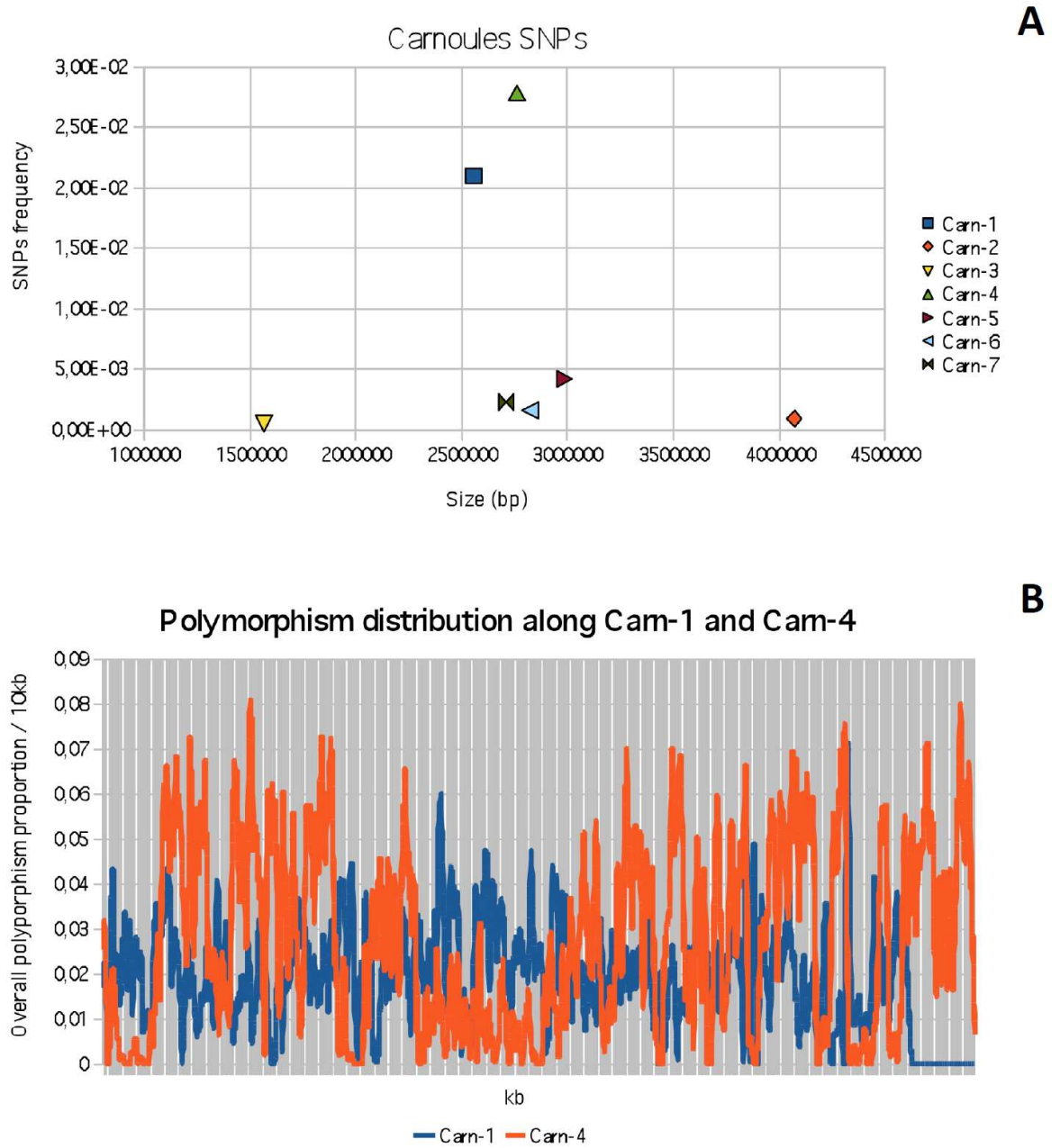


Figure S2

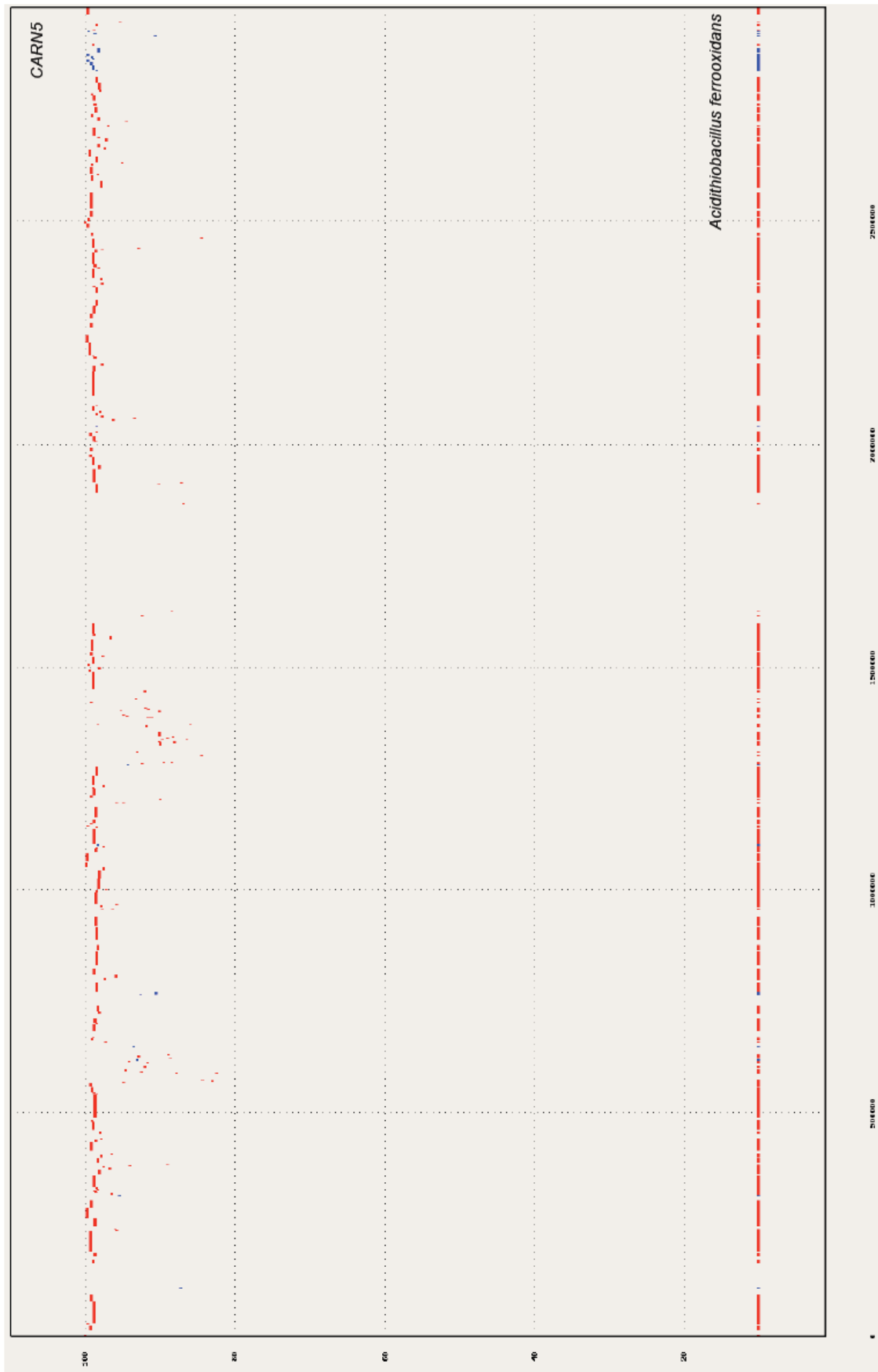


Figure S3

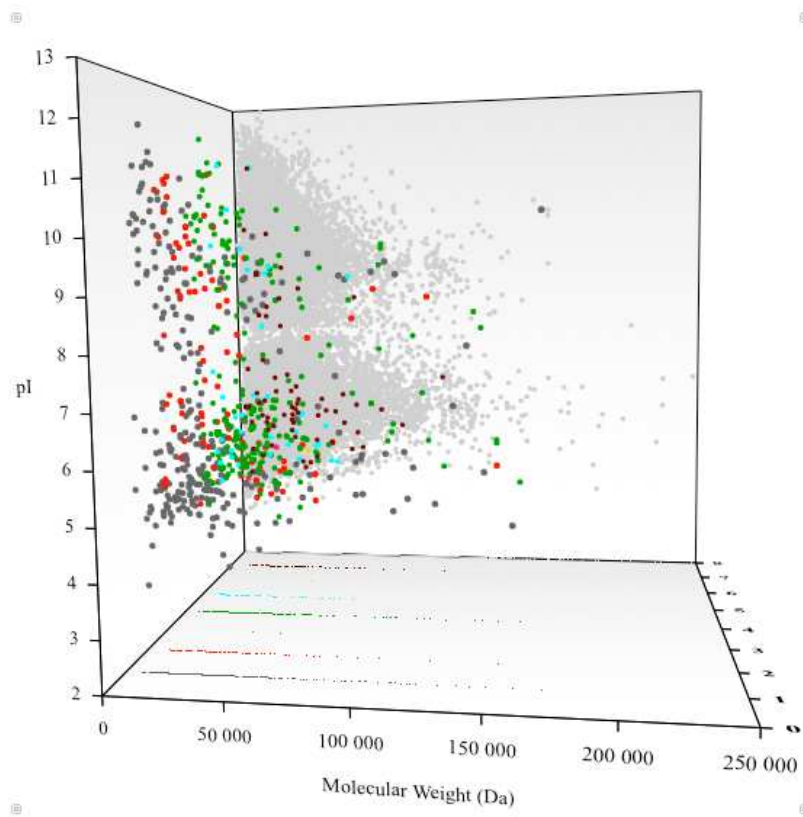


Figure S4

Materials and Methods

Sampling and chemical analysis. Samples were collected in the Reigous creek in May 2007 at the station called COWG (Bruneel et al., 2003) and located 30 m downstream of the spring. Five cm deep white sediments covering the bottom of the creek were collected using a sterile glass, and the running water (i.e. a thin column, less than 10 cm) covering these sediments was filtered (300 mL) through sterile 0.22 µm Nuclepore filters. These filters were transferred into a collection tube, frozen in liquid nitrogen, and stored at -80°C until further analysis. This sampling was done in three replicates.

The main physico-chemical parameters (pH, T°, dissolved oxygen...) were determined in the field at the sampling station and arsenic speciation, Fe(II) and sulfate analyses were performed as previously described (Goulhen-Chollet *et al.*, submitted).

DNA isolation and sequencing. To recover microbial cells, ten grams of sediments were washed in 10 mL of solution 1 (in g.L⁻¹: Na₂SO₄.10H₂O, 0.15 ; (NH₄)₂SO₄, 0.45 ; KCl, 0.05 ; MgSO₄.7H₂O, 0.5 ; KH₂PO₄, 0.05 ; Ca(NO₃)₂.4H₂O, 0.014) and agitated one night at 4°C. After 10 minutes decantation, 7.5 mL of supernatant were added without mixing on 17.5 mL of Nycodenz solution (Axis-Shield, Dundee, Scotland), and then centrifuged 30 minutes at 10 000 x g. The cellular fraction (upper fraction) was removed and washed by adding 2 volumes of solution 1 and centrifuge 15 min at 10 000 x g and 4°C. DNA was extracted from sediments either directly using the UltraClean Soil DNA Isolation Kit according to the recommendation of the manufacturer (MoBio Laboratories Inc., Carlsbad, CA, USA), or after separation of microbial cells by Nycodenz gradient, using the Wizard genomic DNA extraction kit (Promega, U.S.A.). All genomic DNA samples were stored at -20°C until further processing.

DNA extracted from bacterial cells was then fragmented by nebulisation and fragments ranging from 3 and 5 kb were used to construct a genomic library. DNA inserts were sequenced from both ends as previously described (Muller *et al.*, 2007). In parallel, 281,758 DNA reads were obtained by GS-FLEX pyrosequencing using standard procedures. Both procedures produced a total of 430.3 Mb.

Bioinformatics. Binning was performed as follows. First, seven scaffolds containing 16S rRNA gene sequences were characterized. The characterization of conserved gene families (Cicarelli, 2006) was also used for all the bins. Contigs were assigned to a bin if at least 2 contigs from a scaffold could be assigned to the same bin. Isolated contigs were not assigned to any contig as well as scaffolds with only one binned contig. Second, bins 1 and 4 have the same 16S rRNA gene sequence, similar GC% content but could be distinguished by their genomic coverage (6-13 versus 22-35) and bin 1 was used to complete bin 4 by contig sequence similarity. Bin 2 was obtained by similarity with *Thiomonas* sp.,

GC% and coverage. Bins 3 and 5 have a similar GC% and coverage were distinguished because of the similarity of bins 5 with *Acidithiobacillus ferrooxidans*. Finally bins 6 and 7 were separated by the presence of the 16S rRNA gene, conserved gene families and GC%.

Coding sequences were predicted using the AMIGene (Annotation of Microbial Genomes) software (Bocs *et al.*, 2003) and then submitted to automatic functional annotation using a set of tools (Vallenet *et al.*, 2006). Putative orthology relationships between two genomes were defined by gene pairs satisfying either the Bidirectional Best Hit criterion or an alignment threshold (at least 40% sequence identity over at least 80% of the length of the smallest protein). These relationships were subsequently used to search for conserved gene clusters (synteny groups) among several bacterial genomes using an algorithm based on an exact graph-theoretical approach (Boyer *et al.*, 2005). This method allowed multiple correspondences between genes, detection of paralogy relationships, gene fusions, and chromosomal rearrangements (inversion, insertion/deletion). The ‘gap’ parameter, representing the maximum number of consecutive genes that are not involved in a synteny group, was set to five. Manual validation of automatic annotations was performed using the MaGe web interface (Vallenet *et al.*, 2006) in the Arsenoscope relational database (<http://www.genoscope.cns.fr/agc/mage/arsenoscope>).

Clone library and phylogenetic analyses. The bacterial diversity was analyzed by cloning PCR amplified 16S rRNA genes. Bacterial 16S rRNA genes were amplified with 8F and 1489R primers (Weisburg *et al.*, 1991). PCR products were cloned in *Escherichia coli* TOP10 using the pCR2.1 Topo TA cloning kit (Invitrogen, Inc.) following manufacturer’s instructions. Cloned 16S rRNA gene fragments were then amplified using the primers M13F (5'- GTAAAACGACGGCCAG -3') and M13R (5'- CAGGAAACAGCTATGAC-3'), located on the vector. Clones (named CG) were then digested with the enzymes *HaeIII* and *HinfI* (New England Biolabs). Ninety-six clones from the library were analyzed on 3% agarose gel electrophoresis and grouped according to their RFLP patterns (*HaeIII* and *HinfI* digestion). Only one representative of each group was sequenced using the Big Dye® Terminator v3.1 cycle sequencing kit (Applied Biosystems). Sequences were compared with the RDP database (<http://rdp.cme.msu.edu>) by BLAST online searches (Altschul *et al.*, 1998). Multiple sequence alignment of clones and 16S rRNA sequences from CARN bins, except for CARN4, along with closely-related sequences of known phyla chosen on BLAST similarities was performed by using CLUSTALX (Thompson *et al.*, 1997). Phylogenies were constructed with the Molecular Evolutionary Genetics Analysis v4.0 program (Tamura *et al.*, 2007) using Nucleotide: Maximum Composite Likelihood model and Neighbour-joining algorithm. Trees were then annotated with Itol (Letunic and Bork, 2007). Significance of branching order was determined by bootstrap analysis with 1000 resampled data sets. PAST (PAleontological Statistics v1.60) software from <http://folk.uio.no/ohammer/past/> website was

used to perform rarefaction analysis for the clone library with clone phenotype similarity defined at 97% 16S rRNA sequence similarity. Coverage value was calculated to determine how efficient our clone library described the complexity of a theoretical community such as original bacterial community. The coverage (Good, 1953) value is given as $C = 1/(n1/N)$ where $n1$ is the number of clones which occurred only once in the library. The sequences of clones CG determined in this study have been submitted to the EMBL database and assigned Accession Nos. FN391809 to FN391849.

Phylogenomic approach. Molecular phylogenies were inferred using 27 marker genes (Supplementary Table 2) chosen from the reference gene set proposed to reconstruct the tree of life (Cicarelli et al., 2006). For each marker, the corresponding family of homologous genes from the HOGENOM database (Perrière et al., 2000; release 4) was identified. Eukaryotic sequences were removed from each family. Moreover, we completed these families with sequences from 64 new prokaryotic genomes of potential interest in the framework of this study (Supplementary Table 3). Then, each family dataset was aligned with the program MUSCLE (Edgar, 2004) and filtered using the program GBLOCKS (Castresana, 2000; Talavera and Castresana, 2007) in order to select unambiguously aligned positions. Maximum-likelihood phylogenies were reconstructed with PhyML (v2.4.4, Guindon and Gascuel, 2003) using the Jones–Taylor–Thornton (JTT) model of amino-acid substitution (Jones et al. 1992) and performing 1000 bootstrap replicates. Heterogeneities between sites were estimated under a gamma law based model of substitution, with estimation of the alpha parameter by PHYML.

A super-alignment was also performed by concatenating the individual alignments of the 9 marker genes that were most commonly found in the bins (Supplementary Table 4). This approach allows increasing the phylogenetic signal by summing the information contained in each individual dataset. Missing data corresponding to taxonomic sampling differences were coded by gaps. Then, a maximum-likelihood phylogeny was reconstructed using the same options used for individual phylogenies.

Protein extraction and mass spectrometry identification. Cells recovered using Nycodenz gradient were resuspended in a buffer containing SDS concentration (1%, w/v) and β -mercaptoethanol (2.5%) and boiled 5 min before loading. Proteins were separated by SDS-PAGE using 12 % gradient slab gels (PROTEAN II, Bio-Rad laboratories). Electrophoresis was carried out at 50 mA per gel. Proteins were stained with coomassie brilliant blue R-250, systematically excised from the gel and stored at -20 °C until analysis. *In gel* digestion of gel slices was performed as previously described (Weiss et al., 2008).

The resulting peptide extracts were analyzed by nanoLC-MS/MS on a nanoACQUITY Ultra-Performance-LC (UPLC, Waters, Milford, MA) coupled to SYNAPT hybrid quadrupole orthogonal

acceleration time-of-flight tandem mass spectrometer (Waters, Milford, MA). The capillary voltage was set at 3,500V and the cone voltage at 35V. For tandem MS experiments, the system was operated with automatic switching between MS and MS/MS modes. The 3 most abundant peptides, preferably doubly and triply charged ions, were selected on each MS spectrum for further isolation and CID fragmentation with 2 energies set using collision energy profile. The complete system was fully controlled by MassLynx 4.1 (SCN 566, Waters, Milford, MA). Raw data collected during nanoLC-MS/MS analyses were processed and converted with ProteinLynx Browser 2.3 (23, Waters, Milford, MA) into .pkl peak list format.

The MS/MS data were analyzed using the MASCOT 2.2.0. algorithm (Matrix Science, London, UK) to search against a target-decoy protein database composed of predicted protein sequences from the 7 bin-genomes of the Carnoulès microbial community concatenated with reversed copies of all sequences and with common contaminants such as keratins and trypsin. Spectra were searched with a mass tolerance of 30 ppm for MS and 0.1 Da for MS/MS data, allowing a maximum of one missed cleavage with trypsin and with carbamidomethylation of cysteines and oxidation of methionines specified as variable modifications. Protein identifications were validated when at least two peptides with high quality MS/MS spectra (less than 12 points below Mascot’s threshold score of identity at 95% confidence level) were detected. In the case of one-peptide hits, the score of the unique peptide had to be greater (minimal “difference score” of 10) than the 95% significance Mascot threshold. After removal of keratins and trypsin, these thresholds allowed identifying 516 unique proteins with an estimated false discovery rate less than 4.4 % searching the target-decoy database. An additional database search was performed against a target-decoy protein database composed of sequences downloaded from the NCBI database restricted to the closest organisms of the 7 bin-genomes: *Acidobacteria bacterium* Ellin345, *Acidithiobacillus ferrooxidans* ATCC 23270, *Thiomonas* and *Gallionella*, concatenated with reversed copies of all sequences. Protein identifications were validated when at least two peptides with high quality MS/MS spectra (Mascot ion score greater than 35) were detected. In the case of one-peptide hits, the score of the unique peptide had to be greater than 100. These thresholds allowed identifying 4 additional unique proteins with no protein identified in the reversed sequences.

Statistical analysis. The metabolic network of each bin was predicted by the “Pathway Tools” software (Karp *et al.*, 2002) using MetaCyc (Caspi *et al.*, 2008) as a reference pathway database (version 12.0). This software applies selection rules to infer possible metabolic pathways and builds a special database for each bin, called a PGDB (Pathway/Genome Database). From those data, a two dimensional-matrix was built, wherein each line represents a bin (CARN1 to CARN7), each column corresponding to a specific reaction; each value counts how often genes linked to the given

enzymatic reaction were seen within the given bin. This frequency matrix is the starting-point of an exploratory factorial analysis, in order to highlight possible metabolic distinctions between bins. All bins were taken into account in this statistical analysis, but only reactions with non-constant counts could be analysed. The particular method implemented here was a Multiple Factor Analysis (MFA), the specificity of which is to group variables (in this case, normalized quantitative variables) and assign to each group the same weight in the global analysis. In our case, reactions were grouped into the MetaCyc pathways they belonged to (a same reaction belonging to several pathways is thus repeated as many times, one for each pathway/group). This approach favours the discovery of factors which are common to several groups.

In total, 715 distinct reactions were included in the analysis, for a total of 359 pathways. After examination of the amount of inertia captured by the method's resulting factors, 3 were kept for further analysis. For graphical representations, the variable (reaction) plot and the individual (bins) plot were combined, restricting plotted variables to those with a quality of representation greater than 0.75, in order to conserve interpretability.

Solely as an aid to interpretation and to listing readability, for each factorial plane, the reaction variables were hierarchically clustered according to the *local* angles between their vectors. The clustering method used an Euclidean distance and ward's criterion; the number of classes was chosen after manual examination of the cluster tree (Figure S1), and led to 7 and 9 classes for the first and second factorial planes, respectively. As a further aid to interpretation, the variables were plotted in two separate disks, according to the observation (or lack of) of the proteins corresponding to the reactions in the metaproteomic data.

III RÉSULTATS COMPLÉMENTAIRES

1. L’opéron *aox* dans le métagénome

Différents auteurs (Casiot *et al.*, 2003b) (Bryan *et al.*, 2009) ont mis en évidence le rôle prépondérant de la souche *Thiomonas* 3As dans la bioremédiation du site de Carnoulès. Les analyses phénotypiques ont montré sa capacité à oxyder l’As III. D’autre part les gènes *aox* sont retrouvés dans le génome de *Thiomonas* 3As, ainsi que dans d’autres espèces de *Thiomonas* isolées du site, comme récemment décrit par Arsène-Ploetze *et al.*, soumis (Chapitre 1-Partie C).

Nous avons confirmé ce rôle pivot des *Thiomonas* par l’analyse métagénomique des sédiments du Reigous. En effet, de façon très surprenante, nous avons mis en évidence qu’un seul des sept pseudo-génomomes, Carn 2 (*Thiomonas* sp.), possède les gènes codant pour l’arsénite-oxydase (Figure 4.1). Ces gènes ont également été cherchés dans les données non assemblées en pseudo-génomomes, correspondant à des gènes de bactéries non majoritaires⁶. Ainsi seuls quatre gènes *aoxB* et un gène *aoxA* supplémentaire ont été retrouvés, présentant respectivement une identité de plus de 98 % par rapport aux *aoxB* de Carn 2 et 92% d’identité entre les *aoxA*.

⁶ Sur les 550920 lectures (366.4 Mb) issues du séquençage aléatoire de la banque d’ADN métagénomique par la méthode de Sanger, deux tiers ont été assemblés par le programme Arachne (Batzoglou *et al.*, 2002) en 342 scaffolds de 1740 pb à 1.26 Mb. Les pseudo-génomomes Carn 1 à 7 ont été constitués en combinant les analyses d’ADN ribosomal 16S, le pourcentage en GC et le taux de couverture moyen, ainsi que l’aide des génomes de *Thiomonas* 3As et d’*Acidithiobacillus ferrooxidans* qui ont été utilisés comme matrice pour biner les pseudo-génomomes 2 et 5 respectivement. Le tiers de lecture restant ainsi que 210 contigs assemblés par Arachné mais non binés représentent les données non assemblées.

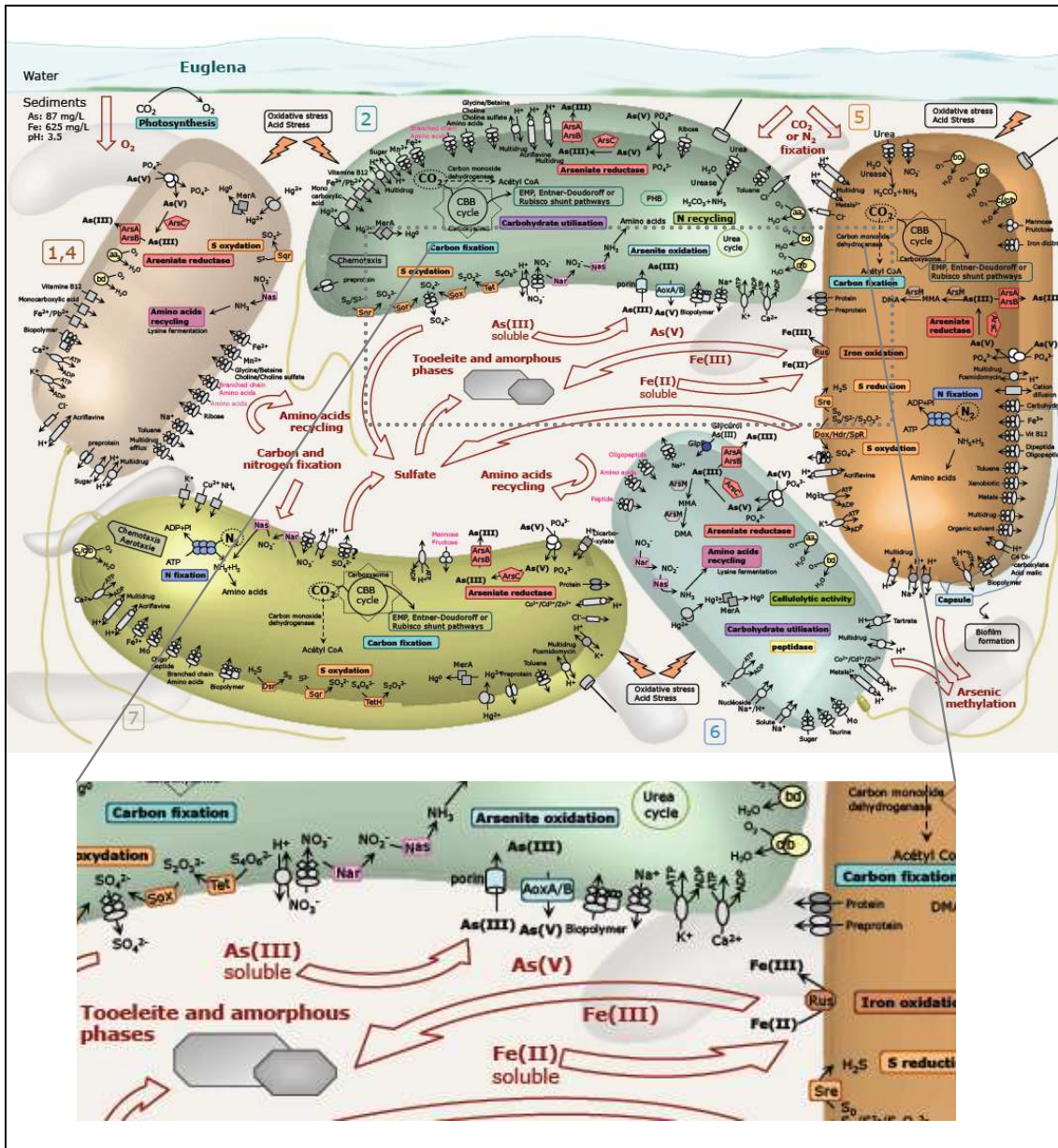


Figure 4.1: Représentation schématique de cinq pseudo-génomés reconstruits à partir des données de métagénomique et leur interaction (à partir de Bertin et al., soumis). Zoom sur l’activité arsénite oxydase de Carn 2 et son implication dans la précipitation de l’arsenic.

En outre, Carn 2 présente la singularité de posséder 3 opérons *aox* distincts (Figure 4.2). Dans aucun organisme décrit jusqu’à présent, il n’a été montré l’existence de copies multiples de cet opéron. Néanmoins Arsène-Plöetz et al., 2009 (Chapitre 1-Partie C) ont récemment montré, grâce à une approche par puces à ADN, que certaines souches de *Thiomonas*, elles aussi issues de Carnoules, possèdent deux copies de cet opéron. Ceci tend à démontrer que la présence d’au moins deux copies de cet opéron n’est pas due à un biais dans la reconstruction des pseudo-génomés, correspond bien à une réalité biologique. Il est cependant impossible d’exclure que ces opérons appartiendraient à

des souches de *Thiomonas* différentes mais trop proches (GC % et couverture) pour être distinguées lors de l’assemblage des contigs en pseudo-génomés.

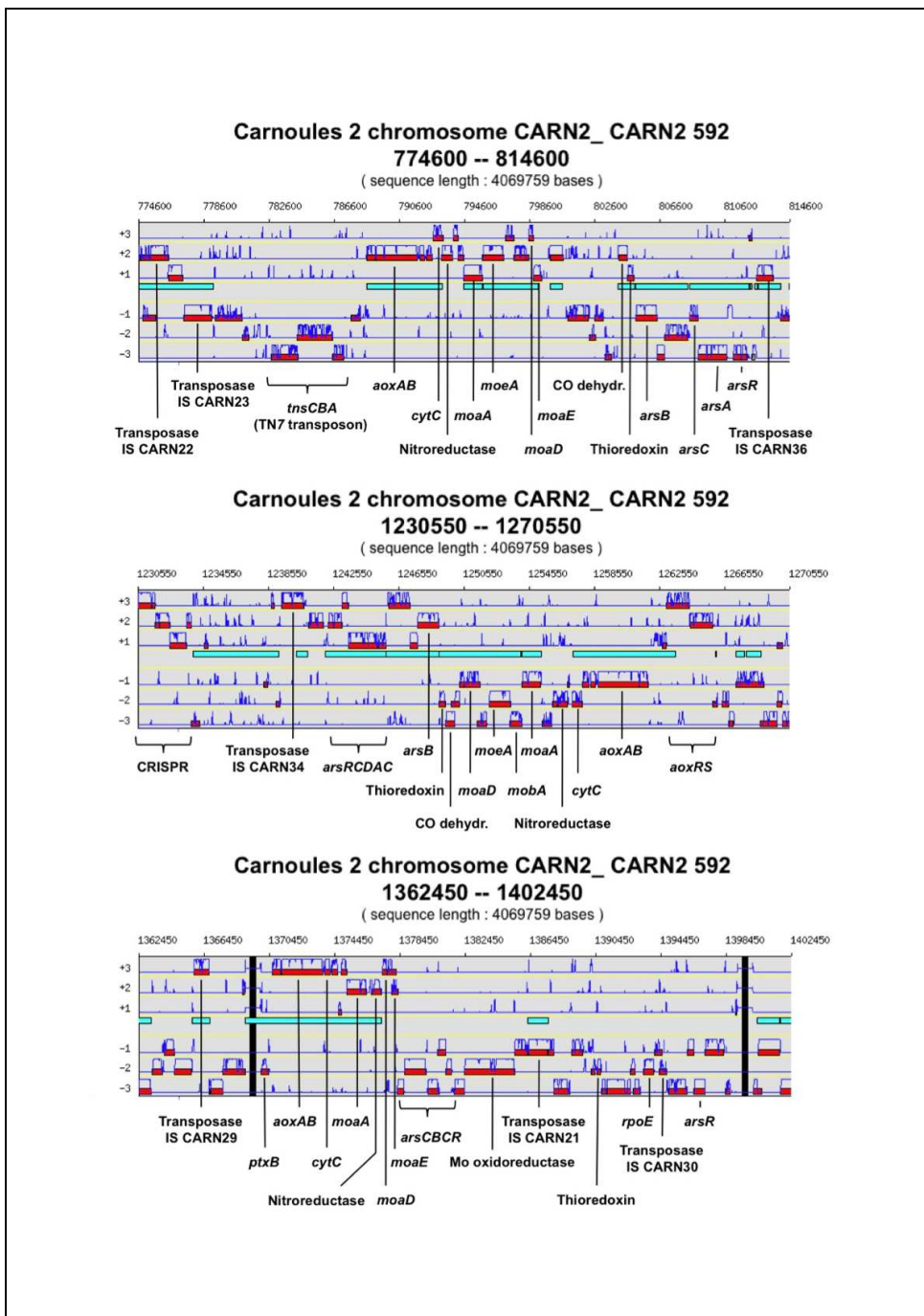


Figure 4.2 : Organisation génétique des régions d’intérêt 1, 2 et 3 du pseudo-génome Carn 2 (Figure issue des données annexes de la publication Bertin et al., 2009)

2. Analyses des régions *aox* de Carn2

a. Les régions aox dans le pseudo-génome Carn2

Le pseudo-génome Carn 2 est composé de 4 346 CDS répartis sur un génome de 4,03 Mb dont la représentation schématique est donnée Figure 4.3. Il est remarquable de noter la forte variabilité en CG par rapport à la moyenne, 64,66 %, ainsi que la présence d’un grand nombre d’éléments transposables. L’importance de ces éléments transposables a été démontrée chez *Thiomonas* 3As dans les différents îlots génomiques (Chapitre 1-Partie C). Ces éléments se trouvent en nombre au moins aussi important chez Carn 2, ce qui laisse suggérer de nombreux transferts horizontaux dans ce pseudo-génome.

Les opérons *aox* sont répartis en différentes régions du génome et situés dans des zones à forte variabilité en GC % (zone en GC pauvre) et à proximité d’éléments transposables. Les régions sont nommées région 1 (774600-814600), région 2 (12300550-1270550) et région 3 (1362450-1402450) et représentées par les encarts gris sur la figure 4.3.

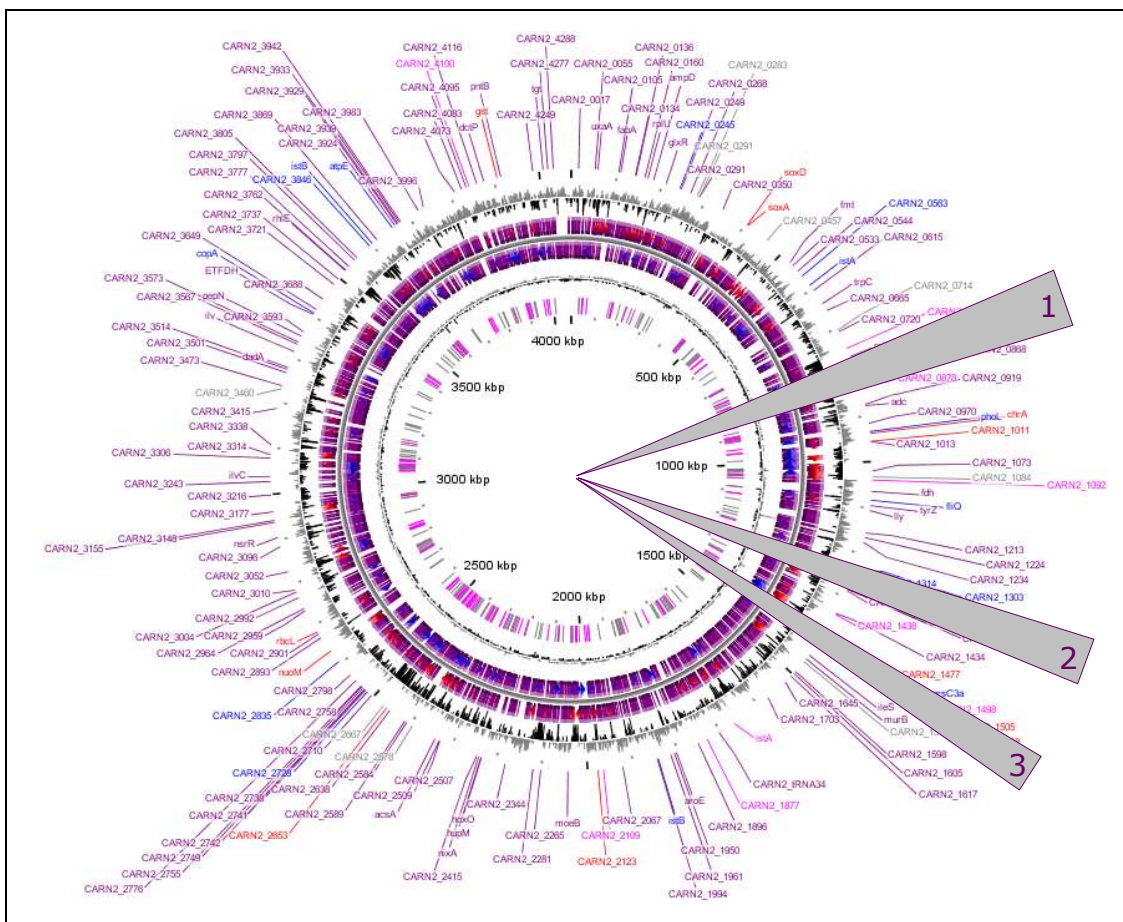


Figure 4.3 : Carte génomique circulaire du pseudo-génome Carn2. De l’extérieur vers l’intérieur : label de certaines CDS (en violet les CDS annotées automatiquement, en rose, rouge [brin +1] et bleu [brin-1] les CDS annotées expertement), courbe de déviation du pourcentage GC déterminée sur des fenêtres de 1000 bp, CDS prédites sur le brin +1, CDS prédites sur le brin -1, la courbe de proportion relative (G+C/G-C) déterminée sur des fenêtres de 1000 bp, éléments transposables en rose et pseudogènes en gris. Les encarts gris notés 1. 2 et 3 matérialisent les régions d’intérêt contenant les trois opérons *aox*.

Chacune de ces trois régions est composée principalement de trois groupes de gènes qui seront détaillés : l’opéron *aox*, les gènes de synthèse de cofacteur à molybdène et l’opéron *ars*.

b. L’opéron *aox* et ces gènes associés

L’opéron *aox* est constitué chez *H. arsenicoxidans* de 4 gènes : *aoxAB* codant respectivement pour la petite sous-unité Rieske à centre [2Fe-2S] et la grande sous-unité possédant un cofacteur à molybdène et un centre [3Fe-4S], de l’arsénite-oxydase. Le gène *aoxC* codant pour une nitroréductase et le gène *aoxD* codant pour un cytochrome c (Muller *et al.*, 2007). On retrouve également un système de régulation à deux composants constitués des gènes *aoxRS* : *aoxR* codant

pour le régulateur transcriptionnel et *aoxS* pour le senseur, comme illustré Figure 4.4 (Koechler *et al.*, 2009).

En amont de l’opéron *aox*, se trouve le gène *ptxB* codant pour un transporteur de type ABC impliqué dans le transport du phosphite. De la même façon que l’arséniate est un analogue structural du phosphate, l’arsénite l’est du phosphite et pourrait donc emprunter le système ABC-transporteur du phosphite, codé par le gène *ptxB* (White & Metcalf, 2004), pour entrer dans la cellule. Meng et collaborateurs (2004) ont montré que l’arsénite est importé au travers de la membrane via les aquaglycéroporines GlpF (protéine de transport du glycérol et autres petites molécules non chargées). Néanmoins l’interruption du gène ne génère pas un gain de résistance à l’As III (Meng *et al.*, 2004). Aussi il doit exister d’autres systèmes d’entrée de l’arsenic dans la cellule, d’autant que les gènes *glpF* n’ont pas été retrouvés chez Carn 2 ni chez *Thiomonas.sp.* 3As (Chapitre 1 – Partie C). Cet autre mécanisme d’entrée dans la cellule pourrait donc être, outre la diffusion passive, l’ABC-transporteur PtxB.

Dans le génome de *Thiomonas* 3As, l’opéron *aox* présente pratiquement la même organisation (Figure 4.4), néanmoins deux particularités peuvent être observées : 2 gènes codant pour des cytochromes-c et deux gènes *arsB*. En aval de cet opéron se trouvent des gènes de synthèse d’un cofacteur à molybdène détaillé ci-après.

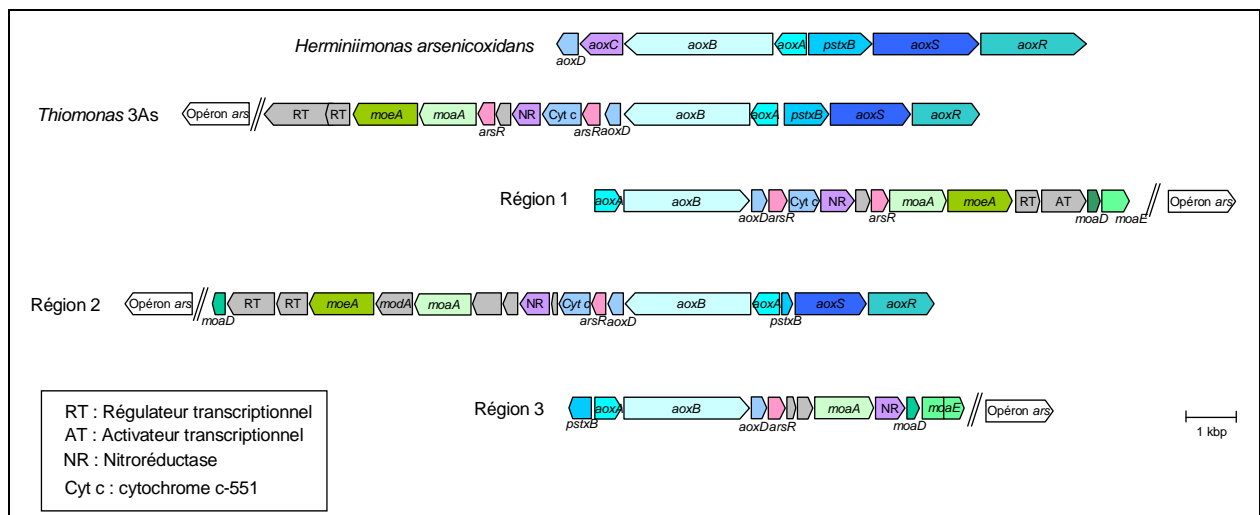


Figure 4.4 : Organisation synténique des opérons *aox* chez *H. arsenicoxidans*, *Thiomonas* 3As et dans les trois régions du pseudo-génome Carn 2.

En ce qui concerne Carn 2, les gènes *aoxAB*, deux gènes codant pour des cytochromes et un gène codant pour une nitroréductase sont retrouvés dans les trois régions d’intérêt ; ces gènes

présentent entre eux un très fort pourcentage d’identité. Néanmoins, seul l’opéron *aox* de la région 2 possède le système à deux composants *aoxRS*.

Il faut rappeler que ces pseudo-génomomes ont été reconstruits par assemblage de super-contigs en fonction du % GC et de la couverture des séquences. Ainsi les barres verticales présentées dans la zone 3 représentent les limites de contigs (Figure 4.4). On peut donc observer que, dans la région 3, le contig débute par le gène *ptxB*, codant pour le système de transport du phosphite, suivi de l’opéron *aox*. Ainsi il est possible d’émettre l’hypothèse de la présence d’un trou dans l’assemblage des contigs, le contig s’arrêtant juste avant les gènes du système de régulation à deux composants.

Dans la région 1 l’absence des gènes *aoxRS* peut être expliquée par l’abondance des gènes codant pour des transposases de part et d’autre de la zone. Ce cluster de gènes pourrait être une zone de transfert horizontal, ce qui expliquerait la perte des gènes *aoxRS* et du gène *ptxB*.

La régulation de ces deux opérons pourrait se faire par le système *aoxRS* présent dans la région 2, les protéines pourraient avoir un rôle de trans-régulateur des opérons des régions 1 et 3. Koechler et collaborateurs (2009) ont montré qu’en présence d’As III, AoxR était phosphorylé par AoxS. AoxR phosphorylé se lie alors à RpoN, le facteur sigma et son coactivateur se lie alors à la séquence -12/-24 du promoteur et activent la transcription (Introduction - Chapitre II) (Koechler *et al.*, 2009). Ainsi pour qu’AoxR ait un rôle de trans-régulateur, il est nécessaire que la séquence consensus de liaison du facteur sigma soit présente et conservée en amont de l’opéron *aox*.

L’alignement des séquences amont des gènes *aoxA* des trois régions avec la séquence consensus de fixation de RpoN montre une variabilité de séquence qui ne permet pas de conclure si les trois régions peuvent être régulées par ce mécanisme.

c. Les gènes impliqués dans la synthèse du cofacteur à molybdène

La grande sous-unité de l’arsénite-oxydase possède un centre à molybdène et un centre [3Fe-4S], comme décrit précédemment. Selon Mukhopadhyay *et al.*, 2002, ce cofacteur à molybdène a un rôle d’oxydant de l’As III en As V : deux électrons sont transférés au centre [3Fe-4S] qui est lui-même re-oxydé en transférant ces électrons au centre [2Fe-2S] de la petite sous-unité (Mukhopadhyay & Rosen, 2002).

La synthèse du cofacteur nécessite en premier lieu l’entrée du molybdène dans la cellule grâce à l’expression des gènes *mod* (*modB* correspondant au gène Carn2_3015). Puis l’activation du molybdate qui est générée par MoeA (MoeA catalyse la formation du complexe Mo-S permettant la synthèse de la molybdoptérine) (Hasona *et al.*, 1998). MoeB quant à lui active la molybdoptérine

synthèse qui permet la conversion du précurseur Z en molybdoptérine. Enfin la protéine codée par le gène *moa* est impliquée dans la conversion du 5'-GTP en précurseur Z de la molybdoptérine (les gènes impliqués *moaACDE* sont répartis sur Carn 2). La figure 4.5 représente un résumé de la synthèse du coeur molybdoptérine.

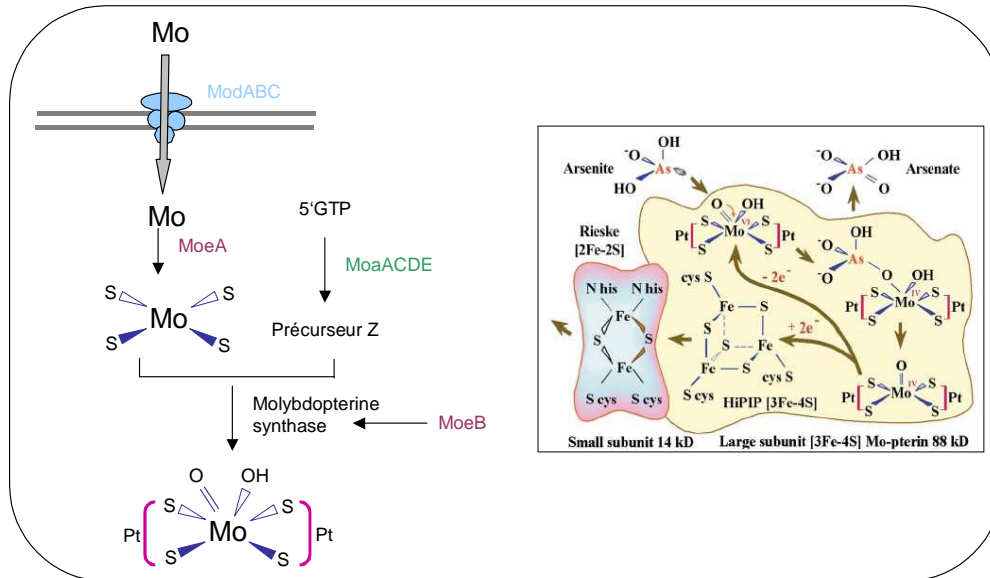


Figure 4.5 : Proposition de représentation schématique de la synthèse du cœur molybdoptérine et représentation schématique et du transfert d'électron de l'arsénite-oxydase selon (Mukhopadhyay et al., 2002) .

Ainsi certains gènes de synthèse de ces cofacteurs sont retrouvés aux abords de la région de l'opéron *aox* chez la souche *Thiomonas* 3As et dans la région 1 de Carn2. C'est le cas des gènes *moeA*, *moaA*, aux abords de la région 2, *moeA*, *modA* et *moaA*. Enfin la région 3 contient les gènes *moaACE*, comme illustré dans la figure 4.5.

Ces gènes essentiels dans l'activité arsénite-oxydase ne sont pas retrouvés en opéron complet, toutefois dans le pseudo-génome pris dans son ensemble, la fonction est conservée.

d. Les gènes *ars*

Le système de résistance à l'arséniate chez les procaryotes est codé par l'opéron *ars*. L'As V est réduit en As III par l'enzyme cytoplasmique l'arséniate réductase ArsC, il est ensuite sorti de la cellule grâce au transporteur ArsB qui peut être couplé à l'ATPase AsrA. La pompe ArsB peut également fonctionner seule comme un antiporteur [As(OH)₃/H⁺]. Il existe pour réguler l'opéron deux répresseurs transcriptionnels ArsR et AsrD. ArsR contrôle le niveau basal de l'expression de l'opéron, c'est-à-dire qu'il réprime l'expression de l'opéron en absence d'inducteur. Le gène *arsD*, présent

uniquement dans certains opérons, code pour un second répresseur ArsD qui contrôle le niveau d'expression maximal de l'*ars*, évitant ainsi la surexpression du transporteur. Il a d'autre part été proposé un rôle de chaperon à cette protéine, en effet ArsD présenterait l'As III à ArsA ce qui induirait l'activité ATPasique et donc l'efficacité du système (Lin *et al.*, 2007b) (Lin *et al.*, 2007a).

Chez Carn 2 trois opérons *ars* sont retrouvés dans le pseudo-génome. Ces trois opérons jouxtent les trois opérons *aox*. Ainsi les gènes *arsBCR* sont retrouvés dans les trois opérons, *arsA* dans les régions 2 et 3 et *arsD* uniquement au sein de la région 2. La présence d'opérons multiples chez des souches issues d'environnements arséniés n'est pas rare, comme par exemple chez *H. arsenicoxydans* (Muller *et al.*, 2007) ou chez *Pseudomonas* S11 (Chapitre 1-Partie A).

3. Conclusions

Les trois régions étudiées ici révèlent être composées des opérons *aox*, *ars* et de gènes impliqués dans la synthèse du cœur molybdoptérique de l'arsénite-oxydase. D'une part, outre les copies multiples de l'opéron *aox*, cette étude a révélé Carn 2 comme le seul pseudo-génome possédant la capacité d'oxyder l'As III en As V dans cet écosystème. Ceci en fait un organisme pivot dans la bioremédiation du site, si on considère que seul l'arsénite-oxydase oxyde l'As III en As V. En effet l'analyse métagénomique se base sur une comparaison par Blast pour l'annotation des gènes, elle ne permet donc pas de mettre en évidence d'autres systèmes résolument nouveaux d'oxydation de l'arsénite.

D'autre part les opérons *aox* et *ars* sont retrouvés aux abords de nombreuses transposases et dans des zones présentant % GC différents du reste du pseudo-génome, ceci pourrait laisser penser à un transfert horizontal ; ces régions pourraient constituer des îlots de résistance et d'oxydation. Ainsi les fonctions les plus importantes de l'écosystème de Carnoulès seraient le fruit de transferts horizontaux d'ADN entre procaryotes.

Perspectives & Conclusions

L'objectif de ce travail de thèse était d'étudier la diversité des mécanismes d'oxydo-réduction de l'arsenic chez les bactéries. L'étude s'est principalement focalisée sur la diversité des gènes *aox* d'isolats bactériens et de communautés complexes *in situ*. Trois aspects majeurs des travaux peuvent être mis en avant et seront discutés ici :

- le premier concerne la diversité des souches arsénite-oxydantes et des gènes *aox* correspondants,
- le second consiste en l'analyse de la diversité des arsénite-oxydases en se basant sur une approche par amorces dégénérées,
- et enfin le troisième a porté sur l'analyse de la diversité des arsénite-oxydases par une approche de génomique environnementale.

Dans des écosystèmes arséniés comme les sites miniers de Carnoulès, de Salsigne ou de Gabe-Gottes près de Sainte-Marie-aux-Mines, présentés dans le corps de ce manuscrit, une flore microbienne variée a pu s'adapter, se développer et coloniser ces milieux grâce à différents mécanismes. Un exemple de ces mécanismes qui a été précédemment décrit (Ellis *et al.*, 2001), consiste en l'oxydation de l'arsénite As III très toxique et mobile en arséniate AS V moins toxique et immobilisé. Ce système est codé par les gènes *aox*. D'autres systèmes de détoxification sont également décrits comme l'oxydation par le système de réduction dissimilatrice Arr (Richey *et al.*, 2009) ou le système de résistance (système *ars*) (Paez-Espino *et al.*, 2009).

Dans ce travail, je me suis principalement focalisée sur le système *aox*. En effet, le système *ars*, très répandu dans les environnements, a très largement été étudié et le système Arr n'a été mis que très récemment en évidence. Ainsi, trois souches arsénite-oxydantes isolées de divers milieux arséniés ont été caractérisées physiologiquement (*Pseudomonas* S11, *Leptothrix* S1.1 et *Thiomonas* 3As). Une analyse génétique a complété cet aspect physiologique pour la souche *Thiomonas sp.* La caractérisation de ces trois isolats oxydant l'arsénite, complète la diversité de souches capables d'oxyder l'arsenic déjà décrites dans la littérature. Cependant, ces microorganismes présentent des mécanismes particuliers d'oxydation de l'arsénite : une expression constitutive de l'arsénite-oxydase *Pseudomonas* S11 et une activité arsénite-oxydase élevée chez *Leptothrix* S1.1. Les analyses génétiques des gènes correspondants, ont révélé des gènes classiques d'oxydation, ce qui ne permet pas d'expliquer ces mécanismes particuliers. Des analyses biochimiques permettraient de mieux connaître l'activité de ces protéines mais aussi leurs interactions. Par exemple, des approches Tap-Tag (Rigaut *et al.*, 1999) qui permettent d'identifier les protéines partenaires en purifiant les complexes sous forme native, devraient mettre en évidence les éventuels partenaires de l'arsénite-oxydase. Le clonage des gènes *aoxABCD* d'*H. arsenicoxydans* chez *E. coli* ne permet pas l'activité de la

protéine malgré l'expression de ces gènes. Il reste donc encore à découvrir des partenaires nécessaires à cette activité.

L'étude du site de Gabe-Gottes a permis d'isoler des souches anaérobies arsénite-oxydantes qui pourraient étoffer la collection de souches arsénite-oxydantes aérobies déjà caractérisées. La recherche des gènes *aox* sur ces souches anaérobies n'a pas permis de révéler la présence de ces gènes. On peut donc émettre deux hypothèses : les gènes codant ces arsénite-oxydases sont trop divergents pour permettre l'utilisation d'amorces dégénérées ou ils ne correspondent pas aux arsénite-oxydases déjà connues. Afin d'analyser ces enzymes, il serait possible de procéder à la migration d'extraits protéiques de ces souches sur gel natif couplé à la détection biochimique de l'activité arsénite-oxydase. Cette approche a par exemple été réalisée lors de l'analyse du gène *aox* de *Pseudomonas* S11. Le séquençage *de novo* des protéines correspondantes par spectrométrie de masse pourrait ainsi permettre de mettre en évidence des arsénite-oxydases résolument nouvelles.

Pour poursuivre l'étude de la diversité des métabolismes/mécanismes d'adaptation à l'arsenic, il serait intéressant d'étendre l'étude à des souches présentant des métabolismes différents liés à l'arsenic comme *Rhizobium* NT-26, chimolithoautotrophe (Santini *et al.*, 2000), séquencée et étudiée au laboratoire ou une souche anaérobie hyperthermophile sporulante et sulfato-réductrice *Desulfotomaculum hydrothermale* (Haouari *et al.*, 2008) en cours de séquençage par le Génoscope.

La seconde partie de ce travail s'est focalisée sur la diversité des arsénite-oxydases en se basant sur une approche reposant sur l'utilisation d'amorces dégénérées. Les études menées sur le site de Sainte-Marie-aux-Mines et celles ayant conduit à la publication Quéménéur et collaborateurs ont permis de mettre en évidence une variabilité des séquences AoxB, en particulier une ségrégation des alpha-protéobactéries par rapport aux beta-protéobactéries et gamma-protéobactéries. Néanmoins, et très logiquement, eu égard à la conservation de fonction, les régions catalytiques de l'enzyme ainsi que les sites de liaison à la sous-unité Rieske de chaque séquence présentent une forte homologie de séquence entre elles. La protéine AoxB peut ainsi être considérée comme marqueur moléculaire des arsénite-oxydases aérobies.

Un grand nombre de séquences *aox* a pu être séquencé dans ce travail de thèse. Les analyses sont cependant basées sur l'utilisation d'amorces et présentent donc un biais puisqu'elles permettent de trouver uniquement des séquences plus ou moins proches de celle déjà connues. Pour mettre en évidence des arsénite-oxydases résolument nouvelles en se détachant des *a priori* de séquences connues, une approche par complémentarité fonctionnelle basée sur le criblage des banques métagénomiques pour l'activité arsénite-oxydase pourrait être menée, sur le site de Gabe-

Gottes par exemple, puisqu'un grand nombre de souches arsénite-oxydantes y a été détectée. Chez *E. coli* complétée par l'opéron *aox*, aucune activité arsénite-oxydase n'a été détectée ; un autre outil de complémentation fonctionnelle est donc requis. Il pourrait s'agir par exemple d'un mutant arsénite-oxydase de la souche *H. arsenicoxydans* à condition de disposer d'un plasmide compatible.

Enfin la troisième partie de ce travail a porté sur l'analyse de la diversité des arsénite-oxydases par une étude de génomique environnementale du site de Carnoulès (Gard). Ce site présente une bioremédiation naturelle de l'arsenic, suggérant la présence de mécanismes d'adaptation et de détoxification *in situ* de l'arsenic. Cette étude a été menée en combinant des approches de métagénomique, de métaprotéomique et des analyses phylogénétiques et chimiques des sédiments du site. Grâce à la reconstruction de sept pseudo-génomomes et à l'analyse fine des potentialités génétiques de chacun d'eux en regard de l'ensemble des protéines exprimées par la communauté bactérienne *in situ*, nous avons exploré les interactions trophiques des différentes bactéries majoritaires au sein de cet AMD et les mécanismes de détoxification *in situ*. L'analyse des séquences codant les arsénite-oxydases au sein de cette communauté bactérienne a révélé qu'elles étaient présentes uniquement dans le pseudo-génome Carn 2, du genre *Thiomonas*. La souche correspondante joue donc un rôle clef dans ce système puisque l'oxydation de l'arsénite en arséniate permet la coprécipitation avec le fer et la formation d'arsénopyrite par exemple, ce qui constitue une étape majeure de la bioremédiation du site. Les gènes *aox* ont d'autre part été recherchés dans l'ensemble des séquences (en dehors de l'assemblage en pseudo-génomomes) et seuls quatre gènes *aoxB* ont été retrouvés. De façon surprenante ce site présente une variabilité de séquence *aox* faible en comparaison du nombre de séquences retrouvées chez les microorganismes du sédiment de Gabe-Gottes. En effet, l'AMD de Carnoulès est un site très contaminé, avec une diversité bactérienne faible et peu de souches possédant les gènes *aox*, contrairement au site de Gabe-Gottes pour lequel la diversité de souches est élevée ainsi que le nombre de séquences *aoxB*, malgré une faible contamination à l'arsenic.

Par ailleurs, un nouveau phylum *Candidatus Fodinabacter aminivorans* (Carn1/4) a été mis en évidence. Ce micro-organisme non cultivé à l'heure actuelle pourrait présenter un intérêt industriel en raison de nombreux gènes codant des enzymes comme des protéases et des cellulases. Son exploration génomique de façon détaillée devrait permettre de mettre en évidence ses exigences nutritionnelles, y compris au niveau des sources de carbone qu'elle pourrait utiliser. Cette exploration permettrait donc de mettre au point un milieu de culture permettant peut être de l'isoler (approche de métagénomique inverse) et de la caractériser physiologiquement.

Cependant, la prudence reste de mise car ces analyses génomiques sont basées sur des homologies de séquence. Une erreur dans l'annotation d'un gène peut se répercuter sur les annotations faites en aval et il est donc important d'annoter de façon experte les gènes d'intérêt. De plus, l'assemblage des séquences en pseudo-génomomes repose entre autres sur le pourcentage en GC et sur la couverture de séquences. Il faut donc considérer les pseudo-génomomes comme pouvant représenter un assemblage de plusieurs micro-organismes. Enfin, lors de l'analyse génétique des pseudo-génomomes une grande proportion des gènes codent pour des fonctions inconnues (de 35% pour Carn2 à 53% pour Carn6). Il est donc nécessaire de coupler ces analyses à des tests physiologiques sur les souches isolées.

Enfin, en se basant sur ces résultats et l'expérience acquise lors de cette étude, une analyse spatio-temporelle des communautés présentes sur le site de Carnoulès peut être envisagée, par exemple par pyroséquençage. Ceci afin de mettre en évidence, entre autres, l'évolution de la population bactérienne et les interactions trophiques entre ces souches et les organismes eucaryotes : en fonction de divers points du Reigous (le ruisseau présente une bioremédiation naturelle) ou en fonction des fluctuations saisonnières (l'absence d'euglènes en saison hivernale qui peut influencer sur la quantité d'oxygène et la matière organique disponible). Des analyses similaires de génomique environnementale pourraient être étendues à d'autres AMDs arséniés (Rio Tinto par exemple) ce qui pourrait permettre de mettre en évidence des micro-organismes marqueurs de ces environnements. Ces approches restent encore coûteuses mais la comparaison des différents résultats devrait permettre de générer une vision globale du fonctionnement de ces écosystèmes.

Les différentes techniques utilisées pour notre étude vont de la microbiologie classique à la génomique environnementale. Elles ont conduit à la caractérisation de nouvelles souches arsénite-oxydantes et de leur enzyme, à l'analyse de la conservation de séquence des AoxB issues de diverses souches ou de communautés complexes. Ce travail a conduit également à l'étude de la diversité des arsénite-oxydases d'une communauté bactérienne complexe par des approches de génomique environnementale. Des analyses de métagénomique inverse pour isoler certaines souches d'intérêt, de la complémentation fonctionnelle sur les banques métagénomiques ou des tests biochimiques sur les arsénite-oxydases des isolats caractérisés dans ce travail, permettront de mettre en évidence de nouveaux partenaires de l'oxydation et des nouvelles arsénites-oxydases qui pourraient présenter un intérêt, y compris en biodépollution.

Monitorat

Suite au concours de l'École Doctorale, j'ai eu l'opportunité d'obtenir un poste de monitrice à l'Université de Strasbourg. J'ai intégré l'équipe pédagogique de L1 (licence première année) en travaux pratiques (TP) et travaux dirigés (TD) de techniques biologiques sous la tutelle de Pascaline Ullmann. Ces TP/TD ont pour objectif d'enseigner quelques techniques de bases nécessaires en biologie, mais aussi d'apprendre aux étudiants l'autonomie, les règles de sécurité et les bonnes pratiques de laboratoire. L'objectif, de tout enseignement pratique ou dirigé de biologie et de science en général, est de provoquer le questionnement, de développer le sens de l'observation, l'esprit critique et surtout la méthode scientifique. C'est-à-dire comment à partir de l'observation des données de l'expérimentation démontrer la véracité ou non des hypothèses de départ. Ce qui nécessite de structurer son cours. Il est essentiel de s'assurer que les étudiants possèdent le vocabulaire adéquat, les notions de base. Les expériences sont ensuite effectuées dans le but de développer la démarche scientifique des étudiants. Les conclusions de ces expériences sont alors rapportées au cours afin de compléter les informations.

Ces TP/TD sont organisés en trois parties intégrées et complémentaires comprenant des techniques de microbiologie (deux TP et un TD), de biologie moléculaire (deux TP et un TD) et d'immunologie (un TP et un TD). De façon plus détaillée les notions abordées sont :

Les bactéries : technique d'identification (identification sur milieu sélectif et différentiel – test de Gram - antibiogramme)

Etude de la croissance bactérienne (suivi de la croissance par mesure de l'absorbance et par technique de dilutions et étalements)

La transformation bactérienne (réalisation de bactéries compétentes - transformation par électroporation et par choc thermique)

Analyse d'ADN par digestion enzymatique et électrophorèse en gel d'agarose (digestion par des enzymes de restrictions - électrophorèse sur gel d'agarose)

Réaction antigène-anticorps (test ELISA – test de double diffusion Ouchterlony)

Lors de ces enseignements j'étais en charge de deux groupes d'étudiants que j'ai suivis tout au long des huit séances, et durant ma dernière année de monitorat j'ai suivi quatre groupes pour les enseignements de microbiologie. Outre l'encadrement j'étais aussi en charge de l'évaluation des étudiants par un contrôle terminal, mais aussi par des exposés présentés en TD.

Au cours de ces trois ans, j'ai allié recherche et enseignement. Ceci à été une expérience endurente mais très enrichissante tant d'un point de vue personnel que professionnel. En effet assurer un monitorat nécessite de consacrer du temps, mais permet de prendre du recul, de

« sortir » de la paillasse et entraîne à être organisé et efficace. Assurer un monitorat implique l'enseignement et ses à-côtés, et de nombreuses formations. Celles-ci m'ont, entre autres, enseigné le placement de la voie, la gestion du stress lors de la prise de parole en public, des éléments de législation... Les notions abordées lors de ces différentes formations m'ont été, pour certaines, utiles lors de mes interventions en TP/TD, mais aussi lors d'interventions scientifiques ou personnelles. Enfin le CIES (Centre d'Initiation à l'Enseignement Supérieur) m'a permis de participer par le biais de l'association Rementa (Réseau des jeunes Enseignants du supérieur, des Moniteurs et Ex-moniteurs National, Transdisciplinaire et Associatif), à la rédaction d'un ouvrage de vulgarisation pluridisciplinaire sur le rire et l'humour. J'y ai contribué par l'écriture d'un chapitre sur le rire et la génétique.

Le monitorat apporte en premier lieu une aisance lors de la prise de parole en public, mais surtout il aide au développement d'un discours structuré et pédagogique. D'autre part un des points fondamentaux du monitorat est le transfert des connaissances et donc le contact avec les étudiants. Il est primordial pour tout chercheur (ou professionnel au sens large) de maîtriser le transfert de ses connaissances, et l'enseignement lors du monitorat est un exercice extrêmement formateur à ce niveau. L'interaction avec les étudiants est un bon reflet de son talent d'orateur et il est aisé de se rendre compte si son discours et ses capacités de pédagogue sont de qualités ou non. Enfin le contact avec les étudiants est réellement très enrichissant d'un point de vue personnel. J'ai pris beaucoup de plaisir à enseigner et à partager mes connaissances.

Le monitorat est une façon d'exercer l'un des aspects importants de la fonction de scientifique qui est de partager ses connaissances, mais aussi de divulguer et transmettre la méthode scientifique. Cette expérience m'a montré la cohésion, le dynamisme et l'interaction entre les activités d'enseignement et de recherche.

Annexes

Annexe 1

Communications scientifiques

PUBLICATIONS SCIENTIFIQUES

M. Quéméneur, **A. Heinrich-Salmeron**, D. Muller, D. Lièvreumont, M. Jauzein, P. N. Bertin, F. Garrido, and C. Joulian. 2008. Diversity surveys and evolutionary relationships of *aoxB* genes in aerobic arsenite-oxidizing bacteria. *Applied and Environmental Microbiology*, 74(14) 4567-73.

Laurence N. Warr, Julia N. Pedrial, Marie Claire Lett, **Audrey Heinrich-Salmeron**, Mohamed Khodja. How clays stimulate marine oil bioremediation. *Applied Clay Science*.46, p337-345.

Florence Arsène-Ploetze, Sandrine Koechler, Marie Marchal, Jean-Yves Coppee, Michael Chandler, Violaine Bonnefoy, Céline Brochier-Armanet, Mohamed Barakat, Valérie Barbe, Fabienne Battaglia-Brunet, Odile Bruneel, Christopher G. Bryan, Jessica Cleiss-Arnold, Stéphane Cruveiller, Mathieu Erhardt, **Audrey Heinrich-Salmeron**, Florence Hommais, Catherine Joulian, Evelyne Krin, Aurélie Lieutaud, Didier Lièvreumont, Caroline Michel, Daniel Muller, Philippe Ortet, Caroline Proux, Patricia Siguier, David Roche, Zoé Rouy, Gregory Salvignol, Djamila Slyemi, Emmanuel Talla, Stéphanie Weiss, Jean Weissenbach, Claudine Médigue, Philippe N. Bertin. Structure, Function and Evolution of the *Thiomonas* spp. Genome. *Plos Genetics* accepté.

P. N. Bertin, **A. Heinrich-Salmeron**, E. Pelletier, F. Goulhen-Chollet, F. Arsène-Ploetze, C. Casiot, S. Gallien, B. Chane-Woon-Ming, L. Giloteaux, M. Barakat, V. Bonnefoy, O. Bruneel, M. Chandler, J. Cleiss, R. Duran, F. Elbaz-Poulichet, N. Fonknechten, B. Lauga, D. Mornico, P. Ortet, P. Siguier, A. A. Thil Smith, D. Vallenet, A. Van Dorselaer, J. Weissenbach, C. Médigue, D. Le Paslier. Diversity of trophic interaction inside an arsenic-rich microbial ecosystem. Soumis à *Science*.

A. Heinrich-Salmeron, F. Chollet Goulhen, Philippe Hammann, Didier Lièvreumont, Philippe N. Bertin, M.-C. Lett. Hypertolerance to arsenic and constitutive arsenite-oxidase activity in the *Pseudomonas* sp. Strain S11 isolated from the gold and arsenic mine waste from Salsigne (France). En preparation.

A. Heinrich-Salmeron, Audrey Cordi, Sébastien Gallien, David Halter, Christophe Pagnout, Florence Goulhen-Chollet, Elham Abbaszadehfard, Alain Van Dorselaer, Christine Schaeffer, Philippe Bertin, Pascale Bauda, Florence Arsène-Ploetze. Characterization of an arsenic contaminated environment, Sainte-Marie-aux-Mines, France, revealed a large prokaryotic diversity with arsenic-specific adaptation capacities. En préparation.

Fabienne Battaglia-Brunet, Caroline Michel, Dominique Breeze, **Audrey Heinrich-Salmeron**, Marie-Claire Lett, Catherine Joulian. First evidence of As(III)-oxidation by a strain belonging to

Leptothrix genus: *Leptothrix arsenicoxydans* sp. nov., isolated from arsenic-rich mine water sediment.
En preparation pour soumission à *Antonie van Leeuwenhoek*.

LIVRE

D. Muller, C. Carapito, S. Koechler, S. Weiss, **A. Salmeron**, N. Perdrial, D. Lièvreumont, F. Arsène-Ploetze, P. N. Bertin & M.-C. Lett (2007) « Arsenic stress in bacteria: involvement in the colonization of arsenic-contaminated environments ? in Zhu Y., Lepp N. & Naidu R. eds, Biochemistry of trace elements : environmental protection, remediation and human health. » Tsinghua University Press. p 544-545, ISBN : 978-7-302-15627-7 (livre).

COMMUNICATIONS ORALES

D. Muller, C. Carapito, S. Koechler, S. Weiss, **A. Salmeron**, N. Perdrial, D. Lièvreumont, F. Arsène-Ploetze, Ph. Bertin & M.-C. Lett : « Arsenic stress in bacteria : involvement in the colonization of arsenic-contaminated environments ? ». 9th ICOBTE, Beijing, 15-19 July 2007 (présenté par M.-C. Lett)

Audrey Heinrich-Salmeron. 2^{ème} colloque Métaux et Microorganismes, ESBS Illkirch, le 22/01/2008 « Evolutionary relationships and diversity survey of *aoxB* genes in aerobic arsenite-oxidizing bacteria »

AFFICHES SCIENTIFIQUES

S. Weiss , **A. Salmeron** , D. Muller, S. Koechler, N. Perdrial, M.-C. Lett, D. Lièvreumont, P.N. Bertin. 2006. Synthèse de biofilms chez les bactéries en réponse à la présence d'arsenic. Les Métaux et le vivant : Aspects Chimiques et Biologiques ESBS Strasbourg 2006.

A. Heinrich-Salmeron, D. Lièvreumont, M.-C. Lett, P.N. Bertin. Arsenophilie ou exploration de la diversité des souches résistantes à l'arsenic. Troisième Colloque d'Ecologie Microbienne 2007.

« L'arsenic et le métabolisme bactérien », « Traitement par voie microbienne des eaux contaminées par l'arsenic », Salon Européen de la Recherche et de l'Innovation (SERI), Paris, 2008.

GDR2909 Métabolisme de l'arsenic chez les procaryotes : de la résistance à la détoxification. Congrès BRG, Strasbourg, 2008.

F. Goulhen-Chollet, **A. Heinrich-Salmeron**, O. Bruneel, D. Halter, C. Casiot, S. Gallien, A. Van Dorsselaer, C. Schaeffer, F. Arsène-Ploetze, P.N. Bertin. Metaproteomic analysis of an arsenic-

contaminated site : functional analysis of the community. Congrès Federation of European Microbiological Societies (FEMS), Gothenburg, Suède, 2009.

P.N. Bertin, **A. Heinrich-Salmeron**, E. Pelletier, F. Goulhen-Chollet, F. Arsène-Ploetze, S. Gallien, A. Van Dorselaer, J. Weissenbach, C. Médigue, D. Le Paslier. Diversité fonctionnelle d'une communauté bactérienne par des approches de métagénomique et de métaprotéomique. Association Francophone d'Ecologie Microbienne (AFEM) section biodiversité, Lyon, France, 2009.

A. Cordi, **A. Heinrich-Salmeron**, F. Goulhen Chollet, D. Halter, C. Pagnout, P. Bertin, F. Arsène-Ploetze, P. Bauda. 2009. Diversité et étude des métabolismes microbiens de l'arsenic d'un ancien site minier. Association Francophone d'Ecologie Microbienne (AFEM), Lyon, France, 2009.

Annexe 2

How clays stimulate marine oil
bioremediation



Contents lists available at ScienceDirect

Applied Clay Science

journal homepage: www.elsevier.com/locate/clay

Clay mineral-enhanced bioremediation of marine oil pollution

Laurence N. Warr^{a,*}, Julia N. Perdrial^b, Marie-Claire Lett^c, Audrey Heinrich-Salmeron^c, Mohamed Khodja^d

^a Institut für Geographie und Geologie, Ernst-Moritz-Armdt Universität, F. Ludwig-Jahn-Str. 17A, D-17487 Greifswald, Germany

^b Centre de Géochimie de la Surface (CNRS-ULP), 1 rue Blessig, 67084-Strasbourg, France

^c Laboratoire de Génétique moléculaire, Génomique et Microbiologie, UMR 7156 ULP-CNRS, 28, rue Goethe 67083-Strasbourg, France

^d Sonatrach/Division Centre de Recherche et Développement, Avenue 1 Novembre Boumerdès 35000, Algeria

ARTICLE INFO

Article history:

Received 5 February 2009

Received in revised form 25 September 2009

Accepted 29 September 2009

Available online 14 October 2009

Keywords:

Smectite

Clay minerals

Bacteria

Bioremediation

Marine hydrocarbon pollution

Beach

ABSTRACT

Three year old laboratory mixtures of Prestige oil, Atlantic sea water and various types of clay mineral powder (Na- and Ca-smectites, palygorskite, synthetic mica-montmorillonite and a low defect kaolinite) contained significantly lower concentrations of extracted oil than the non-clay containing control. As no significant quantities of oil could be detected in the clay samples following solvent extraction, and no expansion of the viscous oil-treated smectite interlayers could be detected by X-ray diffraction, the measured hydrocarbon content is suggested to provide a reasonable estimate of the total oil present. The observed decrease in the concentration of oil over the reaction period is attributed largely to bacterial digestion rather than to hydrocarbon adsorption, with denser populations of cultured oleophilic bacteria in the more strongly digested samples. Overall, the smectites with high specific surface areas and cation exchange capacities induced the highest degree of oil breakdown, and also stimulated the digestion of the heavier aromatic and resin-asphaltene compounds. The presence of bivalent cations at interlayer sites and hydrated surfaces, including abundant Mg^{2+} derived from sea water, is suggested to minimize the thickness of the diffuse ion layer and thus enhance both inorganic and organic nutrient supply from mineral surface to bacterial cell. Adding clay minerals to marine oil pollution is therefore suggested as an effective and economically interesting way of enhancing the bacterial digestion of contaminating spills.

© 2009 Published by Elsevier B.V.

1. Introduction

Marine oil spills resulting from industrial pollution, accidents and acts of war continue to pose a major environmental threat to the ecosystem and represent a loss of billions of Euros to the global economy (Massoud et al., 1996; Patin, 1999). After the November 2003 Prestige spillage off the coast of NW Spain, there was little chance of physically removing, dispersing or isolating the now estimated ~63,000 t of leaked crude oil before being washed up along the Atlantic coastline under the rough sea conditions (Whitfield, 2003). In such cases, attempts to clean up the fuel quickly are often limited and restricted to manual removal from sandy beaches or to using hot pressurized water on boulders and man-made structures.

Although little applied in the Prestige disaster, absorbing or dispersing agents can also be used to help clean up oil spillage (Bayat et al., 2005). Well known are the beneficial properties of clay minerals, which are the most abundant minerals formed during water–rock interaction on Earth. Their charged, large specific surface areas ($m^2 g^{-1}$), and interlayer spaces can potentially adsorb a variety of polar and non-polar organic substances as well as toxic metals (Sposito et al., 1999). As a result, clays and clay minerals find thousands of industrial applications

and are increasingly used for environmental protection and remediation purposes (Murray, 2000; Churchman et al., 2006). In hydrocarbon polluted aqueous systems, colloidal sized clays are additionally known to aid dispersion (Meyers and Quinn, 1973; Owens and Lee, 2003). They increase surface area and accelerate the processes of physical and chemical disaggregation by micelle formation and reduction in the size of oil globules.

An additional property of some clay minerals, which is not so well understood, is their ability to increase the rate of bacterial growth and enhance the degradation of hydrocarbon compounds (Stotzky and Rem, 1966; van Loosdrecht et al., 1990; Chaerun and Tazaki, 2005). It has been suggested that the buffering ability of smectites and their ability to adsorb protons released during hydrocarbon breakdown play an important role in maintaining optimal pH conditions and sustaining bacterial growth (Stotzky and Rem, 1966). Some kaolinites have also been shown to increase the rate of oil breakdown by bacterial digestion (Chaerun and Tazaki, 2005). Here, the formation of C–O–Na–Si complexes on the surfaces of bacterial cell walls associated with particle dissolution is suggested to aid nutrient adsorption by the cell and hence stimulate bacterial activity. Although there is general agreement that mineral surfaces provide favorable conditions for bacteria activity, in many cases the precise mechanisms by which this is achieved remain enigmatic and the experimental observations are not always considered to be consistent (van Loosdrecht et al., 1990).

* Corresponding author. Fax: +49 3834 86 4572.

E-mail address: warr@uni-greifswald.de (L.N. Warr).

In this contribution we report on laboratory tests that demonstrate how adding clay minerals with large specific surface areas ($<55\text{ m}^2\text{ g}^{-1}$) and high cation exchange capacities can significantly enhance the degree of bacterial-mediated hydrocarbon breakdown over significant time periods. In our experimental mixtures of clay minerals, viscous Prestige oil and Atlantic sea water, reacting under static, non-photolytic, poorly oxygenated conditions, 6 out of the 9 clay powders investigated showed $<10\%$ of the original hydrocarbon after a 36 month period. Most samples showed significant reduction in all hydrocarbon fractions, including the toxic aromatic and heavier oil components. This study therefore reveals the enormous potential of available clays as a remedy for some of the most environmentally threatening forms of hydrocarbon beach pollution.

2. Materials and experimentation

For the purpose of this investigation, beach oil was sampled from the North Spanish coastline in early January 2003 (El Bocal beach, 3 km north of Santander) following the Prestige spillage. Being highly viscous and strongly fractionated after ~50 days of chemical and physical degradation during transport and deposition, this oil represents the most toxic variety of coastal pollution. The sample was notably enriched in the heavy aromatic, resin and asphaltene components ($>C_{17}$) and contained a mature consortium of natural oleophilic bacteria, and a small amount of mineral matter and salts derived from the beach.

The laboratory tests were designed to simulate static poorly oxygenated sedimentary refuges that represent the toughest environment for natural degradation where toxic oil is known to persist in significant quantities decades after contamination (Peterson et al., 2003). Experimental mixtures of the oil (5 g), clay (5 g) and sea water (40 ml) were prepared in cylindrical glass tubes. Two sets of experiments were mixed in order to test measurement reproducibility and also to enable more analytical techniques to be used. The effect of clay mineral specific surface area and cation exchange capacity (CEC) was investigated by testing the addition of nine diverse well studied clay powders, most of which are available as industrial materials. Each powder is dominated by a specific type of clay mineral and represents bulk material available at economically realistic prices that would be suitable for large scale application.

With the exception of the Na-montmorillonite^a (IBECO-SEAL-80), distributed by IBECO Bentonit Technologie GmbH (<http://www.ibeco-minerals.com>), all powders were obtained from the Clay Mineral Society source collection (<http://www.clays.org/sourceclays/SourceClays.html>). The powdered samples used were various Na- and Ca-dioctahedral montmorillonites (smectites), a synthetic mica-montmorillonite, a trioctahedral Mg-rich smectite (hectorite), a palygorskite (fibrous),

and a low- and high-defect kaolinite (Table 1). Details of the mineralogy, the specific surface area values, and the CEC data for the mineral powders were taken from published literature (van Olphen and Fripat, 1979; Chipera and Bish, 2001; Hofmann et al., 2004).

The nine samples of clay used as untreated raw powder contain mixed mineral assemblages (Chipera and Bish, 2001). The abundance of clay minerals varies between ~50 and 96%. Other minerals present may include calcite, feldspar and quartz, as well as small amounts of other clay minerals (mica, illite, and talc), gypsum, boehmite and anatase. Both kaolinite samples contain ~1% of the phosphorus mineral, crandallite. The Ca-montmorillonite^d (STX-1) sample is also notable in containing 30% microcrystalline opal-CT ($\text{SiO}_2 \cdot n\text{H}_2\text{O}$). The specific surface area and CEC values of the powders range from $10\text{ m}^2\text{ g}^{-1}$ and $2\text{ meq } 100\text{ g}^{-1}$ for the well crystalline kaolinite^h to $134\text{ m}^2\text{ g}^{-1}$ and $140\text{ meq } 100\text{ g}^{-1}$ for the synthetic mica-montmorillonite^e (van Olphen and Fripat, 1979; Hofmann et al., 2004).

The major element composition of the investigated clay mineral powders (Mermut and Cano, 2001; Hofmann et al., 2004) is presented in Table 2 as % oxides. The main components are SiO_2 (between 43 and 70%) and Al_2O_3 (between 11 and 38%), with the exception of hectorite, which contains ~20% MgO and 14% CaO. All samples contain a range of alkaline (Na_2O and $\text{K}_2\text{O} < 2.20\%$) and alkaline earth elements (MgO and CaO $< 7\%$, excluding hectorite), as well as MnO ($< 0.06\%$), TiO_2 ($< 2\%$) and Fe_2O_3 ($< 5\%$). Other elements present in minor amounts are $\text{P}_2\text{O}_5 < 1\%$ and $\text{SO}_3 < 3\%$. The high loss-on-ignition (LOI) values can be attributed to the large amount of hydrous phases stored in the interlayer and intracrystalline sites of the clay mineral phases. The concentrations of trace and rare earth elements for most of the samples used are also known (Kogel and Lewis, 2001), but are here not presented. Notable is the enrichment of the two kaolinite samples in Ni^{2+} , Cu^{2+} , and Pb^{2+} , with concentrations ranging between ~18 and 45 ppm.

A control mixture of beach oil with sea water was also prepared for comparison without the addition of clay. Following initial mixing of the oil–clay–sea water samples to form more homogeneous biofilm-rich sediment, the glass tubes were closed with plastic caps (not air tight) and stored in a dark undisturbed place. The viscous samples were periodically shaken and occasionally stirred over the experimental period to simulate the poorly oxygenated, non-photolytic conditions of poorly-agitated sediments. Loss of water by evaporation was compensated by adding fresh sea water solution.

Due to the strong degree of sample heterogeneity observed throughout most of the reaction time studied, it was not possible to take representative samples. Only after 36 months were the mineral–oil-rich mixtures considered to be well mixed for reliable investigation. At this stage the experiments were ended and samples analyzed to determine changes in solution pH and composition, the basal lattice

Table 1
Mineralogy of the clay mineral powders.

Sample (name)	Source	Mineral assemblage	Clay surface area $\text{m}^2\text{ g}^{-1}$	Clay CEC value $\text{meq } 100\text{ g}^{-1}$
Na-Mnt ^a (IBECO-SEAL-80)	Commercial product	$<80\%$ Mnt, 8–12% cal, $<3\%$ Fsp, 2% other	55	82
Na-Mnt ^b (Swy-2)	Crook County, Wyoming	~75% Mnt, 16% Fsp, 8% Qtz, 1% Gyp, Mca and/or Ill, Kln (?) and/or Chl (?)	32	76
Ca-Mnt ^c (SAz-2)	Cheto Mine, Arizona	~98% Mnt, 1% Qtz, 1% other	97	120
Ca-Mnt ^d (STX-1)	Gonzales County, Texas	~67% Mnt, 30% Opal-CT ($\text{SiO}_2 \cdot n\text{H}_2\text{O}$), 3% Qtz Fsp, Kln, Tlc (?)	84	84
Mica-Mnt ^e (Syn-1)	Synthetic	~95% Mca-Mnt, 5% Bhm	134	140
Hectorite ^f (SHCa-1)	San Bernardino County, California	~50% Htr, 43% Cal, 3% Dol, 3% Qtz, 1% other	63	44
Palygorskite ^g (PFI-1)	Gadsden County, Florida	79% Pkg, 11% Sme, 6% Qtz, 4% Fsp, 1% other	136	20
Kaolinite ^h (KGa-1b)	Washington County, GA	~96% Kln and trace Dic, 3% Ant, 1% Cdl, Qtz (?)	10	2
Kaolinite ⁱ (KGa-2)	Warren County, GA	~96% Kln, 3% Ant, 1% Cdl, Mca and/or Ill	24	3

Mineral assemblages, mineral specific surface (N_2) area ($\text{m}^2\text{ g}^{-1}$) and ammonium determined CEC ($\text{meq } 100\text{ g}^{-1}$) data are taken from published literature (van Olphen and Fripat, 1979; Hofmann et al., 2004) or otherwise determined in our laboratory.

Mnt = Montmorillonite, Cal = Calcite, Fsp = Feldspar, Qtz = Quartz, Gyp = Gypsum, Mca = Mica, Kln = Kaolinite, Tlc = Talc, Bhm = Boehmite, Htr = Hectorite, Dol = Dolomite, Pkg = Palygorskite, Sme = Smectite, Dic = Dickite, Ant = Anatase, Cdl = Crandallite, Ill = Illite.

Table 2
Major element composition (% oxides) of the clay mineral powders.

Sample	Na ₂ O	MgO	Al ₂ O ₃	SiO ₂	P ₂ O ₅	SO ₃	K ₂ O	CaO	MnO	TiO ₂	Fe ₂ O ₃	LOI
Na-Mnt ^a	2.20	3.00	16.60	52.90	0.80	2.90	0.50	6.10	0.10	0.80	4.60	12.00
Na-Mnt ^b	1.50	2.90	22.10	61.50	0.00	0.05	0.20	1.2	0.01	0.10	4.40	5.76
Ca-Mnt ^c	0.05	6.73	19.98	59.65	0.01	–	0.19	3.15	0.10	0.25	1.77	8.17
Ca-Mnt ^d	0.31	3.79	17.86	70.03	0.01	0.04	0.07	1.73	0.01	0.26	1.20	4.63
Mica-Mnt ^e	0.08	0.18	38.56	49.77	0.00	0.10	0.00	0.05	–	0.03	0.12	10.83
Hectorite ^f	1.35	20.01	0.86	46.66	0.00	0.01	0.14	14.01	0.01	0.03	0.32	16.31
Palygorskite ^g	0.05	10.58	11.13	60.35	0.92	0.11	0.87	1.89	0.06	0.48	3.74	9.52
Kaolinite ^h	0.05	0.040	38.58	43.36	0.37	–	0.00	0.04	0.00	1.67	0.35	15.50
Kaolinite ⁱ	0.06	0.040	38.14	43.49	0.32	0.02	0.02	0.03	–	1.91	1.15	15.03

Chemical data Mermut and Cano, 2001; Hofmann et al., 2004 determined either by X-ray fluorescence, atomic absorption spectrometry or flame emission techniques. LOI = lost-on-ignition.

thickness of clay minerals, the concentration of extractable hydrocarbons and the type and abundance of culturable bacteria. Due to the lack of periodic sampling, this study does not provide information on the rates of reactions involved.

3. Analytical techniques

X-ray diffraction (XRD) analyses were made to quantify the lattice thicknesses (d values of the basal 001 reflection) of clay mineral particles. Here, a Bruker D5000 diffractometer was used to irradiate laboratory dried textured films that were prepared by sedimentation to enhance the peak intensities of basal plane reflections. For comparison, both the lattice thicknesses (d values) of sea water treated and oil-sea water treated clay samples were measured from the position of XRD reflections. The differences between these values give an indication of cation exchange reactions or whether or not larger organic complexes are adsorbed in the interlayer sites. Selected measurements of the interlayer thicknesses in the wet state were also made using a wet-cell reaction chamber (Warr and Hofmann, 2003). Results of the interlayer spacing from montmorillonite in a mixture of sea water and oil were used to develop a nanoscale reaction model.

The pH condition of the sampled sea water solutions was determined using a pH meter. The solution composition of diluted samples (1:100) was measured by combining inductively coupled plasma optical emission spectroscopy (ICP-OES) and inductively coupled plasma mass spectrometry (ICP-MS) techniques for quantification of major and trace elements, including heavy metals.

The total concentration of Prestige oil in both solid and liquid phases was determined by infra-red spectrophotometry after centrifugation separation (3000 rpm for 4 min) following solvent extraction using CCl₄ according to the norm procedure NF XPT 90114 (accuracy of ± 0.01 mg kg⁻¹). The measured concentration of solvent extracted oil was confirmed by thermogravimetric analyses and determination of mass loss differences between pure and oil-contaminated clays. These tests, together with electron microscopy study of clay mineral particles, indicate that no significant quantities of oil remained in the

solid clay fractions following solvent extraction. The concentrations of saturated, aromatic, resin and asphaltene phases were also determined on selected samples using thin layer gas chromatography-flame ionization detector (TLC/FID) techniques. For this analysis, hydrocarbon fractions were extracted using *n*-hexane, toluene and dichloromethane solvents (Karlsen and Larter, 1991).

The number of bacteria (g⁻¹ oil) represents a count of colony forming units (CFU) that grew after 10 days on a culture plate containing a saline Luria Bertani medium (Atlas, 1993). Gram types of different bacteria were recognized microscopically after staining. For the colonies grown from the pure Prestige beach oil sample, the identification of the types of bacteria was done by analysis of the 16S rDNA. The total DNA of bacterial strains was isolated using the Wizard[®] Genomic DNA purification kit (Promega). The 16S rDNA fragments were amplified by polymerase chain reaction (PCR) from DNA extracts by using the eubacterial universal primers specific for 16S rDNA: fD1 (5'-CAACAGAGTTTGATCTGGCTCAG-3') and rD1 (5'-GCTTAAGGAGGTGATCCAGCC-3'). The PCR products were purified (QIAquick PCR purification kit, QIAGEN) and sequenced by using the AmpliTaq FS DNA polymerase and BIGDYE TM terminators. Sequence reactions were analyzed with an Applied Biosystems 373XL sequencer. Database searches and sequence analyses were performed using the RDP-II program (Cole et al., 2007).

Direct observations and microchemical analyses of the clay, bacteria and oil components were made at the nanometre scale on selected samples by transmission electron microscopy (TEM) combined with energy dispersive spectroscopy (EDS). Following dialysis to remove salts, dispersed and diluted suspensions of the samples were sedimented onto carbon films that are stretched over a Cu-grid. This investigation used a JEOL JEM 1210 equipped with an OXFORD LINK EDX micro-analyser, housed at the Geology department of the University of Greifswald.

Although the two experimental sets did not allow us to systematically determine the sampling errors involved in our system, we were able to spot test the repetition of analytical results and to determine a general level of data reliability. Such tests revealed that

Table 3
Elemental composition of the 3 year old experimental solutions.

Sample	Na	Mg	Si	P	K	Ca	Mn	Fe	Cu	Zn	Sr	Ba	Total cation
Coastal sea water	5139	1250	0.413	<0.5	367	401	0.02	<0.02	<0.01	0.017	7.81	0.018	7165
Oil – no clay	–	1336	1.35	<0.5	–	490	0.02	<0.02	<0.01	<0.005	8.69	0.024	–
Na-Mnt ^a	5634	1311	6.4	<0.5	331	380	0.45	0.02	<0.01	0.015	6.14	0.090	7669
Na-Mnt ^b	5613	1297	12.4	<0.5	399	598	0.77	<0.02	<0.01	<0.005	14.7	0.110	7935
Ca-Mnt ^c	–	1172	14.4	<0.5	–	443	10.64	0.08	<0.01	<0.005	18.6	0.044	–
Ca-Mnt ^d	5510	1346	18.0	<0.5	–	898	0.22	0.02	<0.01	<0.005	11.1	0.152	–
Mica-Mnt ^e	–	1172	3.25	<0.5	–	443	0.11	–	–	–	7.6	–	–
Hectorite ^f	–	1408	4.24	<0.5	–	491	0.36	0.02	<0.01	<0.005	44.9	0.113	–
Palygorskite ^g	–	1363	1.56	<0.5	–	596	0.22	<0.02	<0.01	<0.005	8.7	0.123	–
Kaolinite ^h	5077	1302	8.28	<0.5	382	518	0.04	<0.02	<0.01	<0.005	8.8	0.077	7296
Kaolinite ⁱ	5088	1279	3.09	<0.5	385	552	0.12	<0.02	<0.01	<0.005	8.8	0.122	7307

Concentration of cations (mg/l) in the oil- and clay mineral-bearing experimental sea water solutions after 36 months of reaction time. The compositions of the untreated sea water and a control sample containing oil but no clay are given for comparison. Underlined numbers reflect differences of <15%.

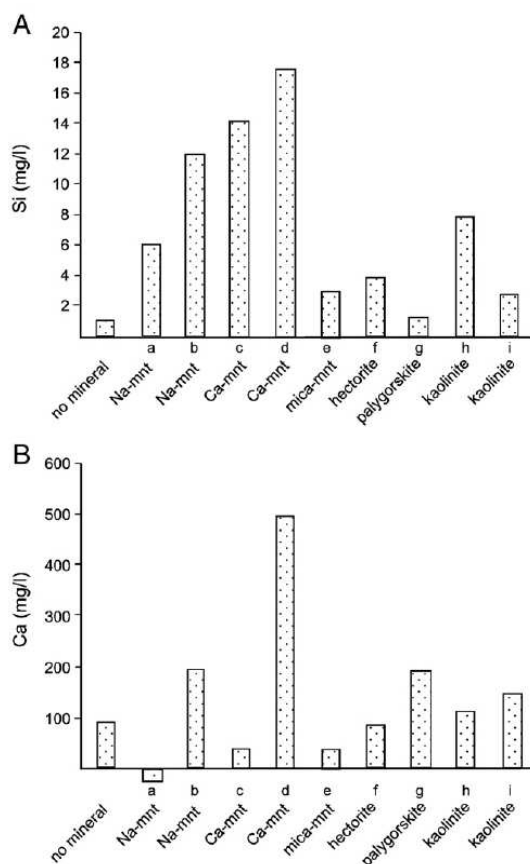


Fig. 1. Concentration (in mg/l) of selected cations in sea water solutions treated with the various clay mineral powders. A) Si⁴⁺ and B) Ca²⁺.

Table 4

Clay mineral, sea water, oil and bacterial characteristics.

Clay mineralogy	<i>d</i> value nm clay-SW (clay-SW-oil)	Solution chemistry		Oil content		Bacteria
		pH ₁ (pH ₂)	Sum of heavy metals ppb V, Ni, Co, Cr, Zn, Pb	Solid g kg ⁻¹ (liquid mg kg ⁻¹)	Total %	CFU mg ⁻¹ oil
No mineral	-	-	88	84.8 (0.6)	67.8	5909
Na-Mnt ^a (IBECO-SEAL-80)	1.541 (1.527)	7.7 (7.3)	20	8.2 (0.4)	6.6	16715
Na-Mnt ^b (Swy-2)	1.402 (1.462)	7.2 (7.2)	20	27.4 (0.6)	21.9	10571
Ca-Mnt ^c (SAZ-2)	1.556 (1.538)	7.0 (6.9)	25	8.0 (0.2)	6.4	19632
Ca-Mnt ^d (STX-1)	1.507 (1.530)	6.9 (7.0)	65	5.5 (1.7)	4.4	16418
Mica-Mnt ^e (Syn-1)	1.210 (1.166)	5.2 (7.3)	56	4.2 (0.4)	3.4	28481
Hectorite ^f (SHCa-1)	1.555 (1.524)	7.4 (7.3)	55	11.4 (0.5)	9.1	8166
Palygorskite ^g (PFI-1)	1.046 (1.050)	7.1 (7.2)	63	11.1 (0.4)	8.9	20548
Kaolinite ^h (KGa-1b)	0.718 (0.718)	6.5 (7.1)	107	26.9 (0.4)	21.5	15872
Kaolinite ⁱ (KGa-2)	0.717 (0.719)	5.9 (7.1)	79	99.1 (0.2)	79.3	6113

d value = basal 001 lattice thickness (nm) of clay minerals (110 for palygorskite) after saturation in sea water (SW) and sea water-oil solutions were measured by X-ray diffraction. pH₁ = of sea water with clay, pH₂ = of sea water with clay and oil. The concentration of heavy metals was measured by ICP-MS. The total concentration of Prestige oil in both solid and liquid phases was determined according to the norm procedure NF XPT 90114 (accuracy of ±0.01 mg kg⁻¹). The number of bacteria (g⁻¹ oil) represents a count of colony forming units (CFU) that grew after 10 days on a culture plate containing a saline Luria Bertani medium (Atlas, 1993).

both sampling and analytical errors combined always lay within the range <15% at 1σ standard deviation. The low precision of the results presented is considered when interpreting the results.

4. Results and discussion

4.1. Solution composition and clay mineralogy

The composition of coastal sea water used in this study is listed in Table 3, together with the sample containing Prestige oil without the presence of clay and the 9 oil- and clay-treated solutions. The composition of the filtered sea water solution used for experimentation is characteristic of shoreline waters diluted by the inflow and mixing of terrestrial waters. Compared to the average composition of non-diluted sea water (Pilson, 1998), the solution contains ~52% less Na and minor but insignificant differences (2.7–5.9% depletion) in the other major cations (Mg²⁺, Ca²⁺, K⁺ and Sr²⁺). Some enrichment in Mn²⁺, Zn²⁺ and Ba²⁺ is evident, which is presumably enhanced by terrestrial contributions.

The oil-contaminated solution (no clay) shows selective enhancement of some cation concentrations compared to the pure coastal sample. This reflects the small amount of mineral and salt contamination observed in the Prestige beach oil sample that could not be effectively removed prior to usage. The enrichment in Si concentration is suggestive of silicate minerals and most likely reflects small amounts of quartz sand contamination derived from the beach. The enhanced amounts of Ca²⁺ and possibly Mg²⁺ come from beach salts concentrated during repeated evaporation of tidal waters.

All clay-treated solutions contain significantly higher concentrations of a number of cations that can be attributed to a combination of mineral dissolution and exchange reactions (Table 3). No precise correlations occur between the concentration of elements in mineral powders and solutions, but some general relationships can be recognized. For example, the Ca-montmorillonite^d with the highest SiO₂ content produced the most Si⁴⁺-rich solution and the hectorite^f and palygorskite^g with the highest MgO released the most Mg²⁺. The lack of correlation between the elemental concentrations of the solids (Table 2)

and the sea water solutions (Table 3) represents the presence of a broad range of mineral phases of varying chemical stability and reactivity.

Despite the complex nature of dissolution and exchange kinetics occurring in these samples, the significant increase in Si^{4+} in all layer silicate dominated clay powders indicates additional dissolution of tetrahedral sheets where this cation is stored (Fig. 1A). In the Ca-montmorillonite^d the high amount of microcrystalline opal-CT present is likely to represent an additional source of Si^{4+} . In contrast, the fibrous palygorskite^e clay showed the least increase, indicating a higher degree of mineral stability in saline solution. Significant enhancements in the concentration of Ca^{2+} , Mn^{2+} , Sr^{2+} and Ba^{2+} occur in most samples (Fig. 1B; Table 3). Due to the polymineralic nature of the samples, it is not possible to identify interlayer cation exchange reactions from these results. The higher concentrations of Ca^{2+} released into solutions subjected to Na-montmorillonite^(b), Ca-montmorillonite^(d) or palygorskite^(e) are of interest (Fig. 1B), but as these clays contain low total concentrations of CaO (<2 wt.%), the presence of calcite and feldspar impurities in the whole rock powders is the probable source of the enhanced Ca^{2+} here detected in solution.

The exchange of interlayer Na^+ cations in smectite layers for bivalent ions (either Ca^{2+} , Mg^{2+} or Sr^{2+}) from the sea water is evident in clay-sea water mixtures from the two-water layer hydration state measured by X-ray diffraction in the air dried condition (range of 1.402–1.556 nm at ~30% relative humidity; Table 4). However, the general lack of additional interlayer expansion with the addition of Prestige oil (range 1.462–1.538 nm; Table 4) and a normal expansive behaviour in ethylene glycol implies that no significant intercalation of hydrocarbon compounds occurred in these reaction sites; a feature confirmed by the lack of carbon detected in chemical analyses of individual grains by TEM. The intercalation of even some of the smallest hydrocarbon molecules such as methanol, ethanol or propanol would cause interlayers to expand to thicknesses larger than >1.6 nm (Brindley et al., 1969). The minor variations detected (± 0.06 nm) are most likely due to small differences in hydration state resulting from largely non-polar, hydrophobic oil coatings of clay particles.

4.2. Oil content

As no significant quantities of organics remained in the solid clay fraction following solvent extraction, the % oil measured is taken to be a reasonable estimate of the total hydrocarbon content. In Fig. 2A, these concentrations are seen to be generally lower in clay mineral samples with high specific surface area and CEC values. These two mineralogical parameters are here plotted separately as the layer charge density of clay minerals is typically heterogeneous and may vary independently of specific surface area. Clay minerals with small specific surface areas ($10\text{--}55\text{ m}^2\text{ g}^{-1}$) and lower CEC values (2–76 meq 100 g^{-1} range) retain between 12 and 80% oil, whereas those with significantly larger specific surface areas ($55\text{--}134\text{ m}^2\text{ g}^{-1}$) and higher CEC values (55–140 meq 100 g^{-1} range) preserve <8% oil. Although the results do not produce good correlations, the experimental changes are best approximated using similar shapes of power law curves. These general relationships imply that the exchange of cations across reactive surfaces plays an important role in the oil breakdown. An interesting exception is the relatively good performance of the low defect kaolinite^b containing vermiform stacked particles, which suggests that the retention of hydrocarbon compounds in micropores may also be aiding the degradation process.

As photolytic reactions can be excluded in our experiments, the significant reduction in hydrocarbons is most likely caused by biodegradation. This interpretation is supported by a correlation between % oil and the number of cultured oleophilic bacterial colonies (g^{-1} oil) observed in Fig. 2B. The higher density of bacteria corresponds well with samples of most extensive degradation and with the clay minerals of larger specific surface area and cation exchange capabilities. However, the ~3 times increase in the number

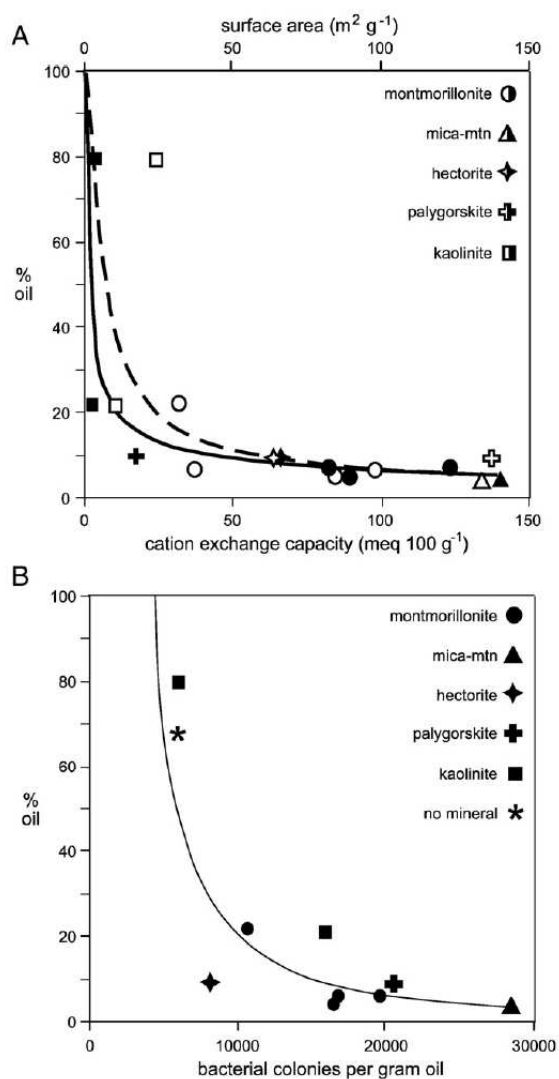


Fig. 2. A) Concentration of remaining oil (expressed as % of original amount) versus clay mineral specific surface area (open circles fitted with dashed line $y = 309.38x^{-0.845}$, $R^2 = 0.58$) and cation exchange capacity (solid circles fitted with solid line $y = 68.3x^{-0.509}$, $R^2 = 0.60$). B) Amount of extracted oil versus number of bacterial colonies (mg^{-1} oil), $y = 1E + 08x^{-1.713}$, $R^2 = 0.71$.

of bacteria cells recorded for the most depleted sample (~3% oil) compared to the least depleted (~80% oil) does not account for all the differences in hydrocarbon concentration. This is primarily attributed to fact that not all oleophilic bacteria species grow on any given culture media and therefore the total population is likely to be underestimated. Secondly, we also quantify the bacteria species cultured after 36 months of activity and therefore have no information on population changes occurring over the complete period of reaction time. Despite the inability to measure the total population of all oleophilic bacteria present in our experiments, the observed trends are considered to reflect the biodegradation process.

Further indication of intense hydrocarbon breakdown by bacteria over the 36 month period is evident from the narrow range of pH

conditions of the sea water solutions (range 6.9–7.3) compared to non-contaminated, sea water–clay mixtures (range 5.2–7.7; Table 4). This stability is suggested to be related to the production of organic pH buffers produced during oil biodegradation, although the presence of such compounds was not analyzed in this study. Microscopic observations and staining tests of the 6 to 8 types of cultured bacteria per sample indicate that each clay–oil mixture is characterized by its own consortia, with varying communities of oil-degrading bacteria. Although these assemblages represent just a small proportion of the true amount of oleophilic species present, the observed differences do imply that varying types of oil-degrading bacterial assemblages develop in the diverse clay environments of the experimental mixtures. Plate testing of the mineral powders in sterile water revealed that these bacteria were not derived from the clay samples used.

4.3. Bacterial analyses

PCR and sequencing of 16SrDNA fragment from 7 types of bacterial colonies cultured from the pure Prestige beach oil sample allowed identification of the well known oil-degrading gram negative *Pseudomonas* sp., together with various gram positive species of *Bacillus*, namely *licheniformis*, *subtilis*, *megaterium* and *firmus* (Fig. 3A–D). A number of the bacteria species seen under the optical microscope contain either gas bubbles in the cell to aid buoyancy or endospores to aid survival. The long cells (up to 10 μm in length) of *Bacillus megaterium* were also notably mobile and contain flagella. The presence of *Alcanivorax*, *Flavobacterium* and *Acinetobacter*, as identified in previously studied Prestige coastal pollution (Alonso et al., 2005; Gallego et al., 2006), were here not detected. The occurrence of *B. licheniformis*, *B. subtilis* and *B. megaterium* is of particular interest in that certain strains have been shown to produce biosurfactants that

reduce the interfacial tension between oil and brine, and hence aid hydrocarbon emulsification (Yakimov et al., 1995).

4.4. Transmission electron microscopy

Observations by TEM of the Ca-montmorillonite^d sample containing oil and bacteria reveal a number of features that reflect the mechanisms of mineral dissolution and hydrocarbon digestion. Coatings of biofilm ~0.5 μm in thickness are commonly observed around bacterial cells (Fig. 4A). The film contains strongly dissolved particles of montmorillonite, as well as nanoparticles that give the biofilm a characteristic mottled appearance. Because of their extremely small size it was not possible to determine whether or not these particles are crystalline or amorphous. As the material typically contains Si^{4+} and occasionally Al^{3+} , Mg^{2+} and Ca^{2+} , they may represent highly leached fragments of montmorillonite or alternatively, a Si-rich gel like material that has been described in other studies of clay mineral–bacteria interaction (Chaerun and Tazaki, 2005). Toward the outer margin of the biofilm, aggregates of montmorillonite and oil occur together. Areas of oil were not always easily visible by TEM, but their distribution could be easily located by elemental mapping of areas of interest. These typically occur as discrete patches or small oil particles attached to montmorillonite grains (Fig. 4A).

Other bacteria observed in the Ca-montmorillonite^d sample appear to lack biofilm and show direct contacts between the basal surfaces of clay particles and the cell membrane (Fig. 4B). The EDX spectra of the montmorillonite particles show higher concentrations of Mg^{2+} and lower Ca^{2+} than present in the original clay material. This indicates that the higher concentration of Mg^{2+} in the sea water has led to exchange of Ca^{2+} in the interlayer sites by this cation. Bacteria cells also contain polyphosphate bodies, which contain some

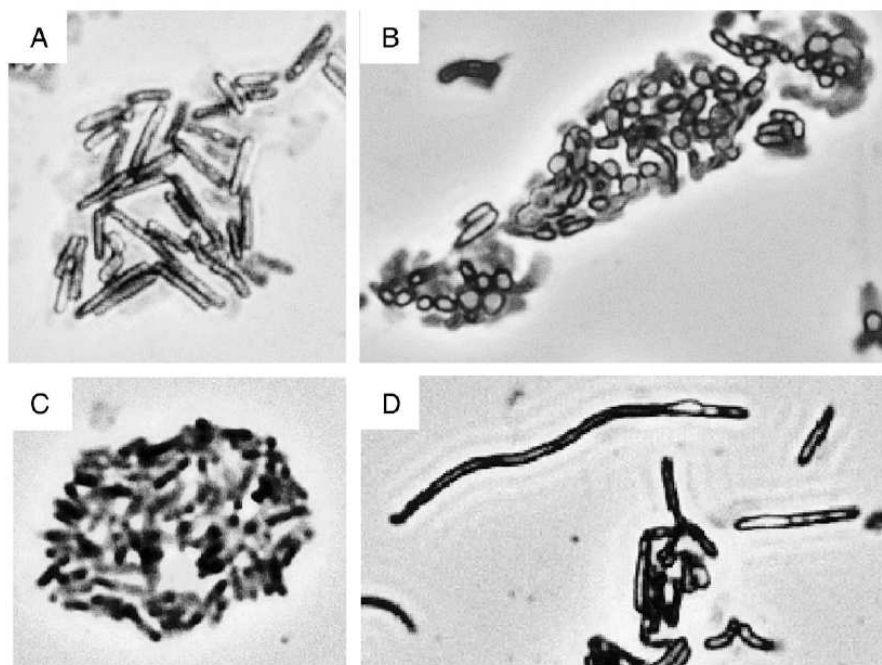


Fig. 3. Optical microscope images of different types of bacteria cultures grown in Luria Bertani medium (after gram staining). A) *Pseudomonas* species (+), B) *Bacillus licheniformis* (+), C) *Bacillus subtilis* (+), D) *Bacillus megaterium* (+). The observed bacteria cells range between 1 and 10 μm in length.

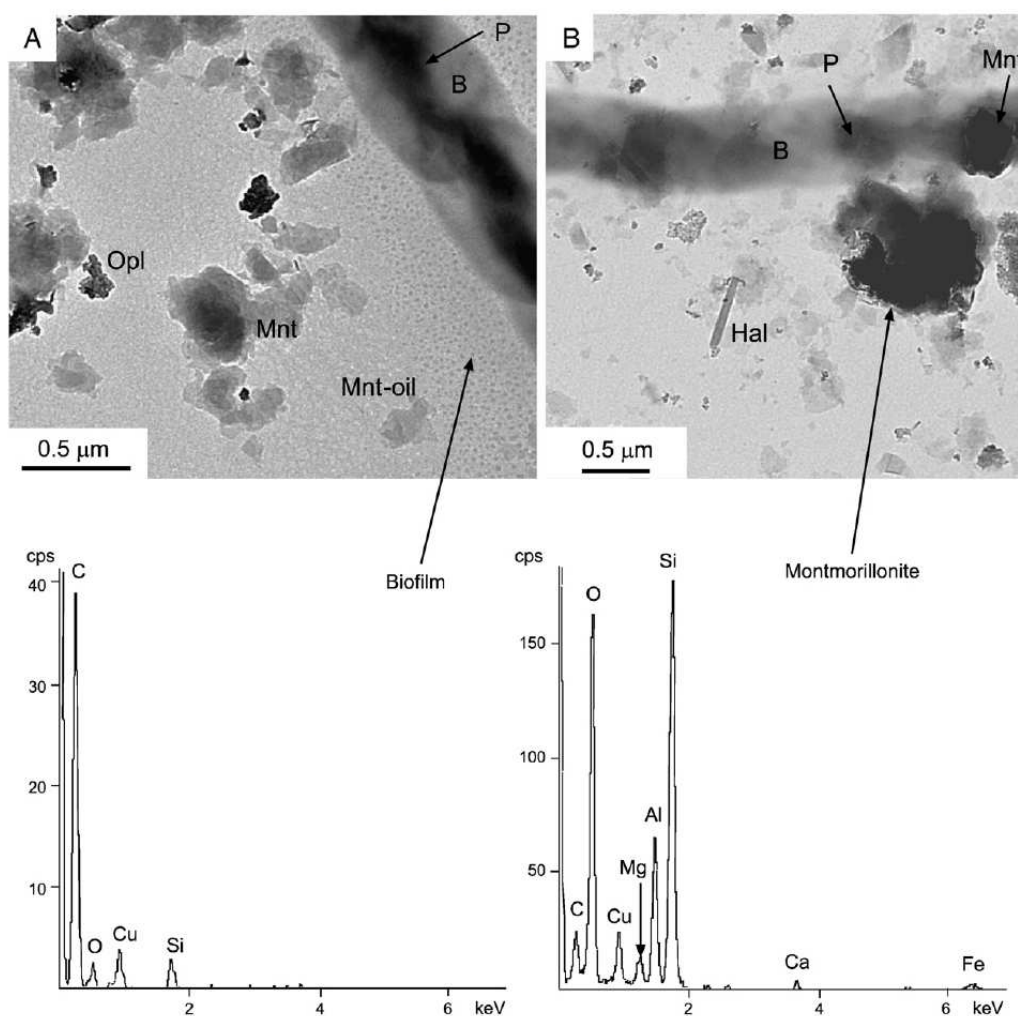


Fig. 4. High resolution transmission electron microscope images and microchemical analyses of the clay, bacteria and oil mixtures of sample Ca-montmorillonite^d (STX-1). A) Image of a long rod shaped bacteria cell surrounded by a coating of biofilm. The EDX pattern of the biofilm is shown below. B) Image of a rod shaped cell with no biofilm and direct contact with montmorillonite particles. The EDX pattern of a typical montmorillonite particle reveals the abundance of Mg²⁺ in the interlayer. B = bacteria, P = phosphorus rich globules occur in the central areas of the cell structure. The presence of Cu is derived from the sample holder grid. Hal = hallyosite, Opl = opal-CT, Mnt = montmorillonite.

Mg²⁺ and Ca²⁺ (Fig. 4), suggestive of metal sequestration and heavy metal detoxification (Baxter and Jensen, 1980; Perdrial et al., 2008).

4.5. Hydrocarbon analyses

To further examine the nature of clay mineral-enhanced bacteria digestion, the concentrations of remaining saturated, aromatic, resin and asphaltene compounds were determined for selected samples. Here, the industrial Na- and Ca-montmorillonites, and the low- and high-defect varieties of kaolinites were analyzed (Fig. 5A). The montmorillonites show successful reduction in all fractions and, in particular, have lower concentrations of heavier resins and asphaltenes. The low defect kaolinite^b sample also performed well in reducing the saturated and aromatic components but was less successful in breaking down heavy fractions. The only clay mineral to degrade less oil than the control sample, with no mineral present, was that of the high-defect kaoliniteⁱ with low specific surface area and CEC properties. This sample managed to reduce successfully the

heavy oil fractions but significantly enhanced the concentrations of both saturated and aromatic compounds.

4.6. Heavy metal content in solution

Consideration of the toxic heavy metals V, Ni, Co, Cr, Zn and Pb detected in sea water solutions indicates no correlation between these elemental concentrations and the degree of hydrocarbon breakdown (Table 4, Fig. 5B). This is clearly emphasized by Wyoming Na-montmorillonite^b and high-defect kaoliniteⁱ both of which show the same degree of oil breakdown (~78%) but have contrasting heavy metal contents in solution. As in most of the adsorbent clay mineral samples, the Na-montmorillonite treated solution shows significant depletion in the measured heavy metals (up to 77% less relative to the pure oil control), whereas the kaolinite sample solution is enriched by ~18%. This is particularly noticeable in the montmorillonite solution concentrations of Ni and Pb, with depletions ranging from -10 to -55 ppb, in contrast to the enrichments of +4 to 12 ppb Pb observed

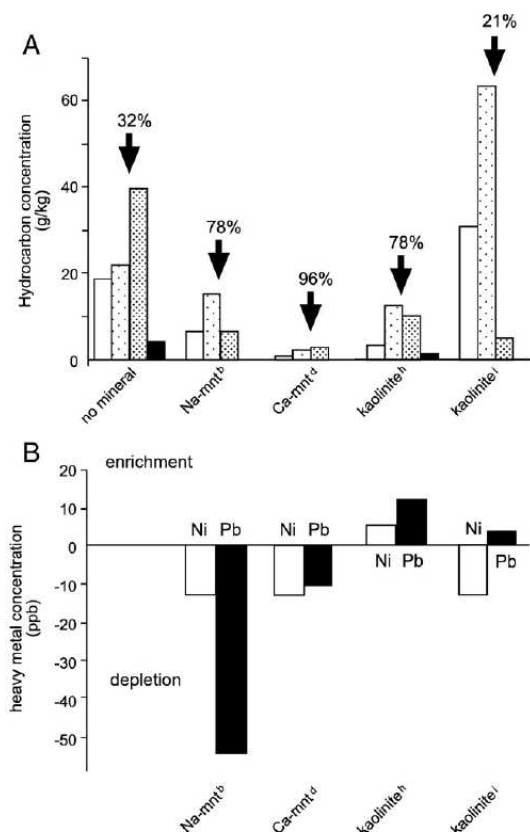


Fig. 5. A) Prestige hydrocarbon fractions measured by thin layer gas chromatography–flame ionization detector after 36 months of bacterial degradation with industrial clay powders. The % of digested oil is calculated based on the initial concentration of 125 g kg^{-1} . B) Concentrations (ppb) of the heavy metals Ni and Pb in solutions of selected clay–oil mixtures, expressed relative to that measured in the pure oil sample. Fields of metal enrichment and depletion are shown.

in kaolinite solutions (Fig. 5B). These elevated heavy metal concentrations can be attributed to metals released during the partial dissolution of kaolinite particles.

4.7. Mechanisms of bacterial stimulation

The sustained bacterial activity (% oil; Fig. 2) in our experiments is not readily explainable by the ability of clay minerals to adsorb protons (Stotzky and Rem, 1966). As the starting pH of the sea water solution was 7.5, most of the experimental mixtures containing clay and oil have lower pH values and thus reflect bulk proton release produced by oil breakdown (Table 4). Although adsorption of these ions on to the surfaces of clay mineral particles is to be expected, which contributes to the buffering process (range 5.9–7.7), the presence and breakdown of hydrocarbon compounds appear to buffer more strongly the solution pH (range 6.9–7.3). This mechanism is evident in all samples irrespective of mineral type, surface area and CEC values. Similarly, the formation of C–O–Na–Si complexes on the surfaces of bacterial cell walls as suggested by Chaerun and Tazaki (2005) may be an active mechanism in these experimental mixtures, with the occurrence of Si-enriched biofilm (Fig. 4). However, this mechanism does not appear consistent with the abundant release of Si^{4+} into solution (Table 3) and does not explain the better performance of oil breakdown by clays with high specific surface areas and CEC capabilities.

Many of the most successful clay minerals in stimulating the bacterial digestion of the Prestige oil are dioctahedral smectites and in particular the Ca-montmorillonite samples. Over 93% of the total concentration of oil originally present in these samples was consumed in the presence of bivalent montmorillonite and presumably incorporated as biomass, released as volatile gases (CO_2 and possibly CH_4) and converted to inorganic carbon. The importance of exchangeable Ca^{2+} is also emphasized by the extensive hydrocarbon breakdown (>90%) observed with palygorskite and hectorite, both of which are abundant in exchangeable bivalent cations.

Based on TEM study of the bacteria, smectite and oil relationships in the Ca-montmorillonite^o sample (Fig. 4A,B), we propose a nanoscale reaction model with bivalent cations (Ca^{2+} and Mg^{2+}) locally bridging negatively charged montmorillonite, bacteria cell and/or polysaccharide biofilm surfaces (Fig. 6). Bivalent cations are known to significantly suppress the electrostatic double layer repulsion between particles and can also induce attractive forces between montmorillonite and negatively charged organic particles of up to 9 mN m^{-1} (Liu et al., 2004). The thickness of the chloride poor diffuse ion layer in sea water is calculated to be <1 nm by DLVO theory (van Loosdrecht et al., 1990) assuming 0.479 M monovalent and 0.063 M bivalent solution. Hamaker composite $6.00\text{E}-21 \text{ J}$, surface potential (zeta) of mineral -0.01 V and bacteria -0.02 V , particle size $2 \mu\text{m}$ and temperature 298 K . The thickness of the hydrated interlayer in sea water was determined experimentally by wet-cell X-ray diffractometry (Warr and Berger, 2007). Surface bivalent Ca^{2+} and Mg^{2+} ions act as local bridges among negatively charged clay minerals, biofilm and bacteria surfaces. The reaction model predicts how trapped and weakly bonded non ionic hydrocarbons can be digested by the adjacent bacterium cell and converted to short chained hydrocarbons with release of volatile components. Hydrocarbons could also be adsorbed as cation- π or charge dipole interactions (Zhu et al., 2004) and toxic heavy metals (X^+)

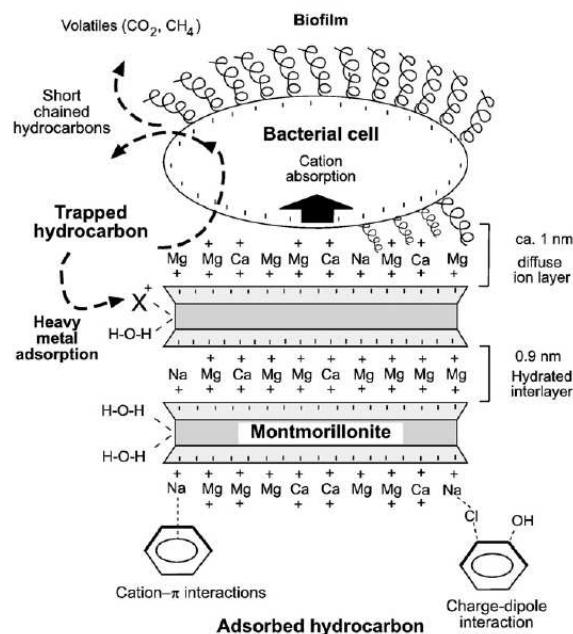


Fig. 6. Proposed mechanism of cation and hydrocarbon assimilation by oleophilic bacteria in the proximity of a weakly charged bivalent-montmorillonite particle surface (bacteria cell not scaled relative to clay particle). The arrow marks the direction of cationic transport to the bacterial cell, which may be sequestered as polyphosphate bodies.

released during the breakdown of the oil can attach to available edge sites (Churchman et al., 2006).

Close juxtaposition of bacteria and the charged mineral surface enables the cell to directly access the nutrient-rich clay mineral surface characterized by i) high concentrations of exchangeable cations, ii) reduced chloride concentrations and iii) weakly bound hydrocarbon complexes. Although many of the clay minerals we studied have the potential of adsorbing and detoxifying toxic hydrocarbons and heavy metals (Cr, Cu, Pb, Ni and V) released during hydrocarbon breakdown (Zhu et al., 2004; Hundal et al., 2001; Andrade et al., 2004–32), our results imply that these mechanisms are not of first importance in the bioremediation of marine spills. Here, we consider that the bivalent cation bridging of charged surfaces, the hydrocarbon retention properties of the porous Ca- or Mg-clays, and the direct contact of bacteria and biofilm communities to nutrient-rich surfaces are primary mechanisms for stimulating the bacterial digestion of heavily fractionated beach oil.

5. Conclusions

1. Experimental mixtures of Prestige oil–clay mineral–sea water mixtures stored for over 3 years in a confined reaction system indicate that the addition of clay mineral powders to beach oil significantly increases the degree of hydrocarbon breakdown.
2. Investigation of the reaction products reveals that the most extensive hydrocarbon breakdown corresponds to the samples that have undergone more bacterial digestion. This breakdown was not restricted to the saturated (alkane) compounds, but includes the heavier aromatic, resin and asphaltene fractions.
3. The electrostatically charged clay mineral surface and in particular the presence of surface bridging bivalent cations (Ca^{2+} , Mg^{2+}) are suggested as two key factors that enhance the supply of cationic nutrients from the mineral surface to the bacterial cell.
4. Despite increasing navigation safety, marine oil spills are expected to continue and new remediation strategies are of major importance. The idea of using industrial clay minerals to accelerate the natural breakdown of hydrocarbons by indigenous bacteria could in the future present a more cost effective alternative to enhance the speed of oil clean up.

Acknowledgements

We thank Fernando Ayllon for his assistance in sampling of the Prestige beach oil. To Amelie Aubert, René Boutin, Jean-Luc Cezard and Fabienne Huber for their support in the laboratory (CGS, Strasbourg). Simone Seefeld is thanked for measuring the composition of the sea water samples and Manfred Zander (IGG, Greifswald) for his help in the TEM investigation.

References

Atlas, R.M., 1993. In: Parks, L. (Ed.), *Handbook of Microbiological Media*. CRC Press, Inc, p. 666.

Alonso, J., Costa, M.M., Nova, B., Albaiges, J., Solanas, A.M., Figueras, A., 2005. Chemical analysis and bacterial diversity associated to marine sediments affected by the Prestige oil-spill. *Vertimar 2005, Symp. Marine Accid. Oil Spills*. extended abstract, <http://otvm.uvigo.es/vertimar2005/orales.html>.

Andrade, M.L., Covelo, E.F., Vega, F.A., Marcet, P., 2004. Effect of the Prestige oil spill on salt marsh soils on the coast of Galicia (Northwestern Spain). *Environ. Qual.* 33, 2103–2210.

Baxter, M., Jensen, T., 1980. Uptake of magnesium, strontium, barium, and manganese by *Plectonema boryanum* (Cyanophyceae) with special reference to polyphosphate bodies. *Protoplasma* 104 (1–2), 81–89.

Bayat, A., Aghamiri, S.F., Moheb, A., Vakili-Nezhaad, G.R., 2005. Oil spill cleanup from sea water by sorbent materials. *Chem. Eng. Technol.* 28 (12), 1525–1528.

Brindley, G.W., Wiewiora, K., Wiewiora, A., 1969. Intracrystalline swelling of montmorillonite in some water–organic mixtures (clay organic studies XVII). *Am. Miner.* 54, 1635–1644.

Chaerun, S.K., Tazaki, K., 2005. How kaolinite plays an essential role in remediating oil-polluted sea water. *Clay Miner.* 40, 481–491.

Chipera, S.J., Bish, D.L., 2001. Baseline studies of the Clay Minerals Society source clays: X-ray diffraction analyses. *Clays & Clay Miner.* 49 (5), 398–409.

Churchman, G.J., Gates, W.P., Theng, B.K.G., Yuan, G., 2006. Clays and clay minerals for pollution control. In: Bergaya, F., Theng, B.K.G., Lagaly, G. (Eds.), *Handbook of Clay Science: Developments in Clay Science*, vol. 1, pp. 625–675.

Cole, J.R., Chai, B., Farris, R.J., Wang, Q., Kulam-Syed-Mohideen, A.S., McCarrell, D.M., Bandela, A.M., Cardenas, E., Garrity, G.M., Tiedje, J.M., 2007. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Res.* 35, D169–D172 (Database issue).

Gallego, J.R., González-Rojas, E., Peláez, A.J., Sánchez, J., García-Martínez, M.J., Llamas, J.F., 2006. Natural attenuation and bioremediation of Prestige fuel oil along the Atlantic coast of Galicia (Spain). *Organ. Geochem.* 37, 1869–1884.

Hofmann, H., Bauer, A., Warr, L.N., 2004. Behavior of smectite in strong salt brines under conditions relevant to the disposal of low- to medium-grade nuclear waste. *Clays & Clay Miner.* 52 (1), 14–24.

Hundal, L.S., Thompson, M.L., Laird, D., Carmo, M., 2001. Sorption of phrenanthrene by reference smectite. *Environ. Sci. Technol.* 35, 3456–3461.

Karlsen, D.A., Larter, S.R., 1991. Analysis of petroleum fractions by TLC/FID: applications to petroleum reservoir description. *Organ. Geochem.* 17, 603–617.

Kogel, J.E., Lewis, S.A., 2001. Baseline studies of the Clay Minerals Society source clays: chemical analysis by inductively coupled plasma-mass spectroscopy (ICP-MS). *Clays & Clay Miner.* 49, No. 5, 387–392.

Owens, E.H., Lee, K., 2003. Interaction of oil and mineral fines on shorelines: review and assessment. *Marine Pollution Bulletin* 47 (9–12), 397–405.

Patin, S., 1999. Environmental impact of the offshore oil and gas industry. *Ecomonitor Publication*, p. 448.

Perdrial, N., Liewig, N., Delphin, J.-E., Elsass, F., 2008. TEM evidence for intracellular accumulation of lead by bacteria in subsurface environments. *Chem. Geol.* 253 (3–4), 196–204.

Peterson, C.H., Rice, S., Short, J.W., Esler, D., Bodkin, J.L., Ballachey, B.E., Irons, D.B., 2003. Long-term ecosystem responses to the Exxon Valdez oil spill. *Science* 302, 2082–2086.

Massoud, M.S., AlAbdali, F., AlGhadban, A.N., AlSarawi, M., 1996. Bottom sediments of the Arabian Gulf. 2. TPH and TOC contents as indicators of oil pollution and implications for the effect and fate of the Kuwait oil slick. *Environ. Poll.* 93, 271–284.

Mermut, A.R., Cano, A.F., 2001. Baseline studies of the Clay Minerals Society source clays: chemical analyses of major elements. *Clays & Clay Miner.* 49 (5), 381–386.

Meyers, P.A., Quinn, J.G., 1973. Association of hydrocarbons and mineral particles in saline solutions. *Nature* 244, 23–24.

Murray, H., 2000. Traditional and new applications for kaolin, smectite and palygorskite: a general overview. *Appl. Clay Sci.* 17, 207–221.

Sposito, G., Skipper, N.T., Sutton, R., Park, S.-A., Soper, A.K., Greathouse, J.A., 1999. Surface geochemistry of the clay minerals. *Proc. Natl. Acad. Sci. U. S. A.* 96, 3358–3364.

Stotzky, G., Rem, L.T., 1966. Influence of clay minerals on microorganisms I. Montmorillonite and kaolinite on bacteria. *Can. J. Microbiol.* 12, 547–563.

van Loosdrecht, M.C.M., Lyklema, J., Norde, W., Zehnder, J.B., 1990. Influence of interfaces on microbial activity. *Microbiol. Rev.* 54, 75–87.

van Olphen, H., Fripat, J.J. (Eds.), 1979. *Data Handbook for Clay Materials and other Non-Metallic Minerals*. Pergamon Press, New York, 346pp.

Warr, L.N., Hofmann, H., 2003. In-situ monitoring of powder reactions in percolating solution by wet-cell X-ray diffraction techniques. *J. Appl. Cryst.* 36, 948–949.

Whitfield, J., 2003. How to clean a beach. *Nature* 422, 464–466.

Pilson, M.E.Q., 1998. *An Introduction to the Chemistry of the Sea*. Prentice-Hall, Inc, Upper Saddle River, NJ, 431pp.

Yakimov, M.M., Timmis, K.N., Wray, V., Fredrickson, H.L., 1995. Characterization of a new lipopeptide surfactant produced by thermotolerant and halo-tolerant subsurface *Bacillus licheniformis* BASS0. *Appl. Environ. Microbiol.* 1706–1713.

Liu, J., Xu, Z., Masliyah, J., 2004. Role of fine clays in bitumen extraction from oil sands. *AIChE Journal* 50, 1917–1927.

Warr, L.N., Berger, H., 2007. Hydration of bentonite in natural waters: application of “confined volume” wet-cell X-ray diffractometry. *Phys. Chem. Earth* 32, 247–258.

Zhu, D., Herbert, B.E., Schautman, A., Carraway, E.R., Hur, J., 2004. Cation–π bonding: a new perspective on the sorption of polycyclic aromatic hydrocarbons to mineral surfaces. *Environ. Qual.* 33, 1322–1330.

Bibliographie

- Abernathy, C. O., Thomas, D. J. & Calderon, R. L. (2003).** Health effects and risk assessment of arsenic. *J Nutr* **133**, 1536S-1538S.
- Achour, A. R., Bauda, P. & Billard, P. (2007).** Diversity of arsenite transporter genes from arsenic-resistant soil bacteria. *Res Microbiol* **158**, 128-137.
- Afkar, E., Lisak, J., Saltikov, C., Basu, P., Oremland, R. S. & Stolz, J. F. (2003).** The respiratory arsenate reductase from *Bacillus selenitireducens* strain MLS10. *FEMS Microbiology Letters* **226**, 107-112.
- Ahmann, D., Krumholz, L. R., Hemond, H. F., Lovley, D. R. & Morel, F. M. M. (1997).** Microbial mobilization of arsenic from sediments of the Aberjona watershed. *Environmental Science and Technology* **31**, 2923-2930.
- Allen, E. E. & Banfield, J. F. (2005).** Community genomics in microbial ecology and evolution. *Nature Reviews Microbiology* **3**, 489-498.
- Baker, B. J. & Banfield, J. F. (2003).** Microbial communities in acid mine drainage. *FEMS Microbiol Ecol* **44**, 139-152.
- Baker, B. J., Tyson, G. W., Webb, R. I., Flanagan, J., Hugenholtz, P., Allen, E. E. & Banfield, J. F. (2006).** Lineages of acidophilic archaea revealed by community genomic analysis. *Science* **314**, 1933-1935.
- Baker, B. J., Tyson, G. W., Goosherst, L. & Banfield, J. F. (2009).** Insights into the diversity of eukaryotes in acid mine drainage biofilm communities. *Appl Environ Microbiol* **75**, 2192-2199.
- Banfield, J. F., Tyson, G. W., Allen, E. E. & Whitaker, R. J. (2005a).** The search for a molecular-level understanding of the processes that underpin the Earth's biogeochemical cycles. In *Reviews in Mineralogy and Geochemistry*, pp. 1-7.
- Banfield, J. F., Verberkmoes, N. C., Hettich, R. L. & Thelen, M. P. (2005b).** Proteogenomic approaches for the molecular characterization of natural microbial communities. *OMICS A Journal of Integrative Biology* **9**, 301-333.
- Battaglia-Brunet, F., Dictor, M. C., Garrido, F., Crouzet, C., Morin, D., Dekeyser, K., Clarens, M. & Baranger, P. (2002).** An arsenic(III)-oxidizing bacterial population: Selection, characterization, and performance in reactors. *J Appl Microbiol* **93**, 656-667.
- Battaglia-Brunet, F., Itard, Y., Garrido, F., Delorme, F., Crouzet, C., Greffié, C. & Joulain, C. (2006a).** A Simple Biochemical Process Removing Arsenic from a Mine Drainage Water. *Geomicrobiology Journal* **23**, 201-211.
- Battaglia-Brunet, F., Joulain, C., Garrido, F., Dictor, M. C., Morin, D., Coupland, K., Barrie Johnson, D., Hallberg, K. B. & Baranger, P. (2006b).** Oxidation of arsenite by *Thiomonas* strains and characterization of *Thiomonas arsenivorans* sp. nov. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology* **89**, 99-108.
- Battaglia-Brunet, F., Dictor, M. C., Lett, M. C. & Bertin, P. N. (2008).** Bioremédiation d'eaux contaminées par l'arsenic. *Technique de l'ingénieur*.

- Batzoglou, S., Jaffe, D. B., Stanley, K., Butler, J., Gnerre, S., Mauceli, E., Berger, B., Mesirov, J. P. & Lander, E. S. (2002).** ARACHNE: a whole-genome shotgun assembler. *Genome Res* **12**, 177-189.
- Baymann, F., Lebrun, E., Brugna, M., Schoepp-Cothenet, B., Giudici-Orticoni, M. T. & Nitschke, W. (2003).** The redox protein construction kit: pre-last universal common ancestor evolution of energy-conserving enzymes. *Philos Trans R Soc Lond B Biol Sci* **358**, 267-274.
- Beloin, C., Ayora, S., Exley, R., Hirschbein, L., Ogasawara, N., Kasahara, Y., Alonso, J. C. & Le Hégarat, F. (1997).** Characterization of an Irp-like (IrpC) gene from *Bacillus subtilis*. *Molecular and General Genetics* **256**, 63-71.
- Benndorf, D., Balcke, G. U., Harms, H. & von Bergen, M. (2007).** Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *Isme J* **1**, 224-234.
- Bentley, R. & Chasteen, T. G. (2002).** Microbial methylation of metalloids: arsenic, antimony, and bismuth. *Microbiol Mol Biol Rev* **66**, 250-271.
- Bertin, P. N., Médigue, C. & Normand, P. (2008).** Advances in environmental genomics: Towards an integrated view of micro-organisms and ecosystems. *Microbiology* **154**, 347-359.
- Bhattacharjee, H., Choudhury, R. & Rosen, B. P. (2008).** Role of conserved aspartates in the ArsA ATPase. *Biochemistry* **47**, 7218-7227.
- Bhattacharya, P., Welch, A. H., Stollenwerk, K. G., McLaughlin, M. J., Bundschuh, J. & Panullah, G. (2007).** Arsenic in the environment: Biology and Chemistry. *Science of the Total Environment* **379**, 109-120.
- Biers, E. J., Sun, S. & Howard, E. C. (2009).** Prokaryotic Genomes and Diversity in Surface Ocean Waters: Interrogating the Global Ocean Sampling Metagenome. *Appl Environ Microbiol* **75**, 2221-2229.
- Bissen, M. & Frimmel, F. H. (2003a).** Arsenic - A review. Part I: Occurrence, toxicity, speciation, mobility. *Acta Hydrochimica et Hydrobiologica* **31**, 9-18.
- Bissen, M. & Frimmel, F. H. (2003b).** Arsenic - A review. Part II: Oxidation of arsenic and its removal in water treatment. *Acta Hydrochimica et Hydrobiologica* **31**, 97-107.
- Blum, J. S., Han, S., Lanoil, B. & other authors (2009).** Ecophysiology of "*Halarsenatibacter silvermanii*" strain SLAS-1T, gen. nov., sp. nov., a facultative chemoautotrophic arsenate respirer from salt-saturated Searles Lake, California. *Appl Environ Microbiol* **75**, 1950-1960.
- Branco, R., Francisco, R., Chung, A. P. & Morais, P. V. (2009).** Identification of an *aox* system that requires cytochrome c in the highly arsenic resistant strain *Ochrobactrum tritici* SCII24. *Appl Environ Microbiol*.
- Bruneel, O. (2004).** Contribution à l'étude des mécanismes couplés géochimiques et bactériologiques de transfert de la pollution minière sur le site de Carnoulès (Gard). Thèse
- Bruneel, O., Personné, J. C., Casiot, C., Leblanc, M., Elbaz-Poulichet, F., Mahler, B. J., Le Flèche, A. & Grimont, P. A. D. (2003).** Mediation of arsenic oxidation by *Thiomonas* sp. in acid-mine drainage (Carnoulès, France). *J Appl Microbiol* **95**, 492-499.

- Bruneel, O., Duran, R., Casiot, C., Elbaz-Poulichet, F. & Personné, J. C. (2006).** Diversity of microorganisms in Fe-As-rich acid mine drainage waters of Carnoulès, France. *Appl Environ Microbiol* **72**, 551-556.
- Bruneel, O., Pascault, N., Egal, M., Bancon-Montigny, C., Goñi-Urriza, M. S., Elbaz-Poulichet, F., Personné, J. C. & Duran, R. (2008).** Archaeal diversity in a Fe-As rich acid mine drainage at Carnoulès (France). *Extremophiles* **12**, 563-571.
- Bryan, C. G., Marchal, M., Battaglia-Brunet, F., Kugler, V., Lemaitre-Guillier, C., Lievremont, D., Bertin, P. N. & Arsene-Ploetze, F. (2009).** Carbon and arsenic metabolism in *Thiomonas* strains: differences revealed diverse adaptation processes. *BMC Microbiol* **9**, 127.
- Butcher, B. G., Deane, S. M. & Rawlings, D. E. (2000).** The chromosomal arsenic resistance genes of *Thiobacillus ferrooxidans* have an unusual arrangement and confer increased arsenic and antimony resistance to *Escherichia coli*. *Appl Environ Microbiol* **66**, 1826-1833.
- Butcher, B. G. & Rawlings, D. E. (2002).** The divergent chromosomal ars operon of *Acidithiobacillus ferrooxidans* is regulated by an atypical ArsR protein. *Microbiology* **148**, 3983-3992.
- Cai, L., Liu, G., Rensing, C. & Wang, G. (2009a).** Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. *BMC microbiology* **9**, 4.
- Cai, L., Rensing, C., Li, X. & Wang, G. (2009b).** Novel gene clusters involved in arsenite oxidation and resistance in two arsenite oxidizers: *Achromobacter* sp. SY8 and *Pseudomonas* sp. TS44. *Appl Microbiol Biotechnol*, 1-11.
- Canovas, D., Cases, I. & de Lorenzo, V. (2003).** Heavy metal tolerance and metal homeostasis in *Pseudomonas putida* as revealed by complete genome analysis. *Environ Microbiol* **5**, 1242-1256.
- Carlin, A., Shi, W., Dey, S. & Rosen, B. P. (1995).** The *ars* operon of *Escherichia coli* confers arsenical and antimonial resistance. *Journal of Bacteriology* **177**, 981-986.
- Casiot, C., Leblanc, M., Bruneel, O., Personné, J. C., Koffi, K. & Elbaz-Poulichet, F. (2003a).** Geochemical processes controlling the formation of As-rich waters within a tailings impoundment (Carnoulès, France). *Aquatic Geochemistry* **9**, 273-290.
- Casiot, C., Morin, G., Juillot, F., Bruneel, O., Personné, J. C., Leblanc, M., Duquesne, K., Bonnefoy, V. & Elbaz-Poulichet, F. (2003b).** Bacterial immobilization and oxidation of arsenic in acid mine drainage (Carnoulès creek, France). *Water Res* **37**, 2929-2936.
- Casiot, C., Bruneel, O., Personné, J. C., Leblanc, M. & Elbaz-Poulichet, F. (2004).** Arsenic oxidation and bioaccumulation by the acidophilic protozoan, *Euglena mutabilis*, in acid mine drainage (Carnoulès, France). *Science of the Total Environment* **320**, 259-267.
- Caspi, R., Foerster, H., Fulcher, C. A. & other authors (2008).** The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* **36**, D623-631.
- Chang, J. S., Yoon, I. H., Lee, J. H., Kim, K. R., An, J. & Kim, K. W. (2009).** Arsenic detoxification potential of aox genes in arsenite-oxidizing bacteria isolated from natural and constructed wetlands in the Republic of Korea. *Environ Geochem Health*.

- Charlet, L. & Polya, D. A. (2006).** Arsenic in shallow, reducing groundwaters in Southern Asia: An environmental health disaster. *Elements* **2**, 91-96.
- Chatain, V., Bayard, R., Sanchez, F., Moszkowicz, P. & Gourdon, R. (2005).** Effect of indigenous bacterial activity on arsenic mobilization under anaerobic conditions. *Environ Int* **31**, 221-226.
- Chauhan, N. S., Ranjan, R., Purohit, H. J., Kalia, V. C. & Sharma, R. (2009).** Identification of genes conferring arsenic resistance to *Escherichia coli* from an effluent treatment plant sludge metagenomic library. *FEMS Microbiol Ecol* **67**, 130-139.
- Cheng, H., Hu, Y., Luo, J., Xu, B. & Zhao, J. (2009).** Geochemical processes controlling fate and transport of arsenic in acid mine drainage (AMD) and natural systems. *J Hazard Mater* **165**, 13-26.
- Chivian, D., Brodie, E. L., Alm, E. J. & other authors (2008).** Environmental genomics reveals a single-species ecosystem deep within earth. *Science* **322**, 275-278.
- Choi, S., O'Day, P. A. & Hering, J. G. (2009).** Natural attenuation of arsenic by sediment sorption and oxidation. *Environmental Science and Technology* **43**, 4253-4259.
- Chowdhury, U. K., Biswas, B. K., Chowdhury, T. R. & other authors (2000).** Groundwater arsenic contamination in Bangladesh and West Bengal, India. *Environmental Health Perspectives* **108**, 393-397.
- Cohen, S. M., Ohnishi, T., Arnold, L. L. & Le, X. C. (2007).** Arsenic-induced bladder cancer in an animal model. *Toxicology and Applied Pharmacology* **222**, 258-263.
- DeLong, E. F. (2002).** Microbial population genomics and ecology. *Current Opinion in Microbiology* **5**, 520-524.
- Dey, S. & Rosen, B. P. (1995).** Dual mode of energy coupling by the oxyanion-translocating ArsB protein. *Journal of Bacteriology* **177**, 385-389.
- Dinsdale, E. A., Edwards, R. A., Hall, D. & other authors (2008).** Functional metagenomic profiling of nine biomes. *Nature* **452**, 629-632.
- Duker, A. A., Carranza, E. J. M. & Hale, M. (2005).** Arsenic geochemistry and health. *Environ Int* **31**, 631-641.
- Duquesne, K., Lebrun, S., Casiot, C., Bruneel, O., Personné, J. C., Leblanc, M., Elbaz-Poulichet, F., Morin, G. & Bonnefoy, V. (2003).** Immobilization of Arsenite and Ferric Iron by *Acidithiobacillus ferrooxidans* and Its Relevance to Acid Mine Drainage. *Appl Environ Microbiol* **69**, 6165-6173.
- Duquesne, K., Lieutaud, A., Ratouchniak, J., Yarz-bal, A. & Bonnefoy, V. (2007).** Mechanisms of arsenite elimination by *Thiomonas* sp. isolated from Carnoulès acid mine drainage. *European Journal of Soil Biology* **43**, 351-355.
- Duquesne, K., Lieutaud, A., Ratouchniak, J., Muller, D., Lett, M. C. & Bonnefoy, V. (2008).** Arsenite oxidation by a chemoautotrophic moderately acidophilic *Thiomonas* sp.: From the strain isolation to the gene study. *Environmental Microbiology* **10**, 228-237.
- Ellis, P. J., Conrads, T., Hille, R. & Kuhn, P. (2001).** Crystal structure of the 100 kDa arsenite oxidase from *Alcaligenes faecalis* in two crystal forms at 1.64 Å and 2.03 Å. *Structure* **9**, 125-132.

- Fleischmann, R. D., Adams, M. D., White, O., & other authors (1995).** Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.
- Gailer, J. (2007).** Arsenic-selenium and mercury-selenium bonds in biology. *Coordination Chemistry Reviews* **251**, 234-254.
- Garelick, H., Jones, H., Dybowska, A. & Valsami-Jones, E. (2008).** Arsenic pollution sources. In *Reviews of Environmental Contamination and Toxicology*, pp. 17-60.
- Gihring, T. M. & Banfield, J. F. (2001).** Arsenite oxidation and arsenate respiration by a new *Thermus* isolate. *FEMS Microbiology Letters* **204**, 335-340.
- Gilbert, J. A., Field, D., Huang, Y., Edwards, R., Li, W., Gilna, P. & Joint, I. (2008).** Detection of Large Numbers of Novel Sequences in the Metatranscriptomes of Complex Marine Microbial Community. *PLoS One* **3**, e3042.
- Goulhen-Chollet, F., Halter, D., Bruneel, O. & other authors. (2009).** *Euglena* sp. adaptative response and its persistent metabolic activity in acid mine drainage microbial community. *Appl Environ Microbiol*. Submitted.
- Green, H. H. (1918).** Isolation and description of a bacterium causing oxidation of arsenite to arsenate in cattle-dipping baths. *Rep Dir Vet Res S Afr* **6**, 593-599.
- Haferburg, G. & Kothe, E. (2007).** Microbes and metals: interactions in the environment. *J Basic Microbiol* **47**, 453-467.
- Hamamura, N., Macur, R. E., Korf, S., Ackerman, G., Taylor, W. P., Kozubal, M., Reysenbach, A. L. & Inskeep, W. P. (2009).** Linking microbial oxidation of arsenic with detection and phylogenetic analysis of arsenite oxidase genes in diverse geothermal environments. *Environmental Microbiology* **11**, 421-431.
- Handelsman, J. (2004).** Metagenomics: Application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews* **68**, 669-685.
- Haouari, O., Fardeau, M. L., Cayol, J. L., Casiot, C., Elbaz-Poulichet, F., Hamdi, M., Joseph, M. & Ollivier, B. (2008).** *Desulfotomaculum hydrothermale* sp. nov., a thermophilic sulfate-reducing bacterium isolated from a terrestrial Tunisian hot spring. *International Journal of Systematic and Evolutionary Microbiology* **58**, 2529-2535.
- Hardouin, J., Duchateau, M., Joubert-Caron, R. & Caron, M. (2006).** Usefulness of an integrated microfluidic device (HPLC-Chip-MS) to enhance confidence in protein identification by proteomics. *Rapid Communications in Mass Spectrometry* **20**, 3236-3244.
- Harrison, J. J., Ceri, H., Roper, N. J., Badry, E. A., Sproule, K. M. & Turner, R. J. (2005).** Persister cells mediate tolerance to metal oxyanions in *Escherichia coli*. *Microbiology* **151**, 3181-3195.
- Harrison, J. J., Ceri, H. & Turner, R. J. (2007).** Multimetal resistance and tolerance in microbial biofilms. *Nat Rev Microbiol* **5**, 928-938.

- Hasona, A., Ray, R. M. & Shanmugam, K. T. (1998).** Physiological and genetic analyses leading to identification of a biochemical role for the *moeA* (molybdate metabolism) gene product in *Escherichia coli*. *J Bacteriol* **180**, 1466-1472.
- He, Z., Xiao, S., Xie, X. & other authors (2007).** Molecular diversity of microbial community in acid mine drainages of Yunfu sulfide mine. *Extremophiles* **11**, 305-314.
- Hedges, R. W. & Baumberg, S. (1973).** Resistance to arsenic compounds conferred by a plasmid transmissible between strains of *Escherichia coli*. *Journal of Bacteriology* **115**, 459-460.
- Hei, T. K., Liu, S. U. X. & Waldren, C. (1998).** Mutagenicity of arsenic in mammalian cells: Role of reactive oxygen species. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 8103-8107.
- Hoefl, S. E., Blum, J. S., Stolz, J. F., Tabita, F. R., Witte, B., King, G. M., Santini, J. M. & Oremland, R. S. (2007).** *Alkalilimnicola ehrlichii* sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. *International Journal of Systematic and Evolutionary Microbiology* **57**, 504-512.
- Hopenhayn, C. (2006).** Arsenic in drinking water: Impact on human health. *Elements* **2**, 103-107.
- Hoven, R. N. V., Santini, J. M. & Streimann, I. C. A. (2004).** Prokaryotic arsenic metabolism - Roles in arsenic cycling and bioremediation. In *Australasian Institute of Mining and Metallurgy Publication Series*, pp. 93-103.
- Hu, J., Fang, J., Dong, Y., Chen, S. J. & Chen, Z. (2005).** Arsenic in cancer therapy. *Anti-Cancer Drugs* **16**, 119-127.
- Hughes, M. F. (2002).** Arsenic toxicity and potential mechanisms of action. *Toxicology Letters* **133**, 1-16.
- Inskeep, W. P., Macur, R. E., Hamamura, N., Warelw, T. P., Ward, S. A. & Santini, J. M. (2007).** Detection, diversity and expression of aerobic bacterial arsenite oxidase genes. *Environmental Microbiology* **9**, 934-943.
- Islam, F. S., Gault, A. G., Boothman, C., Polya, D. A., Chamok, J. M., Chatterjee, D. & Lloyd, J. R. (2004).** Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature* **430**, 68-71.
- Jerez, C. A. (2008).** The use of genomics, proteomics and other OMICS technologies for the global understanding of biomining microorganisms. *Hydrometallurgy* **94**, 162-169.
- Ji, G. & Silver, S. (1992a).** Reduction of arsenate to arsenite by the ArsC protein of the arsenic resistance operon of *Staphylococcus aureus* plasmid pI258. *Proc Natl Acad Sci U S A* **89**, 9474-9478.
- Ji, G. & Silver, S. (1992b).** Regulation and expression of the arsenic resistance operon from *Staphylococcus aureus* plasmid pI258. *Journal of Bacteriology* **174**, 3684-3694.
- Johnson, D. B. & Hallberg, K. B. (2003).** The microbiology of acidic mine waters. *Research in Microbiology* **154**, 466-473.

- Jones, C. A., Langner, H. W., Anderson, K., McDermott, T. R. & Inskeep, W. P. (2000).** Rates of microbially mediated arsenate reduction and solubilization. *Soil Science Society of America Journal* **64**, 600-608.
- Kalia, K. & Flora, S. J. (2005).** Strategies for safe and effective therapeutic measures for chronic arsenic and lead poisoning. *J Occup Health* **47**, 1-21.
- Kashyap, D. R., Botero, L. M., Franck, W. L., Hassett, D. J. & McDermott, T. R. (2006).** Complex regulation of arsenite oxidation in *Agrobacterium tumefaciens*. *Journal of Bacteriology* **188**, 1081-1088.
- Koechler, S., Cleiss, J., Proux, C. & other authors (2009).** Transcriptional and post-transcriptional controls affects arsenite oxidase activity in *Herminiimonas arsenicoxydans*. *Molecular Microbiology* submitted.
- Koffi, K., Leblanc, M., Jourde, H., Casiot, C., Pistre, S., Gouze, P. & Elbaz-Poulichet, F. (2002).** Reverse oxidation zoning at a mine tailing stck generating arsenic-rich acid waters (Carnoulès, France). *Journal of the International mine Water Association* **22**, 7-14.
- Krafft, T. & Macy, J. M. (1998).** Purification and characterization of the respiratory arsenate reductase of *Chrysiogenes arsenatis*. *European Journal of Biochemistry* **255**, 647-653.
- Kumar, S. R., Parameswaran, V., Ahmed, V. P., Musthaq, S. S. & Hameed, A. S. (2007).** Protective efficiency of DNA vaccination in Asian seabass (*Lates calcarifer*) against *Vibrio anguillarum*. *Fish Shellfish Immunol* **23**, 316-326.
- Langley, S. & Beveridge, T. J. (1999).** Metal binding by *Pseudomonas aeruginosa* PAO1 is influenced by growth of the cells as a biofilm. *Can J Microbiol* **45**, 616-622.
- Leblanc, M., Achard, B., Othman, D. B., Luck, J. M., Bertrand-Sarfati, J. & Personné, J. C. (1996).** Accumulation of arsenic from acidic mine waters by ferruginous bacterial accretions (stromatolites). *Applied Geochemistry* **11**, 541-554.
- LeBlanc, M., Casiot, C., Elbaz-Poulichet, F. & Personné, C. (2002).** Arsenic removal by oxidizing bacteria in a heavily arsenic-contaminated acid mine drainage system (Carnoulès, France). In *Geological Society Special Publication*, pp. 267-274.
- Lebrun, E., Brugna, M., Baymann, F., Muller, D., Lièvremont, D., Lett, M. C. & Nitschke, W. (2003).** Arsenite oxidase, an ancient bioenergetic enzyme. *Molecular Biology and Evolution* **20**, 686-693.
- Lievremont, D., Bertin, P. N. & Lett, M. C. (2009).** Arsenic in contaminated waters: Biogeochemical cycle, microbial metabolism and biotreatment processes. *Biochimie*.
- Lin, Y. F., Walmsley, A. R. & Rosen, B. P. (2006).** An arsenic metallochaperone for an arsenic detoxification pump. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15617-15622.
- Lin, Y. F., Yang, J. & Rosen, B. P. (2007a).** ArsD residues Cys12, Cys13, and Cys18 form an As(III)-binding site required for arsenic metallochaperone activity. *Journal of Biological Chemistry* **282**, 16783-16791.

- Lin, Y. F., Yang, J. & Rosen, B. P. (2007b).** ArsD: An As(III) metallochaperone for the ArsAB As(III)-translocating ATPase. *J Bioenerg Biomembr* **39**, 453-458.
- Liu, Z., Boles, E. & Rosen, B. P. (2004a).** Arsenic Trioxide Uptake by Hexose Permeases in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **279**, 17312-17318.
- Liu, Z., Carbrey, J. M., Agre, P. & Rosen, B. P. (2004b).** Arsenic trioxide uptake by human and rat aquaglyceroporins. *Biochemical and Biophysical Research Communications* **316**, 1178-1185.
- Lloyd, J. R. & Oremland, R. S. (2006).** Microbial transformations of arsenic in the environment: From soda lakes to aquifers. *Elements* **2**, 85-90.
- Lorenz, P. & Eck, J. (2005).** Metagenomics and industrial applications. *Nature Reviews Microbiology* **3**, 510-516.
- Lovley, D. R. (2000).** Environmental Microbe-Metal Interactions. *American Society for Microbiology, Washington, Lovley Ed.*
- MacLean, D., Jones, J. D. & Studholme, D. J. (2009).** Application of 'next-generation' sequencing technologies to microbial genetics. *Nat Rev Microbiol* **7**, 287-296.
- Macy, J. M., Santini, J. M., Pauling, B. V., O'Neill, A. H. & Sly, L. I. (2000).** Two new arsenate/sulfate-reducing bacteria: Mechanisms of arsenate reduction. *Archives of Microbiology* **173**, 49-57.
- Malasarn, D., Saltikov, C. W., Campbell, K. M., Santini, J. M., Hering, J. G. & Newman, D. K. (2004).** *arrA* is a reliable marker for As(V) respiration. *Science* **306**, 455.
- Mandal, B. K. & Suzuki, K. T. (2002).** Arsenic round the world: a review. *Talanta* **58**, 201-235.
- Matschullat, J. (2000).** Arsenic in the geosphere - A review. *Science of the Total Environment* **249**, 297-312.
- Meng, Y. L., Liu, Z. & Rosen, B. P. (2004).** As(III) and Sb(III) Uptake by GlpF and Efflux by ArsB in *Escherichia coli*. *Journal of Biological Chemistry* **279**, 18334-18341.
- Meyer, N., Helynck, B., Ledrans, M., Le Goaster, C., Kintz, P. & Michel, A. (1999).** Evaluation of biological impregnation of a population exposed to high arsenic concentration in drinking water, Ferrette (France), 1997. *Evaluation de l'impregnation biologique d'une population exposee a une concentration elevee en arsenic dans les eaux de distribution, Ferrette, 1997* **47**, 315-321.
- Michel, C., Jean, M., Coulon, S., Dictor, M. C., Delorme, F., Morin, D. & Garrido, F. (2007).** Biofilms of As(III)-oxidising bacteria: Formation and activity studies for bioremediation process development. *Appl Microbiol Biotechnol* **77**, 457-467.
- Mirete, S., de Figueras, C. G. & Gonzalez-Pastor, J. E. (2007).** Novel nickel resistance genes from the rhizosphere metagenome of plants adapted to acid mine drainage. *Appl Environ Microbiol* **73**, 6001-6011.
- Moran, M. A. (2009).** Metatranscriptomics: Eavesdropping on complex microbial communities. *Microbe* **4**, 329-335.

- Morin, G., Juillot, F., Casiot, C., Bruneel, O., Personné, J. C., Elbaz-Poulichet, F., Leblanc, M., Ildefonse, P. & Calas, G. (2003).** Bacterial formation of tooeleite and Mixed Arsenic(III) or Arsenic(V) - Iron(III) gels in the carnoulès acid mine drainage, France. A XANES, XRD, and SEM study. *Environmental Science and Technology* **37**, 1705-1712.
- Mukhopadhyay, R. & Rosen, B. P. (2002).** Arsenate reductases in prokaryotes and eukaryotes. *Environmental Health Perspectives* **110**, 745-748.
- Mukhopadhyay, R., Rosen, B. P., Phung, L. T. & Silver, S. (2002).** Microbial arsenic: From geocycles to genes and enzymes. *FEMS Microbiology Reviews* **26**, 311-325.
- Muller, D. (2004).** Analyse génétique et moléculaire du stress arsenic de souches bactériennes isolées d'environnements contaminés à l'arsenic. Thèse.
- Muller, D., Lièvreumont, D., Simeonova, D. D., Hubert, J. C. & Lett, M. C. (2003).** Arsenite oxidase *aox* genes from a metal-resistant β -proteobacterium. *Journal of Bacteriology* **185**, 135-141.
- Muller, D., Médigue, C., Koechler, S. & other authors (2007).** A tale of two oxidation states: Bacterial colonization of arsenic-rich environments. *PLoS Genetics* **3**, 0518-0530.
- Newman, D. K., Beveridge, T. J. & Morel, F. M. M. (1997a).** Precipitation of arsenic trisulfide by *Desulfotomaculum auripigmentum*. *Appl Environ Microbiol* **63**, 2022-2028.
- Newman, D. K., Kennedy, E. K., Coates, J. D., Ahmann, D., Ellis, D. J., Lovley, D. R. & Morel, F. M. M. (1997b).** Dissimilatory arsenate and sulfate reduction in *Desulfotomaculum auripigmentum* sp. nov. *Archives of Microbiology* **168**, 380-388.
- Neyt, C., Iriarte, M., Thi, V. H. & Cornelis, G. R. (1997).** Virulence and arsenic resistance in *Yersinia*. *Journal of Bacteriology* **179**, 612-619.
- Nies, D. H. & Silver, S. (1995).** Ion efflux systems involved in bacterial metal resistances. *Journal of Industrial Microbiology* **14**, 186-199.
- Nikolaev, Y. A. & Plakunov, V. K. (2007).** Biofilm-"city of microbes" or an analogue of multicellular organisms? *Microbiology* **76**, 125-138.
- O'Day, P. A. (2006).** Chemistry and mineralogy of arsenic. *Elements* **2**, 77-83.
- Oremland, R. S., Hoefft, S. E., Santini, J. M., Bano, N., Hollibaugh, R. A. & Hollibaugh, J. T. (2002).** Anaerobic oxidation of arsenite in Mono Lake water and by a facultative, arsenite-oxidizing chemoautotroph, strain MLHE-1. *Appl Environ Microbiol* **68**, 4795-4802.
- Oremland, R. S. & Stolz, J. F. (2003).** The ecology of arsenic. *Science* **300**, 939-944.
- Oremland, R. S., Kulp, T. R., Blum, J. S., Hoefft, S. E., Baesman, S., Miller, L. G. & Stolz, J. F. (2005a).** A microbial arsenic cycle in a salt-saturated, extreme environment. *Science* **308**, 1305-1308.
- Oremland, R. S., Kulp, T. R., Blum, J. S., Hoefft, S. E., Baesman, S., Miller, L. G. & Stolz, J. F. (2005b).** Microbiology: A microbial arsenic cycle in a salt-saturated, extreme environment. *Science* **308**, 1305-1308.

- Oremland, R. S. & Stolz, J. F. (2005). Arsenic, microbes and contaminated aquifers. *Trends in Microbiology* **13**, 45-49.
- Osborne, F. H. & Ehrlich, H. L. (1976). Oxidation of arsenite by a soil isolate of *Alcaligenes*. *Journal of Applied Bacteriology* **41**, 295-305.
- Páez-Espino, D., Tamames, J., De Lorenzo, V. & Cánovas, D. (2009). Microbial responses to environmental arsenic. *BioMetals* **22**, 117-130.
- Park, S. J., Kang, C. H., Chae, J. C. & Rhee, S. K. (2008). Metagenome microarray for screening of fosmid clones containing specific genes. *FEMS Microbiol Lett* **284**, 28-34.
- Patel, P. C., Goulhen, F., Boothman, C., Gault, A. G., Charnock, J. M., Kalia, K. & Lloyd, J. R. (2007). Arsenate detoxification in a Pseudomonad hypertolerant to arsenic. *Archives of Microbiology* **187**, 171-183.
- Phillips, S. E. & Taylor, M. L. (1976). Oxidation of arsenite to arsenate by *Alcaligenes faecalis*. *Appl Environ Microbiol* **32**, 392-399.
- Pi, J., Yamauchi, H., Kumagai, Y., Sun, G., Yoshida, T., Aikawa, H., Hopenhayn-Rich, C. & Shimojo, N. (2002). Evidence for induction of oxidative stress caused by chronic exposure of Chinese residents to arsenic contained in drinking water. *Environmental Health Perspectives* **110**, 331-336.
- Pope, P. B. & Patel, B. K. (2008). Metagenomic analysis of a freshwater toxic cyanobacteria bloom. *FEMS Microbiol Ecol* **64**, 9-27.
- Poretzky, R.S., Hewson, I., Sun, S., Allen, A. E., Zehr, J. P., Moran, M. A. (2009). Comparative day/night metatranscriptomic analysis of microbial communities in the north pacific subtropical gyre. *Environ Microbiol* **11**, 1358-75.
- Qin, J., Rosen, B. P., Zhang, Y., Wang, G., Franke, S. & Rensing, C. (2006). Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2075-2080.
- Qin, J., Lehr, C. R., Yuan, C., Le, X. C., McDermott, T. R. & Rosen, B. P. (2009). Biotransformation of arsenic by a yellowstone thermoacidophilic eukaryotic alga. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5213-5217.
- Quéméneur, M., Heinrich-Salmeron, A., Muller, D., Lièvremon, D., Jauzein, M., Bertin, P. N., Garrido, F. & Joulain, C. (2008). Diversity surveys and evolutionary relationships of *aoxB* genes in aerobic arsenite-oxidizing bacteria. *Appl Environ Microbiol* **74**, 4567-4573.
- Radabaugh, T. R. & Aposhian, H. V. (2000). Enzymatic reduction of arsenic compounds in mammalian systems: Reduction of arsenate to arsenite by human liver arsenate reductase. *Chemical Research in Toxicology* **13**, 26-30.
- Ram, R. J., Verberkmoes, N. C., Thelen, M. P., Tyson, G. W., Baker, B. J., Blake, R. C., 2nd, Shah, M., Hettich, R. L. & Banfield, J. F. (2005). Community proteomics of a natural microbial biofilm. *Science* **308**, 1915-1920.

- Ramírez-Solís, A., Mukopadhyay, R., Rosen, B. P. & Stemmler, T. L. (2004).** Experimental and Theoretical Characterization of Arsenite in Water: Insights into the Coordination Environment of As-O. *Inorganic Chemistry* **43**, 2954-2959.
- Rhine, E. D., Garcia-Dominguez, E., Phelps, C. D. & Young, L. Y. (2005).** Environmental microbes can speciate and cycle arsenic. *Environmental Science and Technology* **39**, 9569-9573.
- Rhine, E. D., Phelps, C. D. & Young, L. Y. (2006).** Anaerobic arsenite oxidation by novel denitrifying isolates. *Environmental Microbiology* **8**, 899-908.
- Richey, C., Chovanec, P., Hoeft, S. E., Oremland, R. S., Basu, P. & Stolz, J. F. (2009).** Respiratory arsenate reductase as a bidirectional enzyme. *Biochemical and Biophysical Research Communications* **382**, 298-302.
- Riesenfeld, C. S., Schloss, P. D. & Handelsman, J. (2004).** Metagenomics: Genomic analysis of microbial communities. In *Annual Review of Genetics*, pp. 525-552.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. & Seraphin, B. (1999).** A generic protein purification method for complex characterisation and proteom. *Nat Biotechnol* **17**, 1030-1032
- Rosen, B. P. (1999).** Families of arsenic transporters. *Trends in Microbiology* **7**, 207-212.
- Rosen, B. P., Bhattacharjee, H., Zhou, T. & Walmsley, A. R. (1999).** Mechanism of the ArsA ATPase. *Biochimica et Biophysica Acta - Biomembranes* **1461**, 207-215.
- Rosen, B. P. (2002a).** Biochemistry of arsenic detoxification. *FEBS Letters* **529**, 86-92.
- Rosen, B. P. (2002b).** Transport and detoxification systems for transition metals, heavy metals and metalloids in eukaryotic and prokaryotic microbes. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology* **133**, 689-693.
- Rosen, B. P. & Liu, Z. (2009).** Transport pathways for arsenic and selenium: A minireview. *Environ Int* **35**, 512-515.
- Roy, P. & Saha, A. (2002).** Metabolism and toxicity of arsenic: A human carcinogen. *Current Science* **82**, 38-45.
- Rusch, D. B., Halpern, A. L., Heidelberg & other authors. (2007).** The SorcererII Global Ocean Sampling Expedition: I, The Northwest Atlantic through the Eastern Tropical Pacific. *PLoS Biol* **5**, 398-431.
- Ruan, X., Bhattacharjee, H. & Rosen, B. P. (2006).** Cys-113 and Cys-422 form a high affinity metalloid binding site in the ArsA ATPase. *Journal of Biological Chemistry* **281**, 9925-9934.
- Salmassi, T. M., Venkateswaren, K., Satomi, M., Nealson, K. H., Newman, D. K. & Hering, J. G. (2002).** Oxidation of arsenite by *Agrobacterium albertimagni*, AOL15, sp. nov., isolated from Hot Creek, California. *Geomicrobiology Journal* **19**, 53-66.
- Santini, J. M., Sly, L. I., Schnagl, R. D. & Macy, J. M. (2000).** A new chemolithoautotrophic arsenite-oxidizing bacterium isolated from a gold mine: Phylogenetic, physiological, and preliminary biochemical studies. *Appl Environ Microbiol* **66**, 92-97.

- Santini, J. M., Sly, L. I., Wen, A., Comrie, D., De Wulf-Durand, P. & Macy, J. M. (2002a).** New arsenite-oxidizing bacteria isolated from Australian gold mining environments - Phylogenetic relationships. *Geomicrobiology Journal* **19**, 67-76.
- Santini, J. M., Stolz, J. F. & Macy, J. M. (2002b).** Isolation of a new arsenate-respiring bacterium - Physiological and phylogenetic studies. *Geomicrobiology Journal* **19**, 41-52.
- Santini, J. M. & Vanden Hoven, R. N. (2004).** Molybdenum-Containing Arsenite Oxidase of the Chemolithoautotrophic Arsenite Oxidizer NT-26. *Journal of Bacteriology* **186**, 1614-1619.
- Schloss, P. D. & Handelsman, J. (2003).** Biotechnological prospects from metagenomics. *Current opinion Biotechnology* **14**, 303-310.
- Schoen, A., Beck, B., Sharma, R. & Dubé, E. (2004).** Arsenic toxicity at low doses: Epidemiological and mode of action considerations. *Toxicology and Applied Pharmacology* **198**, 253-267.
- Schrenk, M. O., Edwards, K. J., Goodman, R. M., Hamers, R. J. & Banfield, J. F. (1998).** Distribution of *Thiobacillus ferrooxidans* and *Leptospirillum ferrooxidans*: Implications for generation of acid mine drainage. *Science* **279**, 1519-1522.
- Sharma, V. K. & Sohn, M. (2009).** Aquatic arsenic: Toxicity, speciation, transformations, and remediation. *Environ Int* **35**, 743-759.
- Shi, J., Vlamis-Gardikas, A., Øslund, F., Holmgren, A. & Rosen, B. P. (1999).** Reactivity of glutaredoxins 1, 2, and 3 from *Escherichia coli* shows that glutaredoxin 2 is the primary hydrogen donor to ArsC-catalyzed arsenate reduction. *Journal of Biological Chemistry* **274**, 36039-36042.
- Silver, S., Budd, K., Leahy, K. M., Shaw, W. V., Hammond, D., Novick, R. P., Willsky, G. R., Malamy, M. H. & Rosenberg, H. (1981).** Inducible plasmid-determined resistance to arsenate, arsenite, and antimony (III) in *Escherichia coli* and *Staphylococcus aureus*. *J Bacteriol* **146**, 983-996.
- Silver, S. & Phung le, T. (2005).** A bacterial view of the periodic table: genes and proteins for toxic inorganic ions. *J Ind Microbiol Biotechnol* **32**, 587-605.
- Silver, S. & Phung, L. T. (2005).** Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. *Appl Environ Microbiol* **71**, 599-608.
- Smedley, P. L. & Kinniburgh, D. G. (2002).** A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry* **17**, 517-568.
- Stewart, P. S. (1994).** Biofilm accumulation model that predicts antibiotic resistance of *Pseudomonas aeruginosa* biofilms. *Antimicrob Agents Chemother* **38**, 1052-1058.
- Stolz, J. F., Ellis, D. J., Blum, J. S., Ahmann, D., Lovley, D. R. & Oremland, R. S. (1999).** *Sulfurospirillum barnesii* sp. nov. and *Sulfurospirillum arsenophilum* sp. nov., new members of the Sulfurospirillum clade of the ϵ -Proteobacteria. *International Journal of Systematic Bacteriology* **49**, 1177-1180.
- Stolz, J. F. & Oremland, R. S. (1999).** Bacterial respiration of arsenic and selenium. *FEMS Microbiology Reviews* **23**, 615-627.

- Stolz, J. F., Basu, P., Santini, J. M. & Oremland, R. S. (2006).** Arsenic and selenium in microbial metabolism. In *Annual Review of Microbiology*, pp. 107-130.
- Streit, W. R. & Schmitz, R. A. (2004).** Metagenomics - The key to the uncultured microbes. *Current Opinion in Microbiology* **7**, 492-498.
- Teitzel, G. M. & Parsek, M. R. (2003).** Heavy metal resistance of biofilm and planktonic *Pseudomonas aeruginosa*. *Appl Environ Microbiol* **69**, 2313-2320.
- Tringe, S. G., Von Mering, C., Kobayashi, A. & other authors (2005).** Comparative metagenomics of microbial communities. *Science* **308**, 554-557.
- Tyson, G. W., Chapman, J., Hugenholtz, P. & other authors (2004).** Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**, 37-43.
- Valenzuela, L., Chi, A., Beard, S., Orell, A., Guiliani, N., Shabanowitz, J., Hunt, D. F. & Jerez, C. A. (2006).** Genomics, metagenomics and proteomics in biomining microorganisms. *Biotechnology Advances* **24**, 197-211.
- Vallenet, D., Labarre, L., Rouy, Z. & other authors (2006).** MaGe: A microbial genome annotation system supported by synteny results. *Nucleic Acids Research* **34**, 53-65.
- Vaughan, D. J. (2006).** Arsenic. *Elements* **2**, 71-75.
- Venter, J. C., Remington, K., Heidelberg, J. F. & other authors (2004).** Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* **304**, 66-74.
- Wang, L., Jeon, B., Sahin, O. & Zhang, Q. (2009).** Identification of an arsenic resistance and arsenic-sensing system in *Campylobacter jejuni*. *Appl Environ Microbiol* **75**, 5064-5073.
- Weeger, W., Lièvremon, D., Perret, M., Lagarde, F., Hubert, J. C., Leroy, M. & Lett, M. C. (1999).** Oxidation of arsenite to arsenate by a bacterium isolated from an aquatic environment. *BioMetals* **12**, 141-149.
- White, A. K. & Metcalf, W. W. (2004).** The *htx* and *ptx* operons of *Pseudomonas stutzeri* WM88 are new members of the Pho regulon. *Journal of Bacteriology* **186**, 5876-5882.
- Williamson, S. J., Rusch, D. B., Yooseph, S. & other authors (2008).** The Sorcerer II Global Ocean Sampling Expedition: metagenomic characterization of viruses within aquatic microbial samples. *PLoS ONE* **3**, e1456.
- Wilmes, P. & Bond, P. L. (2009).** Microbial community proteomics: elucidating the catalysts and metabolic mechanisms that drive the Earth's biogeochemical cycles. *Curr Opin Microbiol* **12**, 310-317.
- Woyke, T., Teeling, H., Ivanova, N. N. & other authors (2006).** Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature* **443**, 950-955.
- Wu, J. & Rosen, B. P. (1993).** Metalloregulated expression of the *ars* operon. *J Biol Chem* **268**, 52-58.
- Wysocki, R., Bobrowicz, P. & Ulaszewski, S. (1997).** The *Saccharomyces cerevisiae* ACR3 gene encodes a putative membrane protein involved in arsenite transport. *J Biol Chem* **272**, 30061-30066.

- Wysocki, R., Chéry, C. C., Wawrzycka, D., Van Hulle, M., Cornelis, R., Thevelein, J. M. & Tamás, M. J. (2001).** The glycerol channel Fps1p mediates the uptake of arsenite and antimonite in *Saccharomyces cerevisiae*. *Molecular Microbiology* **40**, 1391-1401.
- Xu, C., Shi, W. & Rosen, B. P. (1996a).** The chromosomal *arsR* gene of *Escherichia coli* encodes a trans-acting metalloregulatory protein. *Journal of Biological Chemistry* **271**, 2427-2432.
- Xu, C., Shi, W. & Rosen, B. P. (1996b).** Identification of the As(III)/Sb(III) binding domain and the dimerization of the ArsR repressor. *FASEB Journal* **10**.
- Yamauchi, H. & Fowler, B. A. (1994).** Toxicity and metabolism of inorganic and methylated arsenicals. *Arsenic in the environment Part II: Human Health and Ecosystem Effects* New York: John Wiley In: Nriagu JO, ed, 35-53.
- Ye, J., Yang, H. C., Rosen, B. P. & Bhattacharjee, H. (2007).** Crystal structure of the flavoprotein ArsH from *Sinorhizobium meliloti*. *FEBS Letters* **581**, 3996-4000.
- Yoshida, T., Yamauchi, H. & Fan Sun, G. (2004).** Chronic health effects in people exposed to arsenic via the drinking water: Dose-response relationships in review. *Toxicology and Applied Pharmacology* **198**, 243-252.
- Yuan, C., Lu, X., Jie, Q., Rosen, B. P. & Le, X. C. (2008).** Volatile arsenic species released from *Escherichia coli* expressing the *asIII* S-adenosylmethionine methyltransferase gene. *Environmental Science and Technology* **42**, 3201-3206.
- Zehr, J. P., Bench, S. R., Carter, B. J., Hewson, I., Niazi, F., Shi, T., Tripp, H. J. & Affourtit, J. P. (2008).** Globally distributed uncultivated oceanic N₂-fixing cyanobacteria lack oxygenic photosystem II. *Science* **322**, 1110-1112.
- Zhao, C. Q., Young, M. R., Diwan, B. A., Coogan, T. P. & Waalkes, M. P. (1997).** Association of arsenic-induced malignant transformation with DNA hypomethylation and aberrant gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 10907-10912.
- Zhou, T., Radaev, S., Rosen, B. P. & Gatti, D. L. (2000).** Structure of the ArsA ATPase: The catalytic subunit of a heavy metal resistance pump. *EMBO Journal* **19**, 4838-4845.

DIVERSITE DES MECANISMES D'OXYDO-REDUCTION DE L'ARSENIC CHEZ LES BACTERIES : DES SOUCHES CULTIVABLES AUX COMMUNAUTES

Les micro-organismes sont retrouvés dans des milieux très différents allant des plus communs aux plus extrêmes, comme par exemple les environnements arséniés des sites miniers de Carnoulès, de Salsigne ou de Sainte-Marie-aux-Mines. De nombreuses bactéries sont impliquées dans le cycle bio-géochimique de l'arsenic et jouent un rôle crucial dans sa mobilisation/immobilisation et par conséquent dans sa biodisponibilité. Certaines possèdent l'arsénite-oxydase (Aox), codée par les gènes *aox* et qui catalyse l'oxydation de l'As III en As V, ce qui constitue la première étape de détoxification des milieux arséniés.

L'objectif de ce travail est d'étudier la diversité des mécanismes d'oxydo-réduction de l'arsenic chez les bactéries. Mon travail s'est principalement focalisé sur l'analyse de la diversité des gènes *aox* observée chez des isolats en condition de laboratoire et au sein de communautés complexes en conditions *in situ*.

Ces travaux ont conduit dans un premier temps à la caractérisation physiologique de trois souches : *Pseudomonas* S11, *Leptothrix* S1.1 et *Thiomonas* 3As. Une analyse génétique a complété l'analyse physiologique de *Thiomonas* 3As. Ces études ont de plus conduit à la mise en évidence d'arsénite-oxydases particulières dont l'activité est exprimée de manière constitutive ou présentant une activité enzymatique élevée. Dans un second temps l'analyse des séquences AoxB issues de diverses bactéries cultivables et de communautés bactériennes a permis la mise en évidence d'une variabilité de séquences AoxB et nous a conduit à proposer le gène *aoxB* comme marqueur moléculaire des bactéries arsénite-oxydantes aérobies. Enfin l'analyse des gènes *aox* par une étude de génomique environnementale sur la communauté bactérienne du site de Carnoulès, a permis de mieux appréhender les interactions trophiques entre les souches et de mettre en évidence le rôle pivot de *Thiomonas* sp. dans cet environnement.

Mots clefs : arsenic, diversité, bactéries arsénite-oxydantes, gènes *aox*

DIVERSITY OF BACTERIAL ARSENIC OXIDO-REDUCTION MECHANISMS: FROM CULTIVABLE STRAINS TO COMMUNITIES

Microorganisms are found in several environments from the friendliest to the most extreme such as arsenical environment as french mining sites of Carnoules, Salsigne or Sainte-Marie-aux-Mines. Some of them are involved in arsenic biochemical cycle and play a crucial role in its mobilization/immobilization and thus in its biodisponibility. These bacteria contain arsenite oxidase (coded by *aox* genes) catalyzing the oxidation of As III in As V. This oxidation constitutes the first arsenical environment detoxification stage.

The aim of this work is the study of the diversity of bacterial arsenic oxido-reduction mechanisms. This study focuses on the diversity of *aoxB* genes from isolated strains or complex communities.

Physiological characterizations of three strains have been first realized on *Pseudomonas* S11, *Leptothrix* S1.1 and *Thiomonas* 3As. A genetic approach supplements the physiological data of *Thiomonas* 3As. Particular arsenite oxidases demonstrating a high or a constitutive arsenite oxidase activity have been highlighted. In a second time AoxB sequences analyses from cultivable bacteria or bacterial communities have underlined a sequence variability and allowed to consider *aoxB* gene as molecular marker for aerobic arsenite oxidizing strains. Lastly an environmental genomic approach on bacterial community from Carnoulès mining site allowed to decipher the bacterial trophic interactions and the key role of the arsenite oxidizing strain *Thiomonas* sp. in this study site.

Key words: arsenic, diversity, arsenite oxidizing bacteria, *aox* genes