

UNIVERSITÉ MONTPELLIER II
SCIENCES ET TECHNIQUES DU LANGUEDOC

THÈSE

pour obtenir le grade de
DOCTEUR DE L'UNIVERSITÉ MONTPELLIER II

Discipline : Microbiologie/Parasitologie

École Doctorale : Systèmes Intégrés en Biologie, Agronomie, Géosciences, HydroSciences et
Environnement (SIBAGHE)

Présentée et soutenue publiquement par

Aurélie VOLANT

le 12 décembre 2012

ÉTUDE DES COMMUNAUTÉS MICROBIENNES (BACTÉRIES, ARCHAEA ET
EUCARYOTES) ET DE LEURS VARIATIONS SPATIOTEMPORELLES DANS LA MINE DE
CARNOULÈS FORTEMENT CONTAMINÉE EN ARSENIC

JURY

Mme Françoise Elbaz-Poulichet	Directeur de Recherche, CNRS, Montpellier	Directeur de thèse
M. Philippe Bertin	Professeur, Université de Strasbourg 1, Strasbourg	Directeur de thèse
M. Patrick Taillez	Ingénieur de recherche, INRA, Montpellier	Examineur
Mme Catherine Joulian	Chercheur, BRGM, Orléans	Examineur
M. Philippe Normand	Directeur de recherche, CNRS, Lyon	Rapporteur
M. Pierre Peyret	Professeur, Université Blaise Pascal, Clermont-Ferrand	Rapporteur
Mme Odile Bruneel	Chargé de recherche, IRD, Montpellier	Invitée

Remerciements

Ce travail de thèse est le fruit de 3 années d'étude et de recherche au sein de l'axe 1 (Biogéochimie, contaminants et santé) du laboratoire HydroSciences Montpellier (UMR 5569). Je remercie Eric Servat, directeur de l'UMR 5569, pour son accueil au sein de l'unité ainsi que mes co-directeurs de thèse, Françoise Elbaz-Poulichet et Philippe Bertin pour avoir accepté cette tâche malgré la distance géographique ou disciplinaire.

J'exprime mes profonds remerciements à mon encadrante Odile Bruneel, pour m'avoir donné l'opportunité de réaliser ma thèse sous ses conseils avisés et ses encouragements et pour la confiance qu'elle m'a accordée tout au long de ce travail. Merci de m'avoir donné la chance de travailler sur un sujet qui m'a passionné et m'a permis de développer mes compétences au-delà du monde microbien, vers la bioinformatique et les statistiques.

Je tiens également à exprimer toute ma gratitude au reste de l'équipe « microbio », qui s'est bien agrandi depuis mon arrivée où j'ai été accueillie par Christian Personné, qui m'a initié à mon nouvel environnement. Je remercie donc Angélique Desoeuvre, la première à être venue agrandir les rangs, pour sa contribution aux diverses manipulations nécessaires à ce travail et son soutien lors des périodes mutuelles de galère avec des clonages qui ne marchent pas. Merci également à Marina Héry avec qui j'ai eu des discussions enrichissantes et qui a apporté des avis pertinents sur divers sujets d'interrogations contribuant à l'avancée de mes travaux. Merci aux stagiaires Camila, Valentina, Lorinda et Aurélia pour leurs contributions et leurs efforts.

Ma reconnaissance s'adresse aussi à toute l'équipe « chimie » pour leur temps et leur énergie consacrés à l'analyse des divers échantillons. Merci en particulier à Corinne Casiot pour ses explications, son soutien et sa disponibilité ainsi qu'à Sophie Delpoux pour son dévouement, sa bonne humeur et son amitié lors des nombreuses missions de terrain à Carnoulès. Sans oublier les autres membres de l'équipe Sandra, Rémy et Jean-luc pour avoir contribué à l'ambiance chaleureuse qui règne au labo, notamment lors des pauses café.

REMERCIEMENTS

Je tiens également à remercier le laboratoire de Pau et en particulier Béatrice Lauga pour m'avoir accueillie lors de mes divers séjours et avoir mis à ma disposition son expérience, ses connaissances et les moyens techniques nécessaires à mes travaux.

Merci à tous ceux qui m'ont apporté un soutien amical c'est à dire les nombreuses personnes travaillant à HydroSciences, chercheurs, techniciens et étudiants, et en particulier mes collègues de bureau : Cristina dont je me souviens encore des soirées arrosées à la Cachaça un an plus tôt et Marianne avec qui j'ai passé mes derniers mois de rédaction intensive en nous soutenant mutuellement. Je remercie également Stan, pour m'avoir initiée au bodyboard et pour avoir partagé avec moi de nombreuses sorties au Lac du Crès me permettant de me détendre après de longues heures de boulot.

Je suis extrêmement reconnaissante envers l'ensemble des membres du jury d'avoir accepté de juger mon travail. Je remercie M. Pierre Peyret et M. Philippe Normand pour leur travail de rapporteur ainsi que Mme Catherine Joulian et M. Patrick Taillez pour leur présence en tant qu'examineurs.

Enfin je remercie mes parents, qui m'ont toujours aidée et soutenue et sans qui je n'aurais pu en arriver là aujourd'hui.

Table des matières

INTRODUCTION

INTRODUCTION GÉNÉRALE.....	1
SYNTHÈSE BIBLIOGRAPHIQUE.....	5
A. LES FONDEMENTS DE LA DIVERSITÉ ET DE L'ÉCOLOGIE MICROBIENNE	5
1. DÉFINITION ET HISTORIQUE	5
2. LES MICROORGANISMES, SOURCE DE DIVERSITÉ GÉNÉTIQUE	6
3. LES MICROORGANISMES, ACTEURS DES PROCESSUS ENVIRONNEMENTAUX.....	7
3.1. <i>Le cycle du carbone</i>	7
3.2. <i>Le cycle de l'azote</i>	8
3.3. <i>Le cycle du soufre</i>	10
4. LES MICROORGANISMES ET LEURS RÔLES DANS LA TRANSFORMATION DES ÉLÉMENTS TRACES MÉTALLIQUES ..	12
B. UN CAS PARTICULIER : LES DRAINAGES MINIERS ACIDES	13
1. DÉFINITION ET ORIGINE	13
2. MÉCANISMES DE FORMATIONS	14
2.1. <i>Mécanismes physiques : importance des activités humaines</i>	14
2.2. <i>Mécanismes chimiques : importance de l'environnement naturel</i>	15
2.2.1. Les réactions primaires.....	15
2.2.2. Les réactions secondaires	17
2.3. <i>Mécanismes biologiques : importance des microorganismes</i>	18
2.4. <i>Facteurs influençant la formation du DMA</i>	19
2.4.1. Les facteurs primaires.....	19
2.4.2. Les facteurs secondaires	20
3. MÉCANISMES D'ATTÉNUATION	20
3.1. <i>Précipitation</i>	20
3.1.1. La production de sulfures métalliques.....	20
3.1.2. La co-précipitation.....	21
3.2. <i>Séquestration</i>	22
3.2.1. La séquestration extracellulaire	22
3.2.2. La séquestration intracellulaire.....	22
3.3. <i>Biométhylation</i>	22
3.4. <i>Sidérophores</i>	23
3.5. <i>Oxydation de l'arsenic</i>	23
4. MICROORGANISMES PRÉSENTS DANS LES DRAINAGES MINIERS ACIDES	23
4.1. <i>Diversités phylogénétiques et physiologiques</i>	23
4.2. <i>Les procaryotes</i>	25

TABLE DES MATIÈRES

4.2.1.	Les bactéries	25
4.2.2.	Les Archaea	27
4.3.	<i>Les eucaryotes</i>	29
5.	COMMUNAUTÉS MIXTES ET INTERACTIONS MICROBIENNES	32
5.1.	<i>Compétition</i>	32
5.2.	<i>Prédation</i>	33
5.3.	<i>Mutualisme</i>	33
5.4.	<i>Synergisme</i>	33
6.	CYCLES BIOGÉOCHIMIQUES CATALYSÉS PAR LES MICROORGANISMES	34
6.1.	<i>Les microorganismes utilisant le fer</i>	34
6.1.1.	Oxydation du fer	34
6.1.2.	Réduction du fer	35
6.2.	<i>Les microorganismes utilisant le soufre</i>	35
6.2.1.	Oxydation du soufre	35
6.2.2.	Réduction du soufre	36
6.3.	<i>Les microorganismes utilisant l'arsenic</i>	36
6.3.1.	Oxydation de l'arsenic	36
6.3.2.	Réduction de l'arsenic	37
C.	INTÉRÊTS DES APPROCHES –OMICS	40
1.	LA MÉTAGÉNOMIQUE	42
1.1.	<i>La technologie 454</i>	43
1.2.	<i>La technologie Illumina</i>	45
1.3.	<i>Les technologies SOLiD et de troisième génération</i>	46
2.	LA MÉTATRANSCRIPTOMIQUE	49
2.1.	<i>Les premières méthodologies</i>	50
2.2.	<i>Le séquençage de nouvelle génération</i>	50
3.	LA MÉTAPROTÉOMIQUE	51
3.1.	<i>Techniques de séparation des protéines/peptides</i>	53
3.1.1.	Techniques basées sur l'électrophorèse	53
3.1.2.	Techniques basées sur la chromatographie liquide	53
3.2.	<i>Ionisation et spectrométrie de masse</i>	55
4.	LA MÉTA-MÉTABOLOMIQUE	56
4.1.	<i>La spectroscopie par résonance magnétique nucléaire</i>	57
4.2.	<i>La spectrométrie de masse</i>	58

RÉSULTATS

CHAPITRE I – CARACTÉRISATION DE LA COMMUNAUTÉ BACTÉRIENNE ACTIVE .	63
A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE	63
1. CONTEXTE GÉNÉRAL.....	63
2. OBJECTIFS DE L'ÉTUDE.....	66
B. RÉSULTATS	67
C. DISCUSSION ET CONCLUSION	68
D. CONTRIBUTION SCIENTIFIQUE	69
CHAPITRE II – DYNAMIQUE TEMPORELLE DE LA COMMUNAUTÉ D'ARCHAEA	93
A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE	93
1. CONTEXTE GÉNÉRAL.....	93
2. OBJECTIF DE L'ÉTUDE.....	94
B. RÉSULTATS	94
C. DISCUSSION ET CONCLUSION	95
D. CONTRIBUTION SCIENTIFIQUE	95
CHAPITRE III – VARIABILITÉ SPATIALE DE LA COMMUNAUTÉ EUCARYOTE.....	115
A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE	115
1. CONTEXTE GÉNÉRAL.....	115
2. OBJECTIF DE L'ÉTUDE.....	116
B. RÉSULTATS	117
C. DISCUSSION ET CONCLUSION	118
D. CONTRIBUTION SCIENTIFIQUE	119
CHAPITRE IV – DYNAMIQUES SPATIOTEMPORELLES DE LA COMMUNAUTÉ BACTÉRIENNE ET DRIVERS PHYSICOCHIMIQUES	155
A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE	155
1. CONTEXTE GÉNÉRAL.....	155
2. OBJECTIFS DE L'ÉTUDE.....	156
B. RÉSULTATS	156
C. DISCUSSION ET CONCLUSION	158
D. CONTRIBUTION SCIENTIFIQUE	159

CONCLUSIONS GÉNÉRALES ET PERSPECTIVES
--

CONCLUSIONS GÉNÉRALES	ERREUR ! SIGNET NON DÉFINI.
PERSPECTIVES.....	203
RÉFÉRENCES BIBLIOGRAPHIQUES	205
LISTE DES FIGURES.....	225
LISTE DES TABLEAUX	227
LISTE DES ABRÉVIATIONS.....	228
LISTE DES SYMBOLES	229

INTRODUCTION

Introduction générale

PROBLÉMATIQUE

Les microorganismes présentent une grande diversité génétique et fonctionnelle leur permettant de coloniser tous les écosystèmes de notre planète y compris les plus extrêmes comme par exemple les lacs salés (Rodriguez-Valera *et al.*, 1981), les sources hydrothermales profondes (Juniper & Tebo, 1995), les sédiments marins profonds (Parkes *et al.*, 1994), les déserts (Adams *et al.*, 1992), les régions polaires (Vincent & James, 1996) ou les déchets miniers acides riches en métaux et métalloïdes (ou éléments traces métalliques, ETM) (Ledin & Pedersen, 1996; Baker *et al.*, 2004). En plus de leur ubiquité, les communautés microbiennes sont particulièrement abondantes, on estime par exemple que la biosphère contiendrait jusqu'à 6.10^{30} cellules procaryotes (Whitman *et al.*, 1998), représentant ainsi la plus importante biomasse de notre planète. L'ensemble de ces caractéristiques en font des acteurs clés du fonctionnement des écosystèmes notamment par leur implication majeure dans les cycles biogéochimiques indispensables à la vie sur Terre comme ceux du carbone, de l'azote ou du soufre.

Les microorganismes sont également impliqués dans de nombreux processus biogéochimiques au centre desquels se trouvent les transformations des minéraux et des métaux (Gadd, 2010) pouvant affecter la mobilisation des éléments toxiques susceptibles de menacer la santé publique. Il existe des environnements naturellement riches en ETM tels que les sources chaudes ou les lacs acides des régions volcaniques (*e.g.* parc de Yellowstone, U.S.A.), les sources hydrothermales profondes ou les zones d'affleurement naturel de gisements métallifères dont l'altération peut générer des eaux acides chargées en métaux. Cependant ce sont principalement les pollutions métalliques d'origine anthropique qui posent problème aujourd'hui. En exploitant les ressources minérales depuis des millénaires, l'Homme a créé de nouveaux environnements riches en ETM. Lors de l'exploitation minière, les déchets provenant de l'extraction et du traitement des minerais sulfurés (*e.g.* la pyrite ou l'arsénopyrite) sont mis en contact avec de l'eau et de l'oxygène, induisant l'oxydation des sulfures à l'origine du phénomène de Drainage Minier Acide (DMA).

INTRODUCTION GÉNÉRALE

Outre l'acidification de ces eaux de drainage, ce phénomène soulève de nombreuses préoccupations et constitue aujourd'hui le problème majeur de l'industrie extractive mondiale puisqu'il s'accompagne de la lixiviation de métaux et métalloïdes dont la toxicité peut être dommageable pour l'environnement. Malgré leur toxicité, ces écosystèmes sont colonisés par divers procaryotes qui forment des communautés microbiennes stables avec des eucaryotes (champignons, levures, algues et protozoaires) et jouent un rôle crucial dans ces systèmes (Johnson, 1998; Amaral-Zettler *et al.*, 2002). En effet, certains de ces organismes sont capables d'utiliser les réactions d'oxydoréduction telles que l'oxydation du fer et du soufre pour leurs besoins métaboliques (Baker & Banfield, 2003). Ces activités métaboliques se traduisent par la solubilisation (biolessivage) des éléments chimiques à partir des minerais sulfurés et conduisent ainsi à la pollution des eaux superficielles et souterraines alimentées par les eaux de ruissellement. D'autre part, ces mêmes microorganismes interviennent également dans des processus de bioremédiation ou piégeage d'éléments toxiques en solution en rendant les composés toxiques insolubles permettant ainsi leur accumulation dans les sédiments (Johnson & Hallberg, 2005; Hallberg, 2010). Une remédiation naturelle a par exemple été observée à Carnoulès en France ou sur le site du Rio Tinto en Espagne (Casiot *et al.*, 2003a; Sánchez España *et al.*, 2005). Dans ces DMA, acides et riches en fer, les composés toxiques sont accumulés dans les sédiments sous forme d'une variété d'(oxyhydr)-oxydes de fer et d'hydroxysulphates tel que la jarosite, la schwermannite et la ferrhydrite (Ona-Nguema *et al.*, 2005). Ces processus de précipitation impliquent ici l'oxydation et la précipitation du fer, qui est souvent le métal soluble le plus abondant dans les DMA et l'adsorption d'autres métaux et métalloïdes sur les minéraux ferriques formés (Rowe & Johnson, 2008; Hallberg, 2010). Parmi les éléments toxiques des DMA, l'arsenic (As) pose un problème particulier. C'est en effet un composé très toxique qui peut affecter le fonctionnement des cellules par différents mécanismes. D'une manière générale, sa toxicité sous sa forme inorganique est liée à l'analogie structurale de l'As(V) avec le phosphate (considéré comme un élément essentiel à la vie) et à la très forte affinité de l'As(III) pour les fonctions sulfhydryles. Ces deux caractéristiques sont à l'origine de la majorité des dysfonctionnements cellulaires. Il a été montré que la forme réduite As(III), serait de 25 à 60 fois plus toxique que la forme oxydée As(V) pour certains organismes supérieurs (Korte & Fernando, 1991). De plus, l'As(III) se fixe moins facilement sur les oxydes de fer que l'As(V) à faible pH, il est donc plus mobile dans ces eaux. Chez l'être humain, l'arsenic est très toxique car mutagène et carcinogène (Yamanaka & Okada, 1994; Gebel, 1997). Les DMA représentent ainsi une des formes les

INTRODUCTION GÉNÉRALE

plus graves de pollution de l'eau dans les zones industrielles et post-industrielles au niveau mondial (Johnson, 1995). De plus, la production de DMA peut durer des centaines d'années après la fermeture d'un site minier entraînant la pollution de dizaines de milliers de km de rivières et de nappes phréatiques ainsi que des hectares de lacs dans le monde (Johnson & Hallberg, 2005; Natarajan, 2008). De nombreux pays se voient donc maintenant confrontés aux impacts engendrés par des exploitations en activités ou arrêtées depuis longtemps et pour lesquelles aucune prescription réglementaire ou législative n'imposait à cette époque de précautions environnementales dans la gestion des sites. Ainsi, les DMA représentent une source importante de préoccupations depuis plusieurs années, du fait de la menace qu'ils représentent tant pour l'environnement que pour la santé humaine.

OBJECTIFS

Dans ce contexte, les travaux réalisés au cours de cette thèse qui s'inscrit dans une thématique d'écologie microbienne, avaient pour objectif principal de contribuer à une meilleure connaissance des microorganismes présents dans les DMA et impliqués dans les processus de transferts des métaux et métalloïdes (en particulier l'As). La compréhension de ces processus étant d'une grande importance pour le développement à terme de stratégies impliquant ces microorganismes dans des procédés de remédiation. Pour réaliser cet objectif, nous avons axé notre étude sur le DMA de Carnoulès localisé en France, qui constitue un exemple remarquable d'adaptation des organismes vivants à un environnement extrême présentant un pH acide (2-4) et des concentrations très élevées en sulfate et en éléments métalliques comme l'arsenic (jusqu'à 10 g L^{-1} , concentration parmi les plus importantes au monde). De plus, ce site est très approprié pour analyser la façon dont les microorganismes contribuent à la transformation des métaux et métalloïdes in situ, car des processus efficaces de remédiation naturelle ont été mis en évidence sur ce site (Casiot *et al.*, 2003a; Bruneel *et al.*, 2006). Ces caractéristiques particulières en ont fait un site modèle qui est étudié depuis une quinzaine d'années à la fois d'un point de vue géochimique et microbiologique; il est labellisé par l'OSU OREME depuis 2009.

Les objectifs spécifiques de ce travail visaient à :

- étudier la diversité des microorganismes présents (bactéries, Archaea et eucaryotes) au niveau des eaux drainant la mine ainsi que dans les sédiments du ruisseau par des

INTRODUCTION GÉNÉRALE

approches de génomique impliquant des méthodes classiques de clonage/séquençage mais également des méthodes de pyroséquençage à haut débit ;

- caractériser la dynamique spatiale et/ou temporelle des microorganismes pour mieux comprendre leur évolution dans le système et leur lien avec les métaux et métalloïdes ;
- Identifier les paramètres physicochimiques (pH, température, concentrations en fer, sulfate, arsenic...) à l'origine des profils de diversité observés.

ORGANISATION DU MÉMOIRE

Ce manuscrit est organisé en trois parties :

- une **synthèse bibliographique** permettant d'introduire les travaux réalisés au cours de ma thèse présentera tout d'abord la notion d'écologie microbienne et soulignera l'importance des microorganismes pour le fonctionnement des écosystèmes. Une seconde partie fera ensuite un état des connaissances actuelles sur le phénomène de drainage minier acide, ses mécanismes de formation et d'atténuation ainsi que sur les microorganismes présents et leurs rôles dans les cycles biogéochimiques. Enfin, les avancées technologiques récentes ou technologies « -omics » utilisées pour l'analyse des communautés microbiennes feront l'objet de la troisième partie de cette synthèse.
- les **résultats** seront présentés dans les quatre chapitres suivants essentiellement sous forme d'articles scientifiques. Le premier chapitre portera sur l'identification des bactéries présentes et sur les communautés actives dans les eaux et les sédiments du Reigous. Un second chapitre s'intéressera à la diversité phylogénétique des communautés d'Archaea et leur variation temporelle au sein des sédiments. Le troisième chapitre présentera la diversité des communautés eucaryotes également présentes dans les sédiments ainsi que leur répartition spatiale en relation avec les paramètres physicochimiques. Enfin, un quatrième chapitre abordera la dynamique spatiotemporelle des communautés bactériennes présentes dans les eaux du Reigous et les déterminants physicochimiques qui pourraient en être à l'origine.
- une **conclusion générale** permettra d'expliquer l'apport de ce travail à la connaissance fondamentale concernant le biofonctionnement du DMA de Carnoulès et d'ouvrir sur des perspectives nouvelles permettant de compléter les résultats obtenus.

Synthèse bibliographique

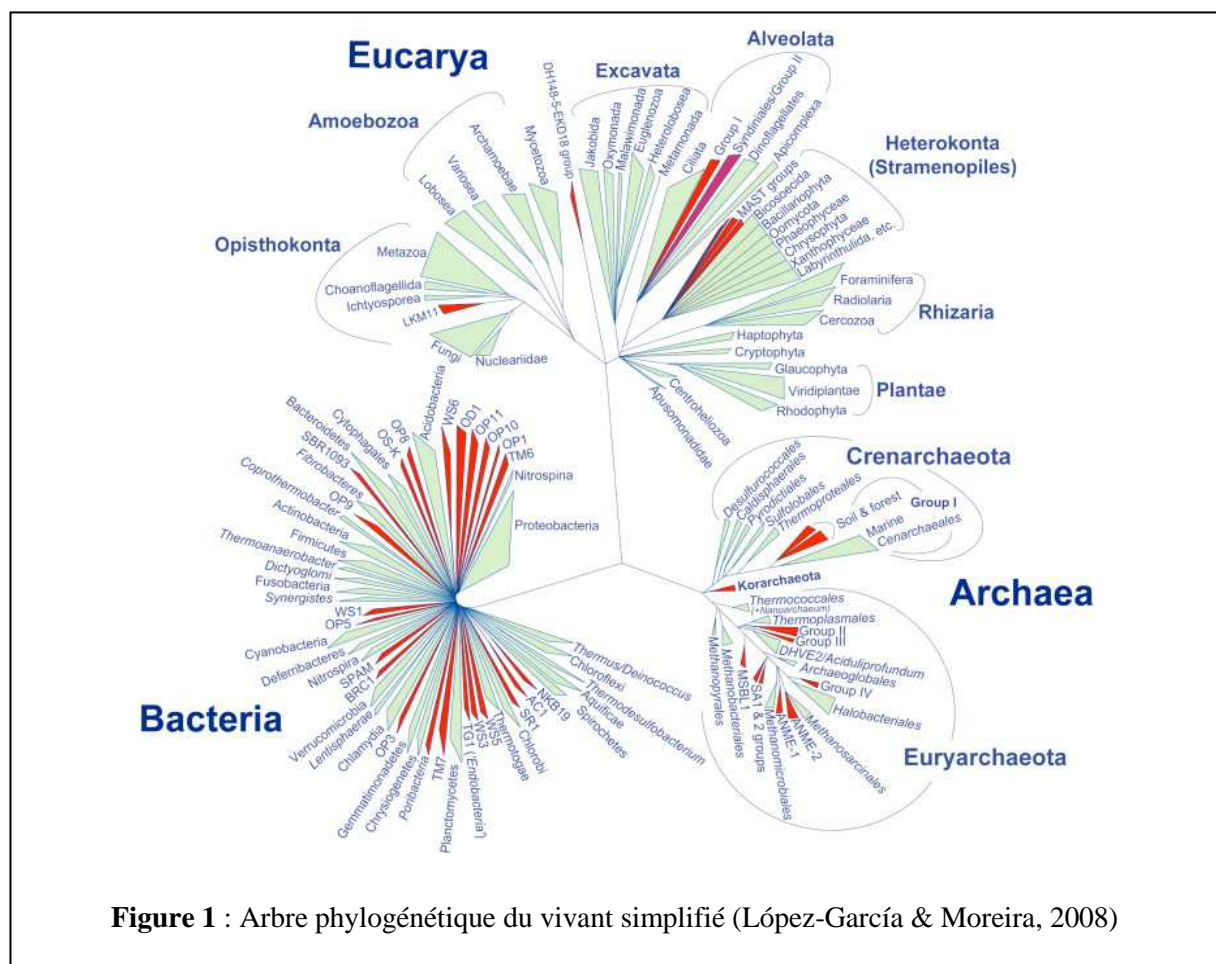
A. LES FONDEMENTS DE LA DIVERSITÉ ET DE L'ÉCOLOGIE MICROBIENNE

1. Définition et historique

A l'origine, la microbiologie a été définie comme l'étude des êtres vivants unicellulaires ou pluricellulaires invisibles à l'œil nu, c'est-à-dire les microorganismes. Cependant certains de ces organismes sont visibles à l'œil nu, en particulier certains microorganismes eucaryotes. Au début du 18^{ème} siècle, Antonie van Leeuwenhoek fut le premier à observer des « animalcules » grâce à des microscopes de sa fabrication (Porter, 1976). Il fallut cependant attendre 200 ans pour que la microbiologie connaisse un réel essor, amorcé avec les travaux fondateurs de Louis Pasteur et Robert Koch. Durant la seconde moitié du 19^{ème} siècle, Louis Pasteur développa la culture pure de souches microbiennes, mettant fin au débat sur la génération spontanée et démontra l'implication de ces organismes dans les processus de fermentation et dans les maladies. Parallèlement, Robert Koch proposa une série de postulats servant à confirmer le rôle étiologique d'un microorganisme dans une maladie infectieuse donnée (O'Brien & Goedert, 1996). Durant cet âge d'or de la microbiologie, Sergei Winogradsky et Martinus Beijerinck ont mis en évidence l'existence de métabolismes variés comme les processus de symbiose plantes/bactéries fixatrices d'azote et l'importance des bactéries dans le cycle du soufre et de l'azote (Kupferberg, 2003; Dworkin, 2012). Ils ont ainsi initié les premiers travaux de microbiologie de l'environnement et d'écologie microbienne en étudiant les communautés de microorganismes du sol et de l'eau et les relations entre ces organismes. Depuis, les connaissances en microbiologie n'ont cessé de s'accroître par le biais d'avancées technologiques, qui ont connu un essor depuis le milieu du 20^{ème} siècle (Prescott *et al.*, 2010). La microbiologie moderne englobe deux sous-disciplines qui se distinguent par leur échelle d'étude. L'écologie microbienne s'intéresse à l'étude du comportement et des activités des microorganismes dans leur environnement naturel immédiat alors que la microbiologie environnementale vise à étudier les processus microbiens globaux c'est-à-dire à plus grande échelle (Lilley *et al.*, 2002).

2. Les microorganismes, source de diversité génétique

Les microorganismes constituent un groupe extrêmement diversifié d'organismes microscopiques, majoritairement unicellulaires et répartis dans les trois domaines du vivant Bacteria, Archaea et Eukaryota (Figure 1). Les eucaryotes qui possèdent un vrai noyau cellulaire délimité par une membrane nucléaire se distinguent des procaryotes (bactéries et Archaea) dont la cellule ne possède pas de noyau cellulaire ni d'autres organites. Les virus, incapables de se reproduire sans détourner la machinerie cellulaire d'un autre organisme, ne sont pas considérés par tous les spécialistes comme vivants. La microbiologie étudie parfois leurs actions sur les microorganismes mais n'a pas pour but de les étudier en tant qu'entités. Cette étude est réalisée dans une autre discipline de la biologie : la virologie.



La diversité microbienne actuelle est la résultante de l'évolution sur presque quatre milliards d'années. En effet, ces organismes présentent un temps de génération relativement court qui a favorisé l'accumulation d'événements de mutation, de délétion ou de recombinaison dans

leurs génomes. A cela s'est ajouté des phénomènes de transferts horizontaux (*i.e.* échanges de gènes entre taxons différents) entre procaryotes (Ochman *et al.*, 2000), de procaryotes à eucaryotes (Hall *et al.*, 2005) mais aussi entre eucaryotes (Xie *et al.*, 2008). La nature très dynamique de leur génome confère ainsi aux microorganismes un fort potentiel adaptatif leur permettant de coloniser divers types d'écosystèmes même les plus extrêmes. La diversité des microorganismes résulte également de phénomènes de radiation évolutive (Cohan & Koeppel, 2008) qui, par la sélection de traits adaptatifs à des niches écologiques différentes, peuvent engendrer un isolement reproducteur entre populations à l'origine de nouvelles espèces. Les microorganismes sont particulièrement abondants et présentent une diversité extrêmement vaste. On estime par exemple, qu'il y a près de 10^6 taxons par gramme de sol (Siggins *et al.*, 2012) et que les cellules procaryotes représentent à elles seules près de la moitié de la biomasse de la planète (Whitman *et al.*, 1998). Ainsi, bien qu'ils soient généralement invisibles à l'œil nu, les microorganismes constituent une composante essentielle du vivant en termes d'abondance et de diversité.

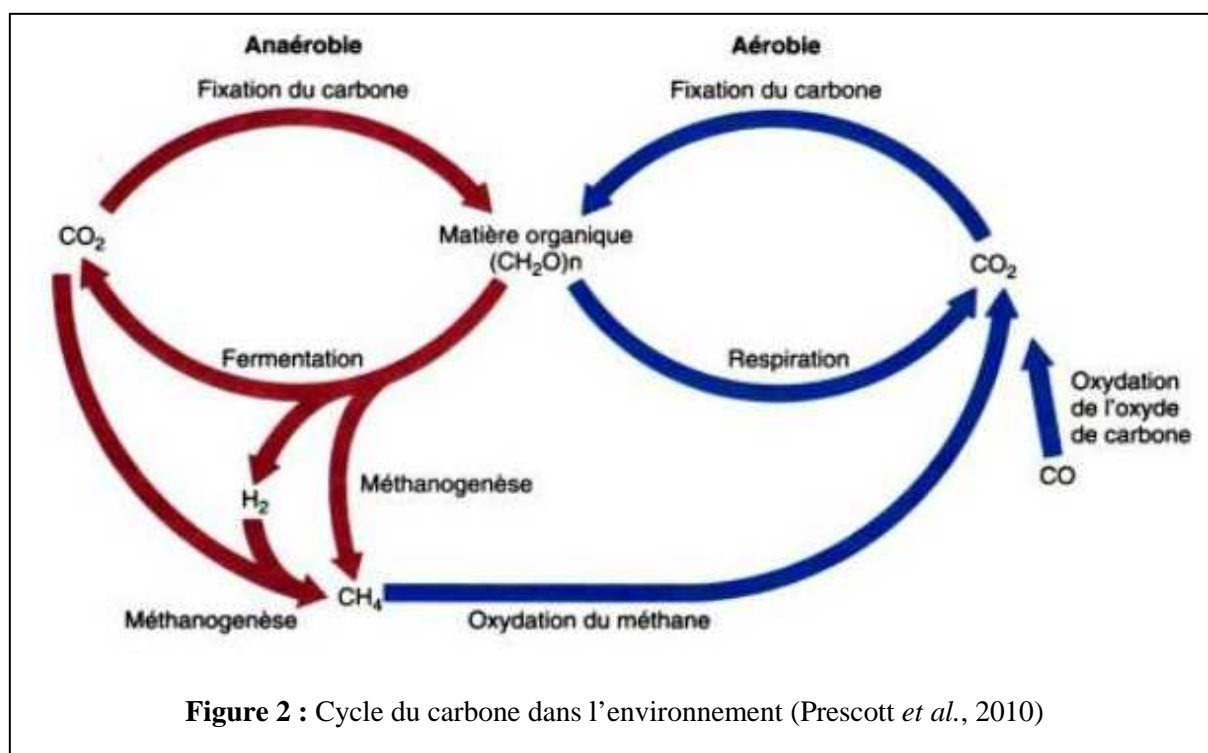
3. Les microorganismes, acteurs des processus environnementaux

Les microorganismes sont à l'origine des conditions environnementales actuelles car ce sont les principaux moteurs de la biogéochimie, des cycles d'éléments nutritifs et de la dégradation des déchets naturels et anthropiques de la planète (Rodriguez-Valera, 2004). Les microorganismes sont en effet les principaux acteurs du fonctionnement des cycles biogéochimiques de divers éléments majeurs (carbone, azote, soufre...) et en traces (fer, mercure, sélénium, arsenic...). Tous ces cycles sont reliés entre eux et les transformations métaboliques de ces nutriments par les microorganismes ont des impacts planétaires.

3.1. Le cycle du carbone

Le carbone constitue moins de 1 % de la masse de l'écorce terrestre mais c'est un des éléments majeurs du monde du vivant. En effet, le passage du CO₂ minéral au monde organique *via* la photosynthèse est une réaction clé de la vie sur terre. Au moins la moitié du carbone sur Terre est fixé par les microorganismes, particulièrement par les procaryotes et les protistes photosynthétiques marins comme les cyanobactéries (*e.g.* *Prochlorococcus* et *Synechococcus*) et les diatomées, respectivement (Prescott *et al.*, 2010). Le carbone peut également être fixé par des chimiolithoautotrophes. Le cycle du carbone est caractérisé par des phases aérobies et

anaérobies de fixation du carbone et de minéralisation de la matière organique *via* les mécanismes de respiration et de fermentation (Figure 2). Les microorganismes assurent des fonctions essentielles au sein de ce cycle en tant que décomposeurs mais également en tant que producteurs de matières organiques. Les microorganismes du sol par exemple, sont les principaux responsables de la minéralisation de la matière organique. Ils contrôlent donc le recyclage des nutriments nécessaires à la croissance des plantes et l'équilibre entre stockage et déstockage du carbone qui a des conséquences sur les émissions de gaz à effet de serre et sur l'érosion des sols (Conrad, 1996; Gyaneshwar *et al.*, 2002).

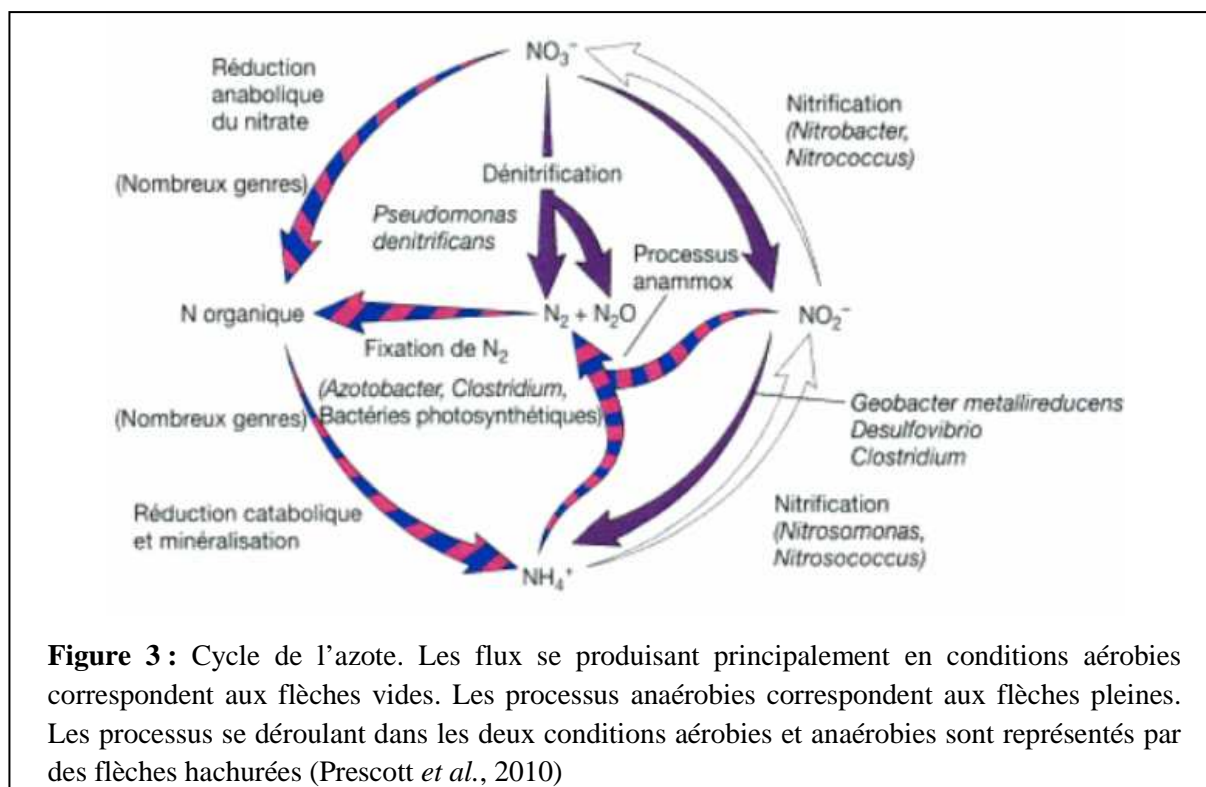


3.2. Le cycle de l'azote

L'azote est un élément essentiel à tous les organismes, il est présent dans la constitution des protéines (acides aminés), des acides nucléiques et divers polymères (Silberbach *et al.*, 2005). Les microorganismes ont un rôle primordial dans ce cycle car la fixation de l'azote est un processus exclusivement procaryote (Raymond *et al.*, 2004). Les nombreuses transformations de l'azote, permettent à cet élément de circuler entre le compartiment atmosphérique et les milieux terrestres et aquatiques et de déterminer en partie la productivité écologique de ces habitats.

Ce cycle se caractérise par différentes étapes (Figure 3) :

- la **fixation de l'azote atmosphérique** produit de l'ammonium qui conduira à la synthèse de composés azotés organiques (acides aminés). Ce processus fait intervenir des communautés bactériennes symbiotiques, spécifiques des végétaux auxquels ces bactéries sont associées (rhizobium, actinomycètes), des communautés bactériennes hétérotrophes aérobies ou anaérobies non symbiotiques ou dites « libres » et des communautés photolithotrophes (cyanobactéries) (Normand & Lalonde, 1986; Franche *et al.*, 2009).
- l'**ammonification ou minéralisation de la matière organique** qui conduit à la formation d'ammonium est réalisée par une très grande partie des communautés bactériennes hétérotrophes aérobies et anaérobies et des communautés fongiques.
- la **nitrification** est un processus en deux étapes au cours desquelles, l'ammoniaque (NH_3) ou les ions ammonium (NH_4^+) sont oxydés en nitrite (NO_2^-) puis en nitrate (NO_3^-). La première étape peut être réalisée par deux groupes de microorganismes, les AOB (Ammonia-Oxidizing Bacteria) caractérisant les bactéries nitrosantes telle que *Nitrosomonas* et les AOA (Ammonia-Oxidizing Archaea) telle que *Nitrososphaera* (Treusch *et al.*, 2005). La seconde étape est réalisée par les bactéries nitrifiantes ou NOB (Nitrite-Oxidizing Bacteria) telle que *Nitrobacter* et *Nitrospira* (Cebren & Garnier, 2005). Ces étapes se déroulent préférentiellement en aérobiose.
- la **réduction assimilatrice ou dissimilatrice des nitrates**. L'assimilation (ou réduction assimilatrice) des nitrates est réalisée, en présence d'oxygène, par un groupe hétérogène de bactéries, champignons, algues et plantes. En conditions anaérobies, lors du processus de réduction dissimilatrice du nitrate ou respiration des nitrates, le nitrate peut devenir accepteur final d'électrons et être converti en ammonium. Des bactéries fermentatives et anaérobies strictes très diverses sont capables de réaliser cette voie dissimilatrice.
- la **dénitrification** est un mécanisme de respiration anaérobie utilisant le NO_3^- ou le NO_2^- comme accepteur final d'électrons. Ce processus fait intervenir des populations bactériennes, surtout hétérotrophes mais aussi autotrophes qui, en conditions anoxiques, assurent la production d'azote sous forme de composés gazeux tels que le N_2O , le N_2 et le NH_3 .



3.3. Le cycle du soufre

Le soufre est un élément très important pour les êtres vivants principalement en tant que constituant des acides aminés soufrés : cystine, cystéine et méthionine. Depuis ces dernières décennies, le cycle du soufre a pris une importance capitale en raison des nombreux problèmes économiques qui lui sont liés comme les pollutions engendrées par les pluies acides contenant de l'acide sulfurique, la pollution atmosphérique au dioxyde de soufre ou la solubilisation et lixiviation des métaux.

Ce cycle comprend trois grandes étapes (Figure 4) :

- la **minéralisation aérobie et anaérobie du soufre organique** en une forme inorganique, le sulfure d'hydrogène (H_2S).
- l'**oxydation des sulfures, du soufre élémentaire et des composés connexes (thiosulfate, sulfite...)** en sulfates. En conditions aérobies, acides ou neutres, des bactéries chimiolithotrophes oxydent ces composés comme source d'énergie, en utilisant l' O_2 comme accepteur d'électrons pour assurer la fixation autotrophe du CO_2 . Ces bactéries appartiennent à des genres divers tels que *Acidithiobacillus*, *Aquaspirillum*, *Aquifex*, *Bacillus*, *Beggiatoa*, *Methylobacterium*, *Paracoccus*, *Pseudomonas*, *Starkeya*, *Sulfolobus*, *Thermithiobacillus*, *Thiobacillus* et *Xanthobacter* (Friedrich *et al.*, 2001). Des membres

des Archaea appartenant aux *Sulfolobales* sont également capables d'oxyder le soufre (Friedrich *et al.*, 2005; Plumb *et al.*, 2007). Quelques espèces phototrophes peuvent aussi croître en anaérobiose en utilisant les nitrates comme accepteur d'électrons et appartiennent aux genres *Allochrochromatium*, *Chlorobium*, *Rhodobacter*, *Rhodospseudomonas*, *Rhodovulum*, et *Thiocapsa* (Friedrich *et al.*, 2001).

- la **réduction des sulfates** en sulfures. En conditions anaérobies strictes, les bactéries sulfato-réductrices utilisent les sulfates comme accepteur d'électrons (respiration anaérobie) pour les réduire en sulfure (H_2S). Ce processus est également désigné par le terme de réduction catabolique des sulfates. Divers microorganismes peuvent réaliser ce processus incluant les bactéries du genre *Desulfovibrio*, *Clostridium*, *Desulfotomaculum* et *Alteromonas* (Muyzer & Stams, 2008) ainsi que des Archaea des genre *Archaeoglobus*, *Thermocladium* et *Caldivirga* (Thevenieau *et al.*, 2007). On distingue par ailleurs une sulfato-réduction dite anabolique qui, par réduction des sulfates en H_2S , conduit à la production de composés organiques soufrés (cystéine, méthionine,...) utilisés pour leur nutrition.

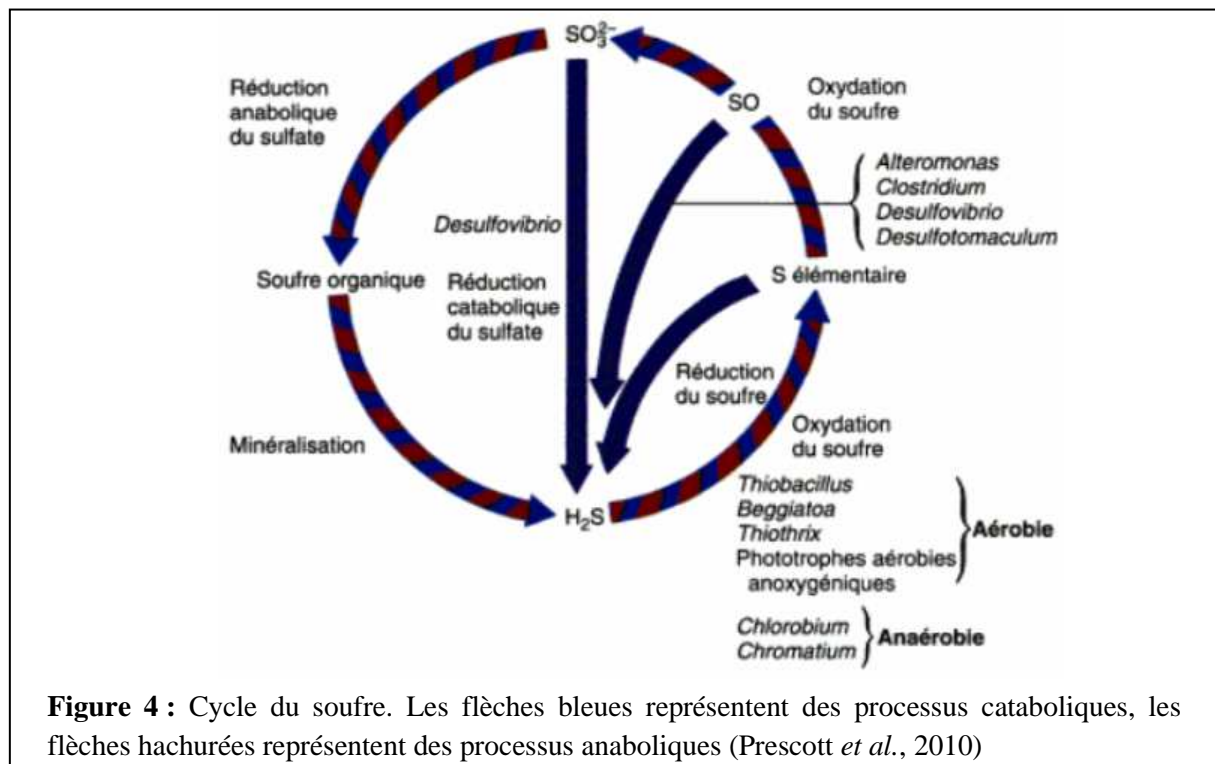


Figure 4 : Cycle du soufre. Les flèches bleues représentent des processus cataboliques, les flèches hachurées représentent des processus anaboliques (Prescott *et al.*, 2010)

Les différents processus microbiens impliqués dans le cycle du soufre assurent des couplages biologiques entre les cycles du carbone, du soufre et du fer et même de l'azote et des couplages chimiques avec les réactions d'immobilisation et de solubilisation de nombreux

éléments traces métalliques (arsenic, plomb, nickel, uranium, sélénium...) susceptibles d'avoir des impacts environnementaux et socio-économiques majeurs.

4. Les microorganismes et leurs rôles dans la transformation des éléments traces métalliques

Les microorganismes sont intimement impliqués dans la biogéochimie des éléments traces métalliques (ETM) par l'intermédiaire d'une variété de processus qui déterminent la toxicité ainsi que la mobilité et par conséquent, la biodisponibilité de ces éléments. L'équilibre entre la mobilisation et l'immobilisation varie selon les organismes impliqués, leur environnement et les conditions physicochimiques (Gadd, 2010). La mobilisation des ETM peut résulter de divers mécanismes impliquant par exemple des réactions d'oxydoréduction, de complexation par des sidérophores et des métabolites ou de méthylation pouvant conduire à la volatilisation. Inversement, des phénomènes de sorption, de séquestration intracellulaire ou extracellulaire ou de chélation peuvent induire l'immobilisation de ces mêmes éléments et ainsi limiter leur disponibilité et donc leur toxicité. Ces processus microbiens permettant d'agir sur la transformation des ETM sont actuellement utilisés dans le cadre des biotechnologies soit dans des procédés de lixiviation des métaux précieux, soit dans des procédés de remédiation d'écosystèmes pollués. Nous verrons par ailleurs plus en détails dans la partie suivante comment les microorganismes peuvent influencer sur ces phénomènes de solubilisation (paragraphe B.2.3) ou d'immobilisation (paragraphe B.3) des ETM à travers l'exemple des drainages miniers acides, qui constituent une des plus importantes sources de pollution des eaux des industries extractives.

Les microorganismes occupent un rôle central dans le vivant de par leur abondance, leur diversité et leur implication dans les processus environnementaux. Leur implication plus particulière dans le cycle des ETM peut avoir des conséquences à la fois bénéfiques et négatives sur les écosystèmes. Dans le contexte des pollutions métalliques il est indispensable d'améliorer nos connaissances sur le compartiment microbien afin de mieux appréhender le fonctionnement de ces écosystèmes et ainsi pouvoir développer des stratégies efficaces de remédiation de ces sites pollués.

B. UN CAS PARTICULIER : LES DRAINAGES MINIERS ACIDES

1. Définition et origine

Les **Drainages Miniers Acides** (DMA) sont générés lorsque les roches contenant des minéraux sulfurés sont exposées à l'érosion (Baker & Banfield, 2003). Les principaux minéraux sulfurés rencontrés dans les gisements et rejets miniers sont énumérés au tableau 1. Il est important de noter que ces eaux acides peuvent apparaître naturellement en l'absence d'activités minières et sont alors connues sous le nom de **Drainage Acides de Roches** (DAR). Cependant, ce phénomène est largement amplifié par les activités humaines et prend une ampleur considérable au sein des activités d'extractions minières. En effet, ces activités (excavation et pompage) perturbent l'équilibre chimique des affleurements de minéraux sulfurés et des gisements de sulfures métalliques profonds en les plaçant brutalement en conditions oxydantes. Ceci conduit à une oxydation rapide des minéraux présents dans les roches par des processus chimiques et biologiques, générant ainsi des effluents souvent acides et riches en métaux et métalloïdes toxiques appelés drainages miniers acides. L'expression « drainage minier acide » est employée pour désigner à la fois le mécanisme d'acidification et les eaux acides qui en découlent. Par extension, cette expression englobe généralement toutes les eaux acides d'origine minière provenant de l'oxydation spontanée des minéraux sulfurés : les eaux produites au fond des mines à ciel ouvert ou dans les galeries ainsi que les eaux de percolation sur les matériaux miniers stockés à l'extérieur de la mine.

Tableau 1 : Principaux minéraux sulfurés

Métal/métalloïde associé	Exemple	Formules
Antimoine	Stibine	Sb₂S₃
Argent	Argentite	Ag₂S
	Proustite	Ag₃AsS₃
Arsenic	Orpiment	As₂S₃
	Réalgar	AsS
Bismuth	Bismuthinite	Bi₂S₃
Cuivre	Chalcopyrite	CuFeS₂
	Énargite	Cu₃AsS₄
Etain	Stannite	Cu₂FeSnS₄
Fer	Arsénopyrite	FeAsS
	Pyrite	FeS₂
	Pyrrhotite	Fe_(1-x)S
Mercure	Cinabre	HgS
Molybdène	Molybdénite	MoS₂

Nickel	Polydymite	Ni₃S₄
	Violarite	FeNi₇S₄
Plomb	Galène	PbS
Thallium	Hutchinsonite	(PbTl)₂As₅S₉
Vanadium	Patronite	VS₄
Zinc	Sphalérite (Blende)	ZnS
	Wurtzite	ZnS

2. Mécanismes de formations

Les minéraux impliqués dans la formation des DMA se trouvent depuis leur formation dans des conditions réductrices qui les préservent de toute réaction. Mais les travaux miniers les placent brusquement en conditions oxydantes qui vont être les conditions initiatrices de la production des DMA. Au-delà de ces modifications physiques que subissent les matériaux d'exploitation, des réactions catalytiques chimiques et biologiques sont à l'origine de la formation, de la propagation et de l'auto-entretien des DMA.

2.1. Mécanismes physiques : importance des activités humaines

L'Homme joue un rôle prépondérant dans la production des DMA car il place d'énormes quantités de matériaux dans des conditions d'instabilité chimique. Les écoulements acides se forment ainsi à partir de la percolation d'eau sur les matériaux solides excavés et/ou dans les galeries en relation avec les mouvements de la nappe phréatique. En effet, les déchets produits, qui représentent généralement des millions de tonnes, sont souvent stockés humides, en conditions oxydantes et généralement exposés aux précipitations. Les phénomènes biologiques et chimiques de formation des DMA pour les déchets miniers s'enclenchent par suintement ou ruissèlement sur le lieu de stockage. D'autre part, dans le cas des exploitations souterraines, il y a formation de plusieurs kilomètres de galeries qui créent des conduits susceptibles de mettre en contact l'eau et l'oxygène avec le minerai. Lors de l'exploitation, si les galeries atteignent la nappe phréatique, l'eau est souvent pompée hors de la mine ce qui induit un abaissement du niveau piézométrique, provoquant l'apparition de conditions oxydantes là où le milieu était réducteur. Après l'abandon de l'exploitation, l'arrêt des pompes d'exhaure provoque la remontée des eaux et l'envoyage des galeries ce qui entraîne la dissolution des minéraux souvent riches en métaux. Au cours du temps, le circuit hydraulique retrouve des conditions réductrices et la qualité de l'eau s'améliore d'abord rapidement (quelques mois à quelques années) puis, plus lentement (quelques dizaines à quelques centaines d'années) (Ledoux, 1996).

2.2. Mécanismes chimiques : importance de l'environnement naturel

2.2.1. Les réactions primaires

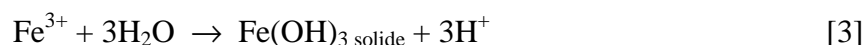
Le DMA est décrit comme un processus d'acidification des eaux de drainage dans l'environnement minier, résultant de l'oxydation de ces minéraux sulfurés sous l'effet des agents atmosphériques. Le minerai sulfuré que l'on retrouve le plus abondamment dans les rejets miniers est la pyrite (FeS_2). Ainsi pour illustrer le mode de génération du DMA, les mécanismes aboutissant à l'oxydation du soufre et du fer contenu dans la pyrite sont fréquemment utilisés. La réaction d'oxydation de la pyrite a été largement décrite par plusieurs auteurs (Singer & Stumm, 1970; Moses *et al.*, 1987; Baker & Banfield, 2003; Akcil & Koldas, 2006). Les minéraux sulfurés peuvent subir une oxydation chimique lente dite **oxydation directe** pour des pH proches de la neutralité (pH~7) (Evangelou, 1995), telle qu'illustrée par l'équation suivante (cas de la pyrite, Akcil & Koldas, 2006).



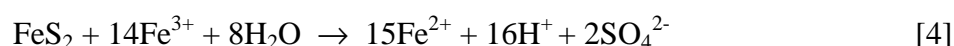
Il y a ainsi formation de fer ferreux Fe(II) et d'acide sulfurique H_2SO_4 ce qui occasionne une diminution du pH si aucun composé alcalin n'est présent dans le milieu. Si l'environnement est suffisamment oxydant (dépend de la concentration en oxygène, du pH et de l'activité des microorganismes), le fer ferreux libéré par l'oxydation de la pyrite va s'oxyder en fer ferrique Fe(III) selon la réaction suivante :



Pour des valeurs de pH > 3,5 le fer ferrique précipite alors sous forme d'hydroxyde de fer ($\text{Fe}(\text{OH})_3$), laissant peu de Fe(III) en solution tout en acidifiant davantage le milieu :

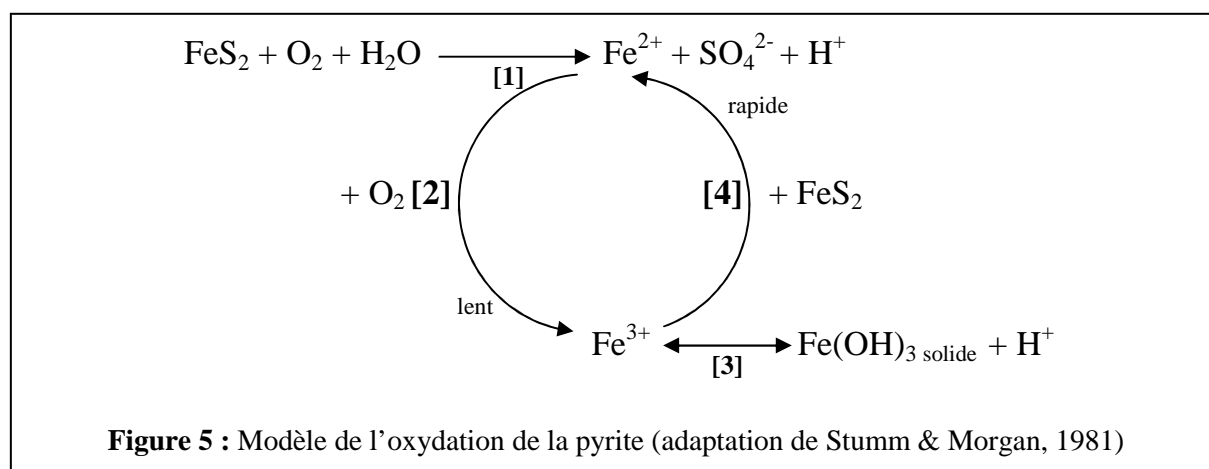


Pour des valeurs de pH < 3,5 la solubilité du fer ferrique augmente, lui permettant ainsi de contribuer à l'**oxydation indirecte** des sulfures :



Cette réaction marque le phénomène d'auto-entretien de la formation du DMA par l'établissement d'un processus cyclique mettant en jeu les réactions [2] et [4] (sans passer par la réaction [3]).

A pH acide (< 3), la réaction d'oxydation de la pyrite par le fer ferrique (réaction [4]) est de 10 à 100 fois plus rapide que l'oxydation directe par l'oxygène (réaction [1]) (Ritchie, 1994), cependant elle est limitée par la vitesse de régénération du fer ferrique qui dépend du pH. En effet, l'oxydation du fer ferreux en fer ferrique se produit naturellement mais très lentement à pH inférieur à 4. On remarquera également que lorsque l'oxygène est l'agent oxydant (réaction [1]), 2 moles d'ions H^+ sont produites par mole de pyrite oxydée alors que l'oxydation par le fer ferrique (réaction [4]) génère seize moles de protons par mole de pyrite. L'acidification du milieu est donc plus importante lorsque l'agent oxydant est le fer. Cependant, l'acidification du milieu peut être limitée par l'alcalinité résultant de la dissolution de minéraux basiques (carbonates et silicates) sachant que cela peut être un important mécanisme de neutralisation seulement à court terme (Johnson & Hallberg, 2003). L'ensemble du processus d'oxydation de la pyrite est schématisé dans la figure 5.

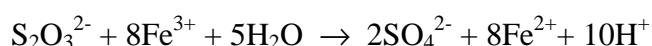
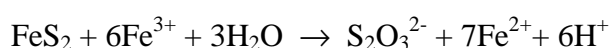


D'autres minéraux sulfurés participent également au phénomène du DMA comme la sphalérite (ZnS), la galène (PbS), la chalcopyrite ($CuFeS_2$), la pyrrhotite (Fe_7S_8) et l'arsénopyrite ($FeAsS$). Il est important de noter que plusieurs de ces minéraux peuvent contenir des métaux en concentrations importantes sous forme de substitution atomique ou encore en traces (exemple : Ni, As, Co dans la pyrite (FeS_2); Fe, Mn, Cd, Hg dans la sphalérite (ZnS)), lesquels pourront passer en solution lors de la dissolution des minéraux et affecter la composition du DMA.

De plus, selon Sand *et al.* (2001) la nature de ces sulfures de métaux permet de distinguer deux voies de lessivage : la voie du thiosulfate et celle des polysulfures.

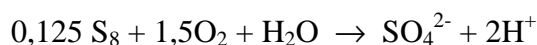
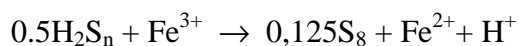
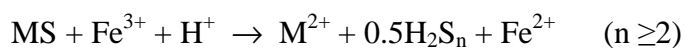
La voie des thiosulfates concerne uniquement les disulfures comme la pyrite (FeS_2), la molybdénite (MoS_2) et la tungsténite (WS_2). Ces sulfures de métaux ne sont dégradables que par une attaque oxydante, c'est à dire par les ions Fe(III) (Gehrke *et al.*, 1998; Schippers & Sand, 1999). La pyrite a été choisie comme modèle pour élucider le mécanisme d'oxydation et la formation de composés soufrés intermédiaires comme le thiosulfate ou le tétrathionate. Les protons sont générés lors de l'ajout d'oxygène aux espèces réduites du soufre. Comme le thiosulfate est le produit clé dans l'oxydation de la partie sulfure de la pyrite, le mécanisme a été défini comme la voie du thiosulfate. Les produits d'oxydation dans le cas de la pyrite et de la molybdénite sont constitués à plus de 90 % de sulfate et d'environ 1 à 2 % de polythionates.

Ce mécanisme est résumé par les équations suivantes :



La voie des polysulfures concerne les sulfures de métaux tels que l'arsénopyrite (FeAsS), le réalgar (As_4S_4), l'orpiment (As_2S_3), la galène (PbS), la chalcopyrite (CuFeS_2), la sphalérite (ZnS) et la hauérite (MnS_2). Ces minéraux sont dégradables par les ions Fe(III) mais également par les protons donc ils peuvent être dissous dans l'acide. La dissolution se fait par l'intermédiaire d'un radical H_2S .

Ce mécanisme est résumé par les équations suivantes :



2.2.2. Les réactions secondaires

La production de fer ferrique lors des réactions précédentes entraîne une acidification supplémentaire par hydrolyse et la formation d'acide sulfurique. Les nouvelles conditions physicochimiques générées par les réactions du DMA donnent lieu ensuite à des réactions

d'oxydation, de dissolution, de précipitation et de recristallisations successives qui produisent de nombreux composés minéraux secondaires. Les deux principaux types de minéraux secondaires formés sont des oxyhydroxydes de fer (hématite, magnétite, goethite) et des hydroxysulfates de fer (jarosite, schwertmannite, copiapite) (Hudson-Edwards *et al.*, 1999; Morin *et al.*, 2003). Ces minéraux secondaires qui contiennent des traces métalliques peuvent précipiter dans le lit des rivières en périodes sèches et constituent ainsi des réserves temporaires de métaux. Au cours de forts événements pluvieux (pluies intenses ou encore inondations) ces éléments peuvent être partiellement solubilisés et entraînés également sous forme particulaire en aval des cours d'eau. Leur dissolution entraîne la libération de polluants métalliques et de sulfates, ainsi qu'une acidification de l'eau. Des « stocks » à plus long terme sont constitués par les oxy-hydroxysulfates et hydroxydes de fer dit « insolubles » comme la jarosite et la goethite (Hudson-Edwards *et al.*, 1999).

2.3. Mécanismes biologiques : importance des microorganismes

L'apparition des DMA est favorisée par la présence de microorganismes qui tirent l'énergie nécessaire à leur croissance des réactions d'oxydoréduction métalliques. Ces microorganismes agissent comme de véritables catalyseurs des réactions chimiques qui se déroulent à la surface des minéraux et en solution. En effet la vitesse d'oxydation de la pyrite dépend de la disponibilité en Fe(III). Or bien que la réaction [3] puisse se faire naturellement d'un point de vue chimique, celle-ci est très lente voire négligeable à pH acide ($\text{pH} < 4$) (Baker & Banfield, 2003; Johnson & Hallberg, 2005). Elle nécessite donc une catalyse qui peut être générée par les microorganismes naturellement présents dans ces milieux. Ainsi, des procaryotes tels que *Acidithiobacillus ferrooxidans* ou *Ferroplasma acidophilum* catalysent l'oxydation du fer ferreux et augmentent ainsi la vitesse de dissolution de la pyrite (Silverman *et al.*, 1964; Nordstrom & Southam, 1997). C'est ce que l'on appelle le **biolessivage** ou **biolixiviation**. Singer & Stumm (1970), ont montré que les microorganismes accélèrent par un facteur 10^6 la vitesse d'oxydation du fer à faible pH par rapport à la vitesse résultant de la réaction chimique seule (réaction [3]). Selon Edwards *et al.* (2000a), l'activité microbienne serait à l'origine d'environ 75 % de la production des DMA.

De nombreux travaux ont été réalisés pour élucider les procédés de dégradation à l'interface des sulfures de métaux (Fowler *et al.*, 1999; Sand *et al.*, 2001; Tributsch, 2001; Kinzler *et al.*, 2003; Zhang & Fang, 2005). Cependant, un débat persiste à propos des mécanismes mis en

jeu lors des interactions se produisant entre le minéral soufré et les microorganismes (Baker & Banfield, 2003). L'oxydation de la pyrite pourrait se dérouler soit par des voies « directes » soit par des voies « indirectes » (Sand *et al.*, 2001). Le mécanisme « **direct** » décrit les interactions enzymatiques des bactéries liées à la surface de la pyrite *via* la présence de substances polymériques extracellulaires (EPS). Larsson *et al.* (1993), ont montré que l'étroite proximité des cellules bactériennes à la surface solide était nécessaire pour des taux de croissance et d'oxydation optimaux. Il a été montré que les taux de dissolution de la pyrite étaient augmentés et accélérés à proximité de cellules de *Ferroplasma acidarmanus* (Edwards *et al.*, 2001) et d'*Acidithiobacillus ferrooxidans* (Fowler *et al.*, 1999; Edwards *et al.*, 2001). Le mécanisme « **indirect** » se traduit par l'oxydation catalytique du fer ferreux en fer ferrique par les bactéries suivie de l'oxydation directe des sulfures par les ions ferriques ainsi produit. Certains effets enzymatiques sont impliqués dans le mécanisme direct, alors que le mécanisme indirect n'inclue que des phénomènes de dissolution chimique des sulfures par les ions ferriques et l'acide sulfurique produit par les bactéries (Zhang & Fang, 2005).

2.4. Facteurs influençant la formation du DMA

La qualité chimique des DMA varie et est dépendante des propriétés physiques, chimiques, minéralogiques et microbiologiques de chaque site. Il existe ainsi de multiples combinaisons de conditions pouvant favoriser la formation des DMA. Cependant il est possible de classer les facteurs influençant l'apparition des eaux acides par ordre d'importance.

2.4.1. Les facteurs primaires

Les facteurs primaires sont à l'origine du phénomène et leurs variations contrôlent les mécanismes chimiques :

- la disponibilité en eau pour l'oxydation et le transport,
- la disponibilité en oxygène,
- les caractéristiques physiques du solide (minéralogie, géologie et altérabilité géochimique).

A noter que d'autres facteurs peuvent intervenir dans les cinétiques de ces mécanismes initiateurs :

- la température,

- l'activité microbiologique.

2.4.2. Les facteurs secondaires

Les facteurs secondaires interviennent une fois le mécanisme de formation d'acidité en place :

- la présence de minéraux susceptibles de neutraliser l'acidité comme les carbonates ;
- l'influence du pH sur l'équilibre Fe(II)/ Fe(III). Inférieur à 3,5 le fer reste en solution et est oxydant, supérieur à cette valeur, il précipite sous forme d'hydroxydes ;
- la pression partielle en l'oxygène. Si elle croît, le taux d'oxydation croît. Cet effet est plus prononcé quand les bactéries sont actives parce que de nombreux microorganismes ont besoin d'oxygène pour leur respiration ;
- la taille des particules dans le cas d'une percolation ;
- les effets galvaniques ayant pour origine les interactions électrochimiques entre les minéraux.

3. Mécanismes d'atténuation

Parallèlement au biolessivage, les microorganismes interviennent également dans des processus de bioremédiation ou piégeage d'éléments toxiques en solution (Edwards *et al.*, 2000a; Hallberg, 2010). Plusieurs mécanismes permettant la réduction de la concentration en éléments métalliques dans les DMA ont été identifiés. Ces mécanismes peuvent être classés en deux catégories, ceux impliquant l'immobilisation des métaux (sorption, séquestration intracellulaire, précipitation, cristallisation) et ceux impliquant leur mobilisation (méthylation, chélation par des métabolites microbiens ou des sidérophores) (Gadd, 2004). Nous ne présenterons ici que quelques exemples de ces mécanismes.

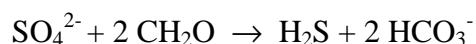
3.1. Précipitation

3.1.1. La production de sulfures métalliques

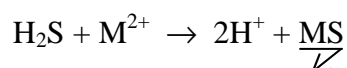
L'un des principaux mécanismes impliqués dans l'atténuation naturelle des DMA est la précipitation de sulfures et de carbonates catalysés par des Bactéries Sulfato-Réductrices (BSR) dans la zone anaérobie (Jong & Parry, 2003; Amils *et al.*, 2007; Qiu *et al.*, 2009). En effet, les ions sulfures produits lors de la sulfato-réduction peuvent interagir avec les ions

métalliques et les métalloïdes présents dans le milieu et précipiter ainsi sous forme de sulfures métalliques fortement insolubles (Neculita *et al.*, 2007).

Lors de la sulfato-réduction, les BSR couplent l'oxydation de composés organiques simples (CH_2O) à la réduction du sulfate (SO_4^{2-}), produisant des sulfures (H_2S) et du bicarbonate (HCO_3^-) selon la réaction suivante :



Le H_2S ainsi produit est très réactif avec les métaux et métalloïdes (M^{2+}) présents dans le milieu et va ainsi permettre leur précipitation sous forme de sulfures métalliques (MS) :



La réaction de précipitation sera favorisée par l'augmentation du pH due à la libération de bicarbonate lors de la réduction des sulfates (Hallberg & Johnson, 2005). Les métaux et métalloïdes toxiques ainsi immobilisés au fond des rivières permettent donc une atténuation de la pollution dans les eaux. Ainsi, de par leur capacité à produire d'importantes quantités d'ions sulfures, les BSR présentent un fort potentiel pour les étapes de bioremédiation (White & Gadd, 1996).

3.1.2. La co-précipitation

Comme nous l'avons vu précédemment, certains microorganismes ont une activité métabolique leur permettant de catalyser l'oxydation du Fe(II) en Fe(III) (réaction [2]). Pour des $\text{pH} > 3,5$, le Fe(III) présent dans l'eau précipite sous forme d'hydroxydes de fer ($\text{Fe}(\text{OH})_3$) (réaction [3]). Or, ces hydroxydes de fer sont capables d'incorporer des éléments toxiques dans leur structure ou de les adsorber à leur surface (Ferris *et al.*, 2000; Lee *et al.*, 2002; Xu *et al.*, 2007). Ce type de mécanisme permet donc d'immobiliser les métaux et métalloïdes ayant une forte affinité pour les oxydes de fer, comme par exemple l'As et le Pb (Casiot *et al.*, 2003a; Morin *et al.*, 2003; Chen *et al.*, 2007). Il a en effet été montré que la précipitation du Fe(III) pouvait entraîner la formation de biominéraux de Fe(III)-As(III) ou -As(V) (Casiot *et al.*, 2003a; Morin *et al.*, 2003).

3.2. Séquestration

3.2.1. La séquestration extracellulaire

La séquestration extracellulaire met en jeu des molécules appelées Exo-Polysaccharides (EPS), un mélange de polysaccharides, de mucopolysaccharides et de protéines. Ces molécules sont synthétisées par des microorganismes, des algues et des champignons et peuvent se lier à des métaux toxiques (Gadd, 2004). En effet, Haack & Warren (2003), ont mis en évidence l'accumulation stable de métaux au sein de biofilms microbiens issus d'un drainage minier acide. Les biofilms sont des assemblages d'une ou plusieurs populations attachées à des surfaces biotiques ou abiotiques par l'intermédiaire d'exo-polymères. Ainsi, ces microorganismes qui sécrètent des polymères sont capables d'immobiliser des composés métalliques par des processus de séquestration passive. C'est le cas notamment d'une bactérie appelée *Herminiimonas arsenicoxydans* capable de séquestrer de l'arsenic dans une capsule exo-polysaccharidique (Muller *et al.*, 2007; Lièvreumont *et al.*, 2009).

3.2.2. La séquestration intracellulaire

La séquestration intracellulaire est un processus actif qui se traduit par l'accumulation de composés métalliques à l'intérieur des cellules via des systèmes de transport membranaires. Le transport membranaire varie en fonction de l'affinité et de la spécificité des éléments métalliques. Par exemple, le césium peut être accumulé par les microorganismes *via les* transporteurs du potassium (Bossemeyer *et al.*, 1989) et de l'ammonium (Avery *et al.*, 1992). Une fois à l'intérieur de la cellule, les espèces métalliques peuvent être liées, précipitées, immobilisées à l'intérieur de structures intracellulaires ou d'organites, ou translocalisées vers des structures spécifiques (Gadd, 2004). Ceci varie en fonction de l'élément concerné et de l'organisme impliqué. Les groupes carboxyles du peptidoglycane sont par exemple les principaux sites de liaison chez les bactéries Gram positif tandis que chez les bactéries Gram négatif ce sont les groupements phosphate et la chitine pour les champignons.

3.3. Biométhylation

La méthylation est l'encrage ou la substitution d'un groupement méthyle (-CH₃) sur un substrat. La méthylation d'éléments métalliques tels que l'arsenic, le mercure, l'étain, le sélénium, le tellure et le plomb peut être induite par action enzymatique de différentes bactéries et champignons dans des conditions aérobies et anaérobies. Les composés

métalliques méthylés formés par ce processus diffèrent dans leur solubilité, leur volatilité et leur toxicité selon le métal de base et l'action enzymatique microbienne (Gadd, 2004). Cette modification enzymatique permet de rendre certains éléments métalliques moins toxiques (Casiot *et al.*, 2004) ou plus volatiles. C'est le cas par exemple de certaines formes d'arsenic qui sont moins toxiques sous forme méthylée que sous forme inorganique. Il a ainsi été montré que plusieurs espèces de bactéries et de champignons étaient capables de méthyler des composés arséniés tels que l'arséniate (As(V), AsO_4^{3-}), l'arsénite (As(III), AsO_2^-) et l'acide méthylarsonique ($\text{CH}_3\text{H}_2\text{AsO}_3$) en diméthyl- ou triméthyl-arsine ($(\text{CH}_3)_2\text{HAs}$, $(\text{CH}_3)_3\text{As}$) volatiles (Tamaki & Frankenberger, 1992). Cependant, il faut également noter que certains éléments deviennent plus nocifs une fois méthylés. C'est le cas du mercure (Hg) et de certaines formes de l'As comme les méthylarsines volatiles.

3.4. Sidérophores

Les sidérophores sont des ligands spécifiques du Fe(III). Ces molécules de coordination sont excrétées pour aider à l'assimilation du fer en facilitant le contact avec le substrat par attachement aux oxydes de fer (Gadd, 2004). Cependant, bien qu'initialement produit dans le but d'obtenir du fer, ces molécules peuvent également se lier à d'autres métaux tels que le magnésium, la manganèse, le chrome (III) ou le galium (III) (Birch & Bachofen, 1990).

3.5. Oxydation de l'arsenic

D'autres réactions permettant de réduire la toxicité des DMA interviennent également comme l'oxydation de l'arsenic. En effet, l'arsenic est considéré comme plus toxique sous sa forme As(III) et il est également plus soluble et donc plus mobile que sous la forme As(V) (Johnson & Hallberg, 2005). Il a par exemple été montré que le Fe(III) se lie préférentiellement à l'As(V) à pH acide, entraînant la formation de précipités permettant ainsi la diminution de la biodisponibilité de l'As dans l'environnement (Casiot *et al.*, 2003a). Les microorganismes oxydant l'arsenic présentent ainsi un grand intérêt environnemental en termes de biorémediation des hydrosystèmes et sites contaminés.

4. Microorganismes présents dans les drainages miniers acides

4.1. Diversités phylogénétiques et physiologiques

Les DMA présentent des caractéristiques physicochimiques (pH, température, concentration en oxygène, en sulfate...) très variables qui dépendent de leur origine et du type de minerai extrait (Table 2). En général, en raison de la teneur relativement élevée en carbonate dans la roche mère, les drainages issus des mines de charbon ont tendance à être moins acides et plus pauvres en éléments métalliques toxiques que les drainages des mines de métaux (Hallberg, 2010). D'autre part, la variabilité des pH affecte les concentrations en métaux dissous. En effet, plus le pH est acide, plus la concentration en métal est importante (e.g. mine de Ynysawred versus mine de Richmond, Tableau 2). De telles variations ont un impact important sur la diversité microbienne qui diffère elle aussi d'un site à l'autre.

Tableau 2 : Caractéristiques physicochimiques de quelques sites miniers répartis dans le monde (adapté de Johnson & Hallberg, 2003 et Hallberg, 2010)

	Ynysawred, Pays de Galles ^a	Wheal Jane, Angleterre ^a	King, Norvège ^b	Parys, Pays de Galles ^a	Río Tinto, Espagne ^c	Richmond, USA ^d	Carnoulès, France ^e
Type de mine	Charbon	Étain	Cuivre	Cuivre	Cuivre-Argent-Or	Cuivre	Plomb-Zinc
pH	6,2	3,4	2,7	2,5	2,2	0,5-1	1,2-8
Eh (mV)	257	462	—	285	450	—	367-877
Fe total	160	290	172	650	$2,3 \times 10^3$	$13-19 \times 10^3$	$1-10 \times 10^3$
Fe ²⁺	140	250	—	650	$1,5 \times 10^3$	$13-19 \times 10^3$	$0,8-6,6 \times 10^3$
SO ₄ ²⁻	464	400	668	1550	10×10^3	$20-100 \times 10^3$	$3-29 \times 10^3$
Al	—	27	22,5	70	—	$1,4-6,7 \times 10^3$	0,4-170
Cu	—	1	16	40	109	120-650	0,1-13
Zn	—	132	25	60	225	700-2600	20-60

Note : toutes les concentrations sont données en mg L⁻¹ ; — : non déterminé

(a) Johnson *et al.*, 2002 ; (b) Johnson *et al.*, 2001 ; (c) López-Archilla *et al.*, 2001 ; (d) Nordstrom *et al.*, 1999 ; (e) Casiot *et al.*, 2003a; Casiot *et al.*, 2005; Casiot *et al.*, 2009.

Malgré leur toxicité (acidité, métaux, métalloïdes), les DMA présentent de nombreuses niches microbiennes mais la structure des communautés est généralement réduite à un nombre limité d'espèces (Baker & Banfield, 2003). Ceci est probablement le résultat des sources d'énergies disponibles qui sont peu nombreuses et des variations en température, pH et force ionique présentes dans ces environnements. Parmi les microorganismes présents dans les DMA, on retrouve des eucaryotes et des procaryotes (bactéries et Archaea) aux capacités métaboliques particulières en raison des conditions extrêmes présentes au sein de ces environnements. De part leurs intérêts biotechnologiques (biolixiviation et bioremédiation), de nombreuses études

se sont focalisées sur leur isolement, leur identification, leur diversité ainsi que sur les mécanismes permettant leur survie dans ces milieux extrêmes (Johnson & Hallberg, 2003; Golyshina & Timmis, 2005; Bruneel *et al.*, 2006; Bruneel *et al.*, 2008; Baker *et al.*, 2009; Qiu *et al.*, 2009). Il est important de souligner que c'est principalement l'introduction des techniques de biologie moléculaire basées sur les gènes codant l'ARNr 16S et 18S qui a permis d'accroître nos connaissances sur la diversité microbienne au sein des DMA. De même, les récentes techniques de séquençage à haut débit devraient encore affiner ces connaissances en permettant d'obtenir une évaluation plus exhaustive et plus précise des microorganismes présents.

4.2. Les procaryotes

Les microorganismes procaryotes métaboliquement actifs dans les environnements acides tels que les DMA appartiennent aux domaines *Bacteria* et *Archaea*. La diversité physiologique et phylogénétique de ces acidophiles a été très étudiée pour les bactéries (Hallberg & Johnson, 2001; Baker & Banfield, 2003; Johnson & Hallberg, 2003). Les acidophiles ont été classés selon différents critères prenant en compte leur optimum de pH et de température mais également la source de carbone qu'ils utilisent. La température optimale de croissance permet de distinguer les mésophiles ($T_{opt} < 40\text{ °C}$), les thermophiles modérés ($T_{opt} 40\text{-}60\text{ °C}$) et les thermophiles extrêmes ($T_{opt} > 60\text{ °C}$) alors que la source de carbone utilisée différencie les autotrophes qui assimilent le CO_2 , les hétérotrophes qui assimilent le carbone organique et les mixotrophes qui utilisent les deux.

4.2.1. Les bactéries

Parmi les lignées bactériennes les plus communes, on trouve au sein des DMA de nombreux représentants appartenant au phylum des *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *Nitrospira* et *Firmicutes* (Figure 6). Le groupe le plus étudié à ce jour appartient à la classe des *Gammaproteobacteria*, comme *Acidithiobacillus ferrooxidans* et *A. thiooxidans* qui ont été isolés.

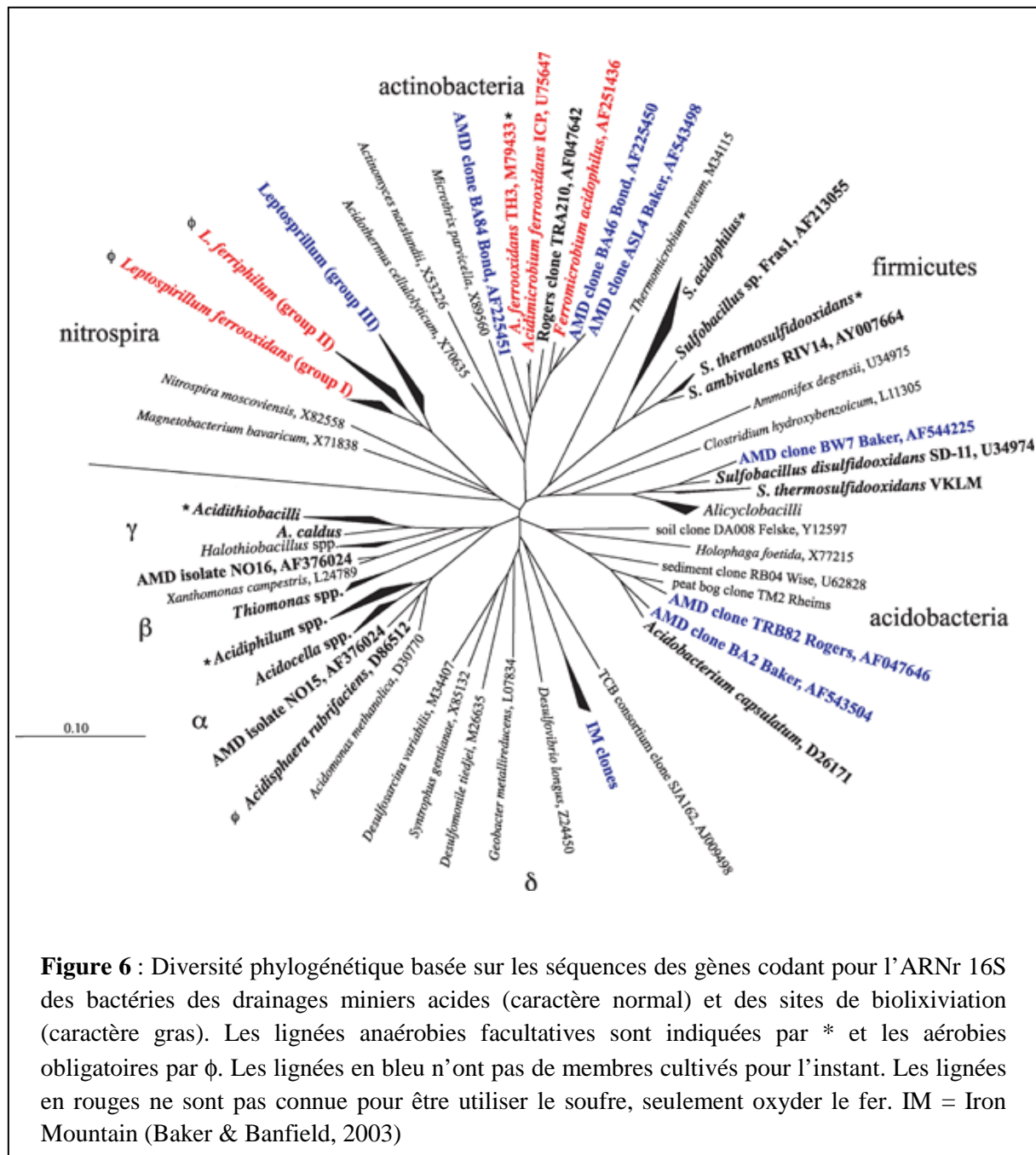


Figure 6 : Diversité phylogénétique basée sur les séquences des gènes codant pour l'ARNr 16S des bactéries des drainages miniers acides (caractère normal) et des sites de biolixiviation (caractère gras). Les lignées anaérobies facultatives sont indiquées par * et les aérobies obligatoires par φ. Les lignées en bleu n'ont pas de membres cultivés pour l'instant. Les lignées en rouges ne sont pas connue pour être utiliser le soufre, seulement oxyder le fer. IM = Iron Mountain (Baker & Banfield, 2003)

La multiplication des études de diversité réalisées au sein des DMA basées à la fois sur des approches moléculaires et culturales ont permis ces dernières années de détecter et d'isoler de nombreux microorganismes procaryotes aux capacités physiologiques diverses dont une liste non exhaustive est présentée au tableau 3.

Tableau 3 : Bactéries détectées au sein de drainages miniers acides

Phylum/Classe	Genre/espèces	Oxydation :			Réduction :		Références :
		As	Fe ²⁺	S	Fe ³⁺	SO ₄ ²⁻	
Proteobacteria							
<i>α-proteobacteria</i>	<i>Acidiphilium cryptum</i>			+	+		(a)
	<i>Acidiphilium</i> spp.			+	+		(b);(c);(d)
	<i>Acidocella</i> spp.				+		(b);(c);(d);(e)
	<i>Acidisphaera rubrifaciens</i>						(f)
<i>β-proteobacteria</i>	<i>Ferrovum myxofaciens</i>		+				(g)
	<i>Thiomonas</i> spp.		+	+			(h)
	<i>Thiomonas arsenivorans</i>	+	+	+			(i)
	<i>Gallionella ferruginea</i>		+				(q); (y)
	<i>Thiobacillus</i> spp.		+	+			(h)
	<i>Ferribacter polymixa</i>		+				(k)
<i>γ-proteobacteria</i>	<i>Acidithiobacillus thiooxidans</i>			+			(l);(m)
	<i>Acidithiobacillus ferrooxidans</i>		+	+	+		(l);(m)
	<i>Acidithiobacillus ferrivorans</i>		+	+			(n)
	<i>Acidithiobacillus caldus</i>			+			(o)
	<i>Halothiobacillus</i> sp.			+			(n)
<i>δ-proteobacteria</i>	<i>Desulfosarcina</i> spp.					+	(p)
	<i>Desulmonile tiedjei</i>					+	(q)
	<i>Desulfuromonas svalbardensis</i>				+	+	(g)
	<i>Syntrophobacter</i> spp.					+	(r)
	<i>Desulfurella</i> spp.					+	(r)
Acidobacteria							
	<i>Acidobacterium capsulatum</i>				+		(s)
	<i>Acidobacterium</i> spp.				+		(e)
Actinobacteria							
	<i>Acidimicrobium</i> spp.		+				(o);(i)
	<i>Acidimicrobium ferrooxidans</i>		+		+		(o)
	<i>Ferrimicrobium acidiphilum</i>		+		+		(t)
	<i>Ferrithrix thermotolerans</i>		+		+		(t)
Nitrospira							
	<i>Leptospirillum ferrooxidans</i>		+				(l);(o)
	<i>Leptospirillum ferrodiazotrophum</i>		+				(u)
	<i>Leptospirillum</i> spp.		+				(b);(c);(d);(g)
Firmicutes							
	<i>Alicyclobacillus aeris</i>		+	+	+		(j)
	<i>Alicyclobacillus disulfidooxidans</i>		+	+	+		(o)
	<i>Clostridium aciditolerans</i>						(v)
	<i>Sulfobacillus thermotolerans</i>		+	+			(w)
	<i>Sulfobacillus</i> spp.			+	+		(x)
	<i>Desulfosporosinus</i> spp.		+		+	+	(r)

(a) Küsel *et al.*, 1999 ; (b) Johnson *et al.*, 2001 ; (c) Johnson *et al.*, 2002 ; (d) López-Archilla *et al.*, 2001 ; (e) Lear *et al.*, 2009 ; (f) Hiraishi *et al.*, 2000 ; (g) Bruneel *et al.*, 2011 ; (h) Bruneel *et al.*, 2003 ; (i) Battaglia-Brunet *et al.*, 2006 ; (j) Guo *et al.*, 2009 ; (k) Heinzl *et al.*, 2009 ; (l) Natarajan, 2008 ; (m) Campos Paulino *et al.*, 2001 ; (n) Hallberg *et al.*, 2010 ; (o) Baker & Banfield, 2003 ; (p) Bruneel *et al.*, 2005 ; (q) Bruneel *et al.*, 2006 ; (r) Sánchez-Andrea *et al.*, 2011 ; (s) Kishimoto *et al.*, 1991 ; (t) Johnson *et al.*, 2009 ; (u) Tyson *et al.*, 2005 ; (v) Lee *et al.*, 2007 ; (w) He *et al.*, 2007 ; (x) Schippers *et al.*, 2010 ; (y) Hallbeck *et al.*, 1993.

4.2.2. Les Archaea

Alors qu'un grand nombre de connaissances ont été acquises concernant les bactéries acidophiles indigènes aux DMA, relativement peu d'informations sont disponibles concernant les Archaea se développant dans ces écosystèmes. Ces organismes, connus à l'origine pour

leurs capacités à se développer dans des environnements extrêmes et des niches spécialisées (pH~0, température > 100 °C), sont également présents au sein d'environnements non-extrêmes comme les sols marins ou terrestres (Chaban *et al.*, 2006). Leurs exceptionnelles capacités d'adaptations permettent ainsi à ces microorganismes de se développer dans les DMA où ils peuvent constituer plus de 50 % et même jusqu'à 85 % de la population procaryote totale (Edwards *et al.*, 2000b).

Les Archaea identifiées au sein des DMA incluent le phylum des *Euryarchaeota* et le phylum récemment décrit des *Thaumarchaeota* (Brochier-Armanet *et al.*, 2008). Parmi les *Euryarchaeota*, l'ordre des *Thermoplasmatales* est le plus fréquemment rencontré au sein des DMA. En effet, des Archaea appartenant au genre *Ferroplasma* ont été identifiées dans différents DMA à travers le monde (Hallberg, 2010), parmi lesquelles *F. acidiphilium* et *F. acidarmanus* (Edwards *et al.*, 2000b; Golyshina *et al.*, 2000). Ces Archaea mésophiles dépourvues de paroi cellulaire sont résistantes à l'arsenic (Gihring *et al.*, 2003) et sont capables d'oxyder le fer. D'autres Archaea appartenant également aux *Thermoplasmatales* ont été détectées, cependant elles ne présentaient pas d'affiliations suffisamment importantes avec des espèces connues pour pouvoir en déduire des informations à propos de leur propriétés physiologiques (Hallberg, 2010). Une espèce appartenant aux *Sulfolobales* (*Metallosphaera prunae*) a été isolée d'une mine située en Allemagne (Fuchs *et al.*, 1995). Il s'agit d'une Archaea thermophile extrême capable de se développer à des températures comprises entre 55 et 80 °C par oxydation du soufre, de la pyrite, de la sphalérite, de la chalcopyrite ou de l'hydrogène moléculaire. La présence d'Archaea méthanogènes appartenant aux *Methanosarcinales* et aux *Methanobacteriales* a également été mise en évidence dans les sédiments du DMA du Río Tinto (Sanz *et al.*, 2011). Cette étude a permis d'identifier trois espèces, *Methanosaeta concilii*, *Methanosarcina barkeri* et *Methanobacterium bryantii*, qui semblent prospérer au sein de micro-niches caractérisées par des conditions réductrices et acides peu importantes alors que ces sédiments sont immergés dans un environnement oxydant extrêmement acide. Enfin, concernant les *Thaumarchaeota*, une étude récente a mis en évidence des Archaea appartenant à l'ordre des *Nitrososphaerales* (Halter *et al.*, 2011). L'espèce détectée, *Candidatus Nitrososphaera gargensis* est impliquée dans l'oxydation de l'ammonium (AOA), une étape clé du cycle de l'azote (Brochier-Armanet *et al.*, 2011). L'implication de certaines Archaea dans les cycles du fer et du soufre laisse supposer que ces

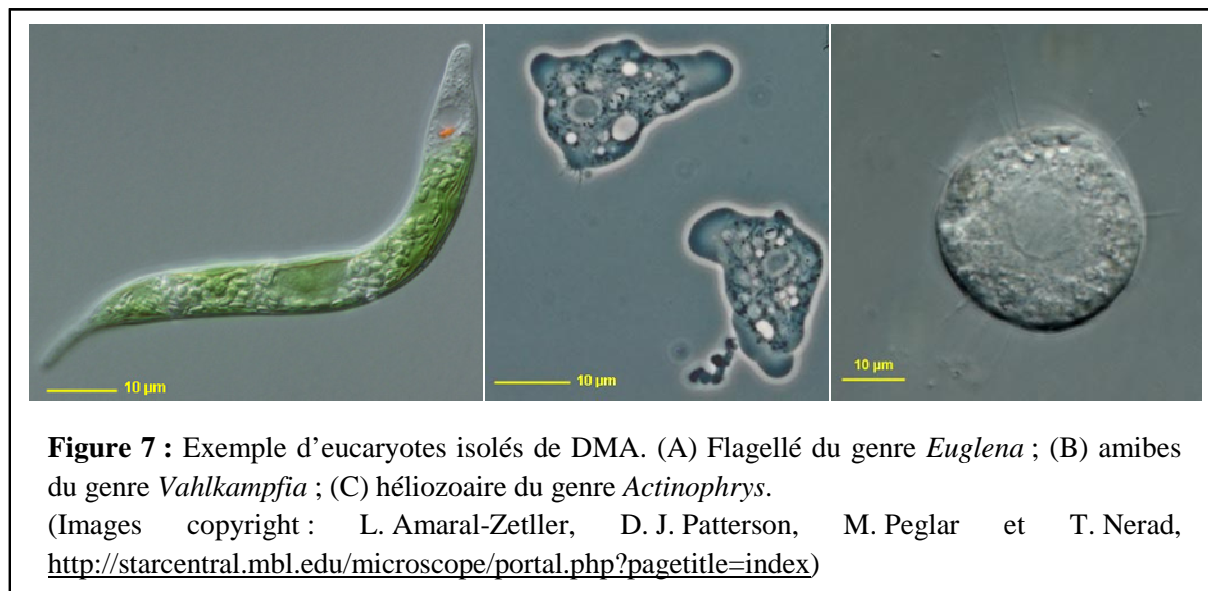
organismes pourraient par conséquent jouer un rôle majeur dans la génération et la remédiation des DMA.

4.3. Les eucaryotes

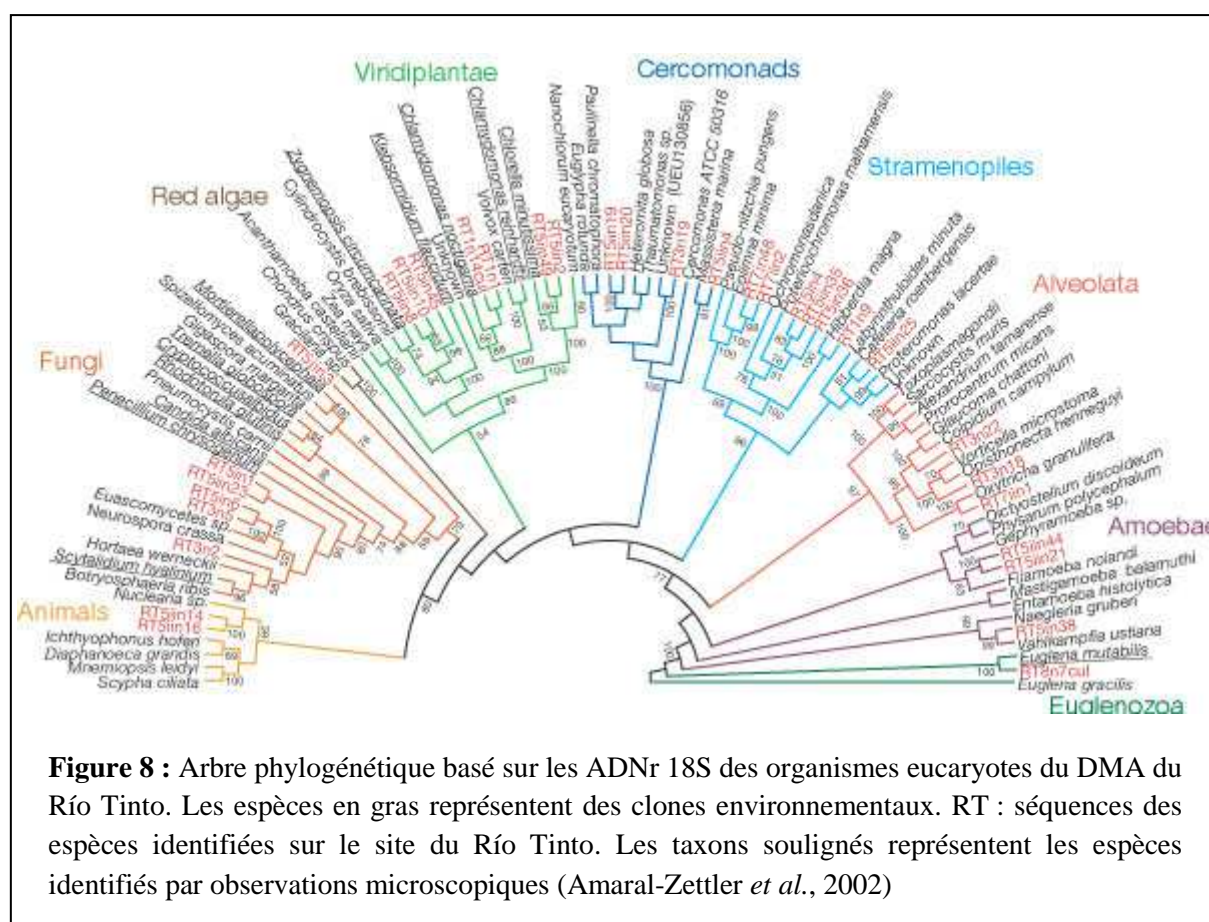
La diversité des eucaryotes au niveau des DMA est encore mal connue car jusqu'à présent très peu d'études se sont intéressées à ces microorganismes qui sont pourtant susceptibles de jouer un rôle important au sein de ces écosystèmes. Des études ont en effet montré, que certains de ces microorganismes sont capables de modifier les conditions physicochimiques de ces écosystèmes en générant de l'alcalinité (de façon directe ou indirecte) ou bien en modifiant la toxicité et/ou la mobilité de certains métaux et métalloïdes (Das *et al.*, 2009). D'autre part, certains protozoaires photosynthétiques comme *Euglena mutabilis* sont capables de stocker le fer sous forme de précipités intracellulaires de lépidocrocite et d'accumuler/adsorber d'autres ETM (As, Pb, Hg, Cd) (Devars *et al.*, 1998; Casiot *et al.*, 2004). Les processus de croissance et de photosynthèse de ce protozoaire sont également susceptibles de modifier les conditions physicochimiques du milieu en contribuant à la saturation en oxygène dissous jusqu'à 200 % (Brake *et al.*, 2001). Les eucaryotes photosynthétiques comme *Euglena mutabilis* jouent également un rôle primordial en tant que producteurs primaires car ils sont responsables de l'apport de carbone organique permettant le maintien des populations procaryotes hétérotrophes comme les BSR (Halter *et al.*, 2012a). Les algues produisent des zones anoxiques utiles au métabolisme des BSR et aident à la génération biogénique d'alcalinité (Das *et al.*, 2009). De même, les champignons peuvent avoir un rôle important dans l'organisation et l'ancrage des biofilms microbiens sur la pyrite (Baker *et al.*, 2009) mais aussi dans l'absorption de métaux dans leur paroi cellulaire ou par des exopolysaccharides.

Les principales études portant sur les eucaryotes ont été réalisées sur le site du Río Tinto situé au sud-ouest de l'Espagne (López-Archilla *et al.*, 2001; Amaral-Zettler *et al.*, 2002; Aguilera *et al.*, 2006, 2007) et sur le site de Richmond situé en Californie (Baker *et al.*, 2009). En raison de difficultés de mise en culture, peu de microorganismes eucaryotes ont été isolés à partir des DMA. Des études ont cependant permis l'isolement de ciliés appartenant au genre *Cinetochilium*, des amibes du genre *Vahlkampfia* et des flagellés euglénoides du genre *Eutreptia* se nourrissant de bactéries acidophiles oxydant le fer (Baker & Banfield, 2003). Trois isolats d'un champignon unicellulaire appartenant au genre *Acidomyces* ont été mis en évidence sur le site de Richmond (Baker *et al.*, 2004). De même, des flagellés du genre

Euglena et des actinopodes héliozoaires du genre *Actinophrys* ont été isolés sur le site du Río Tinto (Figure 7) (Amaral-Zettler *et al.*, 2003). Le genre *Euglena* a également été mis en évidence sur le site de Carnoulès où la souche *Euglena mutabilis* a été isolée (Halter *et al.*, 2012b).



En complément de ces méthodes culture-dépendantes, des études basées sur des techniques de biologie moléculaire ont par la suite permis d'approfondir les connaissances sur la diversité eucaryote présente au sein des DMA. Ces études ont montré que les communautés eucaryotes présentent une certaine variabilité selon les sites, les saisons et/ou les conditions physicochimiques des DMA. Ainsi, une étude couplant analyse de séquences d'ADNr 18S, observations microscopiques et isolements a permis de révéler une diversité insoupçonnée sur le site du Río Tinto (pH~2) où ces microorganismes représentent plus de 65 % de la biomasse totale (Amaral-Zettler *et al.*, 2002). Les principales espèces identifiées dans cette étude appartiennent à divers groupes phylogénétiques présentés dans la figure 8.



La communauté eucaryote décrite dans ce DMA est représentée par de nombreux biofilms qui constituent la plus grande partie des communautés microbiennes phototrophes de cet écosystème (Souza-Egipsy *et al.*, 2011). Les espèces photosynthétiques les plus fréquentes sont représentées par des chlorophytes (algues vertes) du genre *Chlamydomonas*, *Dunaliella*, *Chlorella* et *Stichococcus* et par des euglénides du genre *Euglena*. On trouve ensuite des streptophytes du genre *Klebsormidium*, *Zygnemopsis* et *Mesotaenium*. Des représentants des rhodophytes (*Cyanidium*) et des diatomées du genre *Pinnularia* sont également présents. Concernant les décomposeurs on trouve une grande diversité de champignons incluant des levures et des formes filamenteuses. La communauté mixotrophe est dominée par des cercomonades et des straménopiles appartenant aux genres *Bodo*, *Ochromonas*, *Labyrinthula* et *Cercomonas*. Les protistes sont caractérisés par deux espèces de ciliés appartenant aux genres *Oxytricha* et *Euplotes*. Enfin, des amibes des genres *Valhkampfia* et *Naegleria* sont fréquemment retrouvés ainsi qu'une espèce d'héliozoaire appartenant au genre *Actinophyris* (Amils *et al.*, 2007). Il semblerait par ailleurs que ce soit la présence d'organismes photosynthétiques qui soit à l'origine de l'existence de cette importante diversité de microorganismes eucaryotes (Baker *et al.*, 2004). Paradoxalement, sur le site de

Richmond (pH < 0,9), une faible diversité d'eucaryotes essentiellement dominée par des champignons a été mise en évidence par l'analyse phylogénétique d'une banque d'ADNr 18S (Baker *et al.*, 2009). Des protistes affiliés au clade des protistes acidophiles (Acidophilic Protist Clade, APC) ainsi que des percolozoaires de la classe des *Heterolobosea* ont également été identifiés. Cette faible diversité eucaryote serait quand à elle liée à une faible diversité de substrats organiques dérivés des bactéries et des Archaea et à un manque de niches photosynthétiques (Baker *et al.*, 2004, 2009). Étonnamment, la plupart des eucaryotes mis en évidence dans cette étude (sauf les APC) sont phylogénétiquement proches d'organismes neutrophiles ce qui suggère une adaptation récente à ce type de milieu extrême, contrairement aux procaryotes qui y seraient adaptés depuis des millions, voire des milliards d'années (Baker *et al.*, 2009).

5. Communautés mixtes et interactions microbiennes

Les microorganismes acidophiles existent sous forme de populations mixtes à la fois dans les environnements naturels et anthropiques. Les interactions entre les membres d'un consortium microbien sont primordiales pour l'optimisation de l'activité des communautés microbiennes des DMA. Différentes interactions ont déjà été décrites parmi lesquelles les interactions suivantes.

5.1. Compétition

Il existe une compétition entre les microorganismes acidophiles pour les substrats, c'est à dire leurs sources d'énergie incluant les donneurs d'électrons inorganiques et organiques. Différents paramètres environnementaux tels que la température, le pH, les concentrations en substrats et en métaux dissous, ont un impact important sur ces phénomènes de compétition et sur l'organisme le mieux adapté quelque soit la situation. Par exemple, en raison de sa plus grande affinité pour le fer ferreux et sa meilleure tolérance aux faibles pH (< 1,8) et au fer ferrique, *Leptospirillum ferrooxidans* présente une plus grande efficacité dans la lixiviation de minerai riche en pyrite qu'*Acidithiobacillus ferrooxidans* (Johnson, 1998). De même, la thermotolérance de *L. ferrooxidans* lui donne un avantage compétitif à des températures assez élevées (35-45 °C) alors que *A. ferrooxidans* est plus efficace à des températures plus faibles (25 °C). D'autre part, le taux de croissance plus rapide de *A. ferrooxidans*, permet

généralement à cet organisme de dominer dans des écosystèmes où le fer ferreux est en concentration élevée et/ou le pH est supérieur à 2.

5.2. Prédation

La prédation de bactéries hétérotrophes et chimiolithotrophes par des protozoaires acidophiles obligatoires a été observée en conditions de laboratoire. Un flagellé proche des *Eutreptia* se nourrit préférentiellement d'*A. ferrooxidans* plutôt que de *L. ferrooxidans* en culture contenant les deux chimiolithotrophes (Mc Guinness & Johnson, 1992; Johnson & Rang, 1993). De même, le nombre de ces acidophiles présents dans une usine qui éliminait le soufre du charbon a considérablement diminué dans un laps de temps très court du fait de la présence d'un flagellé acidophile (Johnson & Rang, 1993). Ces eucaryotes peuvent ainsi réguler l'activité biologique et la biomasse des communautés bactériennes.

5.3. Mutualisme

Les interactions entre les microorganismes acidophiles peuvent se traduire par le gain mutuel d'un bénéfice pour chacun des partenaires, tel qu'illustré par exemple par les relations de feedback qui interviennent entre les chimiolithotrophes oxydant le fer ferreux (utilisant le fer comme donneur d'électron) et les hétérotrophes réduisant le fer ferrique (utilisant le fer comme accepteur d'électron). Un autre exemple concerne *L. ferrooxidans* et *A. ferrooxidans* qui sont tous deux sensibles aux acides et composés organiques relargués dans le milieu par leur propre métabolisme induisant une inhibition de leur croissance. Les bactéries hétérotrophes sont capables de lever cette inhibition en métabolisant ces composés. Ce phénomène explique pourquoi les co-cultures d'organismes oxydant le fer et d'hétérotrophes entraînent une dissolution plus efficace des minerais que les cultures pures (Johnson, 1998).

5.4. Synergisme

L'association de deux ou plusieurs espèces de microorganismes acidophiles rend leur activité complémentaire plus efficace que celle obtenue par un organisme seul. C'est le cas par exemple, des co-cultures de *L. ferrooxidans* et des bactéries oxydants le soufre comme *A. thiooxidans* ou *A. caldus* qui permettent une dissolution plus efficace de la chalcopryrite par rapport aux cultures pures (Norris, 1990). De même, les cultures mixtes de *Ferromicrobium*

acidophilus et *A. thiooxidans* sont capables d'oxyder la pyrite contrairement aux cultures pures (Johnson, 1998).

6. Cycles biogéochimiques catalysés par les microorganismes

Les groupes métaboliques primaires détectés dans les DMA sont les lithoautotrophes qui oxydent le Fe(II) et le S libérés lors de la dissolution de la pyrite, les organohétérotrophes qui utilisent le carbone produit par les lithoautotrophes et les microorganismes anaérobies qui couplent l'oxydation du soufre ou du carbone organique à la réduction du Fe(III) (Baker & Banfield, 2003). Le sulfate peut également constituer un important accepteur d'électron pour le développement de ces microorganismes anaérobies. En plus de ces métabolismes impliquant le fer et le soufre souvent présents majoritairement dans ces écosystèmes, il peut exister d'autres voies métaboliques impliquant des éléments tels que l'arsenic.

6.1. Les microorganismes utilisant le fer

6.1.1. Oxydation du fer

De part les fortes concentrations en fer retrouvées dans ces systèmes, les bactéries capables d'oxyder le fer sont généralement favorisées. Ainsi, on retrouve dans les DMA, un certain nombre de bactéries qui tirent leur énergie de l'oxydation du Fe(II) en Fe(III). La première bactérie capable d'oxyder le fer ayant été isolée dans les DMA est *Acidithiobacillus ferrooxidans*. Cette bactérie a été largement étudiée : c'est une bactérie à Gram négatif, acidophile, chimiolithoautotrophe stricte, anaérobie facultative qui est également capable d'oxyder le soufre (Duquesne *et al.*, 2003). Beaucoup d'autres bactéries peuvent également oxyder le fer, comme *Leptospirillum ferrooxidans*, *L. ferriphilum*, *Acidithiobacillus ferrivorans*, *Acidimicrobium ferrooxidans*, *Thiomonas* spp., *Ferrovum myxofaciens* ou *Gallionella ferruginea* pour ne citer que les plus connues. De même quelques Archaea présentent également ce métabolisme, comme *Ferroplasma acidiphilum* un autotrophe obligatoire capable d'oxyder le Fe(II) comme seule source d'énergie (Golyshina *et al.*, 2000) ou bien *F. acidarmanus* qui peut se développer hétérotrophiquement (Edwards *et al.*, 2000b). Ces différentes espèces sont retrouvées au niveau de sites miniers divers où elles ont un rôle central dans la dissolution oxydative des minéraux contenant du fer (pyrite, jarosite...).

6.1.2. Réduction du fer

La respiration anaérobie basée sur la réduction du fer ferrique est une stratégie métabolique très répandue chez les bactéries acidophiles en conditions anoxiques (Johnson & Bridge, 2002). Les espèces thermophiles modérées comme *Sulfobacillus thermosulfidooxidans*, *Sulfobacillus acidophilus* ou *Acidimicrobium ferrooxidans* peuvent utiliser le fer ferrique comme seul accepteur d'électrons quand elles se développent de façon hétérotrophe en conditions limitées en oxygène et en présence d'extrait de levure et de glycérol (Baker & Banfield, 2003). *Acidiphilium acidophilum* (anciennement *Thiobacillus acidophilus*) est capable de réduction mixotrophe du Fe(III) et *A. cryptum* peut coupler la réduction du Fe(III) à l'oxydation de diverses substances incluant le H₂ et le glucose.

D'autres bactéries, comme *Sulfobacillus acidophilus*, sont capables en conditions anaérobies d'utiliser le Fe(III) même s'il n'est pas en solution. Ceci permet d'accélérer la dissolution réductrice de minéraux contenant du fer, tels que la jarosite et la goethite (Bridge & Johnson, 1998). *Acidithiobacillus ferrooxidans* peut également se développer en conditions anoxiques en utilisant le fer ferrique comme accepteur d'électrons et le H₂ comme donneur d'électrons (Pronk *et al.*, 1992). Cependant, les conditions anoxiques ne sont pas indispensables pour la réduction du Fe(III) par des mélanges de populations acidophiles. Küsel *et al.* (2002) ont montré que sous des conditions oxiques, la souche JF-5 d'*A. cryptum* réduit le Fe(III) soluble et la schwertmannite dans le sédiment à pH 3.

6.2. Les microorganismes utilisant le soufre

6.2.1. Oxydation du soufre

La plupart des microorganismes des DMA qui oxydent le soufre sont également capables de fixer le CO₂. En effet, l'oxydation autotrophe du soufre a été décrite chez des cultures pures d'*Acidithiobacillus ferrooxidans*, d'*A. thiooxidans*, de *Sulfobacillus disulfidooxidans*, de *Thiobacillus albertis* et de *Sulfobacillus acidophilus* (Baker & Banfield, 2003). On retrouve des bactéries sulfo-oxydantes strictes telles qu'*Acidithiobacillus thiooxidans* ou *Thiomonas cuprina*, mais il existe également plusieurs oxydants du soufre qui peuvent oxyder le carbone organique et/ou le fer ferreux. C'est le cas notamment d'*Acidithiobacillus ferrooxidans* et de certaines espèces de *Sulfobacillus* (*S. acidophilus* et *S. thermosulfidooxidans*) qui sont capables d'oxyder à la fois le fer et le soufre et de réduire le fer (Baker & Banfield, 2003;

Johnson & Hallberg, 2003). Certaines souches de *S. acidophilus* peuvent également coupler l'oxydation autotrophe du tétrathionate à la réduction de Fe(III) (Norris *et al.*, 1986; Bridge & Johnson, 1998). On trouve également quelques Archaea capables d'oxyder le soufre telle que *Metallosphaera prunae* (Fuchs *et al.*, 1995).

6.2.2. Réduction du soufre

Le sulfate est un autre soluté important dans les DMA et peut servir d'accepteur d'électron pour le développement anaérobie des microorganismes. Cependant, peu d'informations sont disponibles concernant la réduction du sulfate dans les DMA. Ceci est probablement dû au fait que les caractéristiques physicochimiques propres aux DMA (acidité, faible concentration en carbone organique...) n'étaient pas considérées comme propices au développement des BSR. En effet les conditions optimales pour les BSR sont représentées par des milieux anaérobies avec un pH proche de la neutralité et la présence de donneurs d'électron organiques (Widdel, 1988). L'oxygène est par ailleurs connu pour inhiber la réduction microbienne des sulfates, probablement à travers l'inactivation ou l'oxydation de plusieurs enzymes et protéines utilisées dans les processus de réduction. Pourtant, des BSR ont été isolées de drainages acides à pH inférieur à 3 (García *et al.*, 2001) et il a été montré que certaines souches sont capables de survivre plusieurs heures en présence d'oxygène (Dilling & Cypionka, 1990). De plus, la réduction aérobie des sulfates a été montrée dans des zones photosynthétiques très oxygénées (Canfield & Des Marais, 1991). Au sein des DMA, on retrouve des espèces telles que *Desulfosarcina variabilis*, *Desulfosporosinus orientis*, des *Desulfovibrio* et des *Citrobacter* (Johnson & Hallberg, 2003; Qiu *et al.*, 2009). La réduction des sulfates en sulfures entraîne la précipitation de certains métaux et métalloïdes conférant aux BSR un rôle central dans la bioremédiation.

6.3. Les microorganismes utilisant l'arsenic

6.3.1. Oxydation de l'arsenic

L'oxydation microbiologique de l'As(III) en As(V) est un processus connu depuis 1918 (Green, 1918) qui peut avoir un impact sur la mobilité et la toxicité de l'arsenic dans l'environnement, limitant par exemple sa biodisponibilité comme nous l'avons vu dans le paragraphe B.3.5. C'est pourquoi l'oxydation de l'arsénite est étudiée comme un mécanisme pouvant être utilisé pour la biorémediation de systèmes contaminés.

Plus de 30 souches représentant au moins 9 genres sont capables d'oxyder l'As incluant des α -, β - et des γ -protéobactéries, des *Deinocci* (*Thermus*) et des *Crenarchaeota* (Stolz *et al.*, 2006). Physiologiquement divers, ces microorganismes incluent à la fois des bactéries hétérotrophes oxydant l'arsénite (Heterotrophic Arsenite Oxidizers, HOAs) et des chimiolithoautotrophes (Chemoautotrophic Arsenite-Oxidizers, CAOs). Les CAOs couplent l'oxydation de l'arsénite à la réduction de l'oxygène ou des nitrates et utilisent l'énergie dérivée pour fixer le CO₂. L'oxydation hétérotrophe de l'As(III) est perçue comme une réaction de détoxification qui convertit l'As(III) à l'extérieur de la membrane de la cellule en sa forme moins toxique As(V), qui entre probablement moins facilement dans la cellule. Pour les HOAs, l'oxydation de l'As(III) est catalysée par une enzyme périplasmique dont la description a été faite par différents auteurs (Ellis *et al.*, 2001; Stolz & Basu, 2002). Par contre, la structure de l'arsénite oxydase des CAOs n'est pas encore bien établie (Oremland & Stolz, 2005). Quelques procaryotes bien connus sont capables d'oxyder l'arsenic en condition aérobie (*Herminiimonas arsenicoxydans*, *Thiomonas* spp., or *Rhizobium* sp. strain NT26) ou anaérobie (*Alkalilimnicola ehrlichii*) dans le cadre d'un mécanisme de détoxification (Heinrich-Salmeron *et al.*, 2011). Certains chimiolithoautotrophes utilisent également l'arsénite comme donneur d'électron (*Rhizobium* sp. strain NT26 ou *Thiomonas arsenivorans*) (Santini *et al.*, 2000; Battaglia-Brunet *et al.*, 2006). Concernant les Archaea, l'oxydation de l'arsenic a été rapportée chez *Sulfolobus acidocaldarius* (Sehlin & Lindström, 1992) et des homologues de l'arsénite oxydase ont été identifiés dans les génomes de *Sulfolobus tokodaii* et *Aeropyrum pernix* (Lebrun *et al.*, 2003). Chez les eucaryotes, seule l'algue rouge *Cyanidioschyzon* sp. semble être capable d'oxyder l'As(III) mais rien n'indique qu'il s'agisse d'une réaction catalysée enzymatiquement (Qin *et al.*, 2009).

6.3.2. Réduction de l'arsenic

Bien que l'arsenic soit généralement toxique pour les organismes vivants, il a été montré que certains microorganismes sont capables d'utiliser des composés d'arsenic comme accepteur d'électron ou comme mécanismes de détoxification (Anderson & Cook, 2004). Il existe deux types d'organismes réduisant l'arsenic, les DARPs (Dissimilatory Arsenate-Respiring Prokaryotes) et les ARMs (Arsenate-Resistant Microbes) (Oremland & Stolz, 2005).

Les DARPs utilisent l'As(V) comme accepteur d'électron pour leur respiration anaérobie en couplant sa réduction à l'oxydation de substrats organiques ou inorganiques, conduisant à la

formation d'As(III). Ce processus est réalisé *via* l'opéron *ars* (Silver & Phung, 2005). Jusqu'à présent aucun DARPs obligatoire n'a été trouvé car l'ensemble des souches examinées peuvent également réduire d'autres composés, comme par exemple *Sulfurospirillum barnesii* capable de réduire le séléniate, le nitrate, le nitrite, le fumarate, le Fe(III), le thiosulfate, le soufre élémentaire, le dimethylsulfoxyde et l'oxyde de triméthylamine (Oremland & Stolz, 2003). Cette diversité métabolique peut être un facteur écologique important, parce que les espèces chimiques du soufre, du fer et des nitrates (S(-II), FeOOH, NO₃⁻) interagissent avec l'arsenic dans l'environnement.

Les ARMs utilisent la réduction de l'arséniate comme mécanisme de résistance et ne gagnent pas d'énergie par ce processus. L'arséniate qui entre dans le cytoplasme via les pompes à phosphate est convertit en As(III) par un processus contrôlé par une enzyme arséniate réductase cytoplasmique (ArsC). L'As(III) est ensuite expulsé hors de la cellule grâce à un transporteur spécifique ArsB.

Il existe au moins 24 espèces isolées capables de réaliser la réduction dissimilatrice ou respiration de l'As(V). Elles appartiennent à divers groupes phylogénétiques incluant des membres des *Crenarchaeota*, *Aquificae*, *Chrysiogenes*, *Deferribacteres*, *Halanaerobacter*, ainsi que des γ -, δ - et ϵ -protéobactéries (Stolz *et al.*, 2006). On trouve entre autre des espèces bactériennes telles que *Sulfurospirillum barnesii*, *S. arsenophilum*, *Bacillus arsenicoselenatis*, *B. selenitireducens*, *Desulfotomaculum auripigmentum*, *Chrysiogenes arsenatis* et *Desulfomicrobium* strain Ben-RB (Stolz & Oremland, 1999; Macy *et al.*, 2000) ainsi que deux Archaea hyperthermophiles, *Pyrobaculum arsenaticum* et *P. aerophilum* (Huber *et al.*, 2000). Concernant le mécanisme de résistance, des homologues de l'opéron *ars* ont été trouvés au sein des genres *Pseudomonas*, *Bacillus*, *Klebsiella*, *Enterobacter*, *Citrobacter*, *Staphylococcus*, *Salmonella*, *Thiobacillus*, *Yersinia* et *Escherichia* (Macur *et al.*, 2001). Au sein des environnements miniers riches en arsenic, diverses bactéries principalement du type ARMs ont été détectées. C'est le cas des genres *Bacillus*, *Enterobacter* et *Ochrobactrum* dans une mine d'étain (Jareonmit *et al.*, 2012), des genres *Caulobacter*, *Sphingomonas* et *Rhizobium* dans une fonderie de cuivre (Macur *et al.*, 2001) ou des genres *Stenotrophomonas* et *Serratia* dans une mine d'antimoine (Botes *et al.*, 2007). Ces résultats montrent qu'une grande diversité phylogénétique de microorganismes réduisant l'arsenic peut être trouvée

dans ces environnements extrêmes et qu'ils sont ainsi susceptibles de jouer un rôle majeur dans le cycle biogéochimique de l'arsenic.

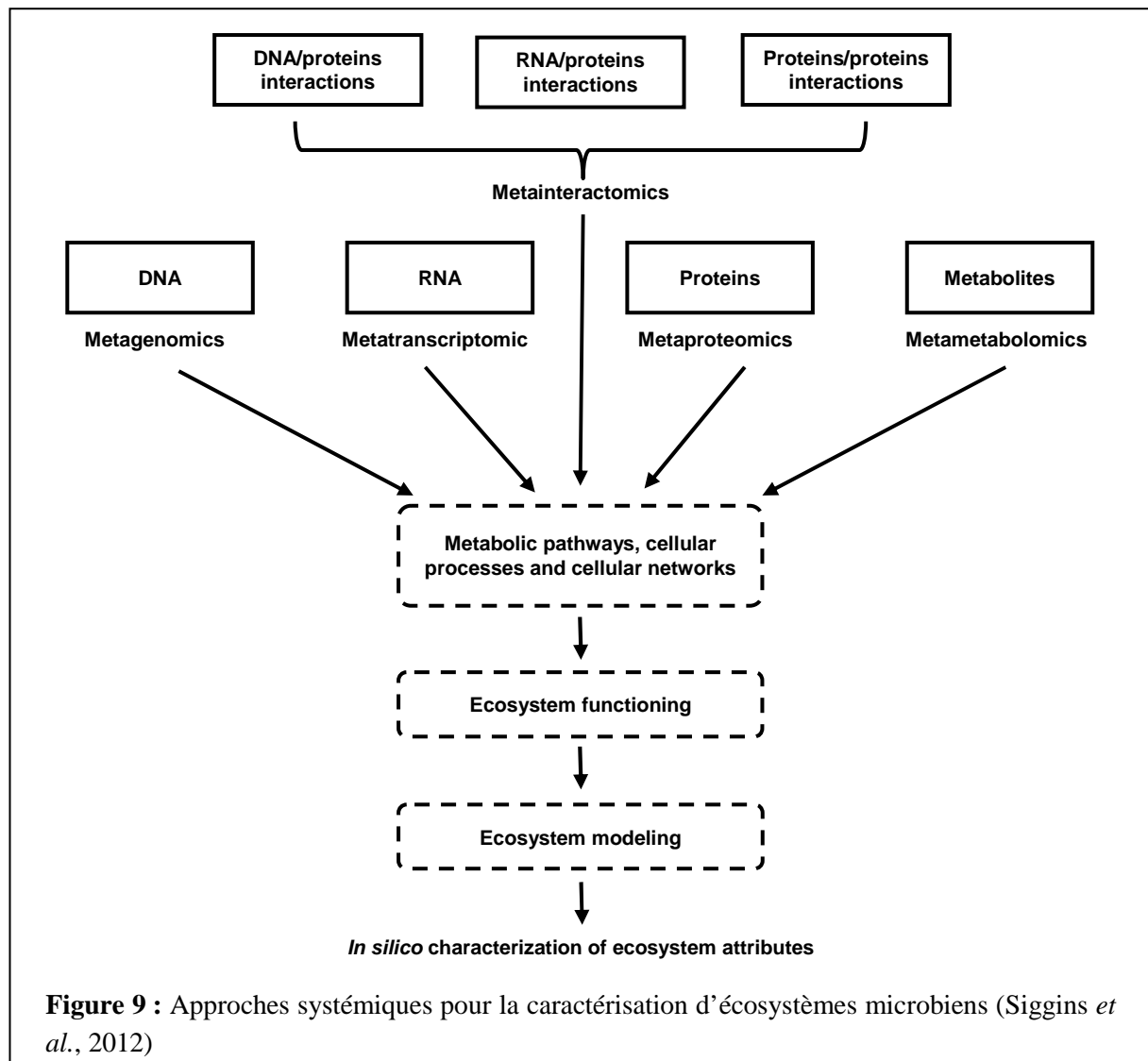
Les microorganismes présents dans les DMA jouent un rôle central dans ces écosystèmes de part leur implication dans les processus à l'origine de leur génération mais également de leur remédiation. Une meilleure caractérisation de la diversité génétique et fonctionnelle de ces communautés microbiennes ainsi que leurs interactions entre elles et avec leur environnement est une étape indispensable à la compréhension du fonctionnement de ces écosystèmes pour pouvoir développer à terme des stratégies de remédiation. Il est donc nécessaire de développer des méthodes d'analyses permettant d'appréhender l'ensemble des communautés microbiennes présentes dans ces écosystèmes.

C. INTÉRÊTS DES APPROCHES –OMICS

Comme nous l'avons vu précédemment, les microorganismes comprenant les bactéries, Archaea et eucaryotes sont omniprésents et jouent un rôle essentiel dans les cycles biogéochimiques indispensables au fonctionnement des écosystèmes comme les DMA. Dans leur habitat naturel, ces microorganismes coexistent en communautés mixtes dont la complexité est spécifique à chaque environnement, avec des estimations allant de 6 taxons présents dans un biofilm issu d'un DMA à 10^6 taxons par gramme de sol (Siggins *et al.*, 2012). Cependant, comme la majorité des microorganismes présents dans l'environnement (dont 99 % des procaryotes), ne sont pas cultivables à ce jour, leur investigation nécessite l'utilisation de techniques moléculaires qui contournent les méthodes traditionnelles de culture et d'isolement d'espèces individuelles (Amann *et al.*, 1995). De plus, afin de pouvoir étudier une communauté microbienne dans son ensemble et ses réponses aux facteurs environnementaux, il s'est avéré indispensable de prendre en compte la complexité des consortiums microbiens présents ainsi que leurs interactions.

L'ensemble de ces constatations s'est traduit par l'apparition d'un nouveau domaine de recherche en écologie microbienne qui englobe les approches systémiques (Figure 9) par lesquelles tous les niveaux de l'information biologique sont étudiés (ADN, ARN, protéines et métabolites). Les avancées technologiques récentes ont rendu possible de telles approches par le développement de méthodes « -omics » à haut débit, où les communautés microbiennes mixtes sont vues comme un seul méta-organisme. La métagénomique, la métatranscriptomique, la métaprotéomique et la méta-métabolomique sont employées pour déterminer respectivement, les séquences d'ADN, les ARN transcrit collectivement, les protéines traduites et les métabolites résultant de processus cellulaires et ceci pour l'ensemble des microorganismes présents dans l'écosystème étudié. L'ensemble des données générées permet ainsi d'identifier les voies métaboliques et les processus cellulaires mis en jeu dans cet écosystème. Un autre niveau d'information est également utilisé, la métainteractomique, qui permet d'explorer le réseau cellulaire de l'écosystème par analyse des interactions moléculaires entre l'ensemble des microorganismes présents. Enfin, l'intégration des données résultant de ces différentes approches systémiques permet d'obtenir un aperçu du fonctionnement de l'écosystème et est également utilisée pour développer des modèles

pouvant servir à prédire le comportement d'un écosystème en réponse à des changements environnementaux (Raes & Bork, 2008; Røling *et al.*, 2010).



L'utilisation de ces différentes techniques en complément des outils bioinformatiques permet donc d'obtenir une image globale de la réponse biologique d'un organisme ou d'une communauté microbienne dans un écosystème. Cependant, chacune de ces méthodes présente des limites qui lui sont propres et qu'il est nécessaire de prendre en compte lors de l'analyse des résultats afin de limiter les erreurs et abus d'interprétation. De plus, bien qu'une quantité énorme d'informations soit générée, il reste indispensable de s'appuyer également sur les méthodes plus traditionnelles afin de pouvoir compléter les résultats obtenus. L'ensemble des atouts et limites de ces différentes techniques permettant l'analyse des communautés microbiennes d'un écosystème est présenté dans les paragraphes suivants.

1. La métagénomique

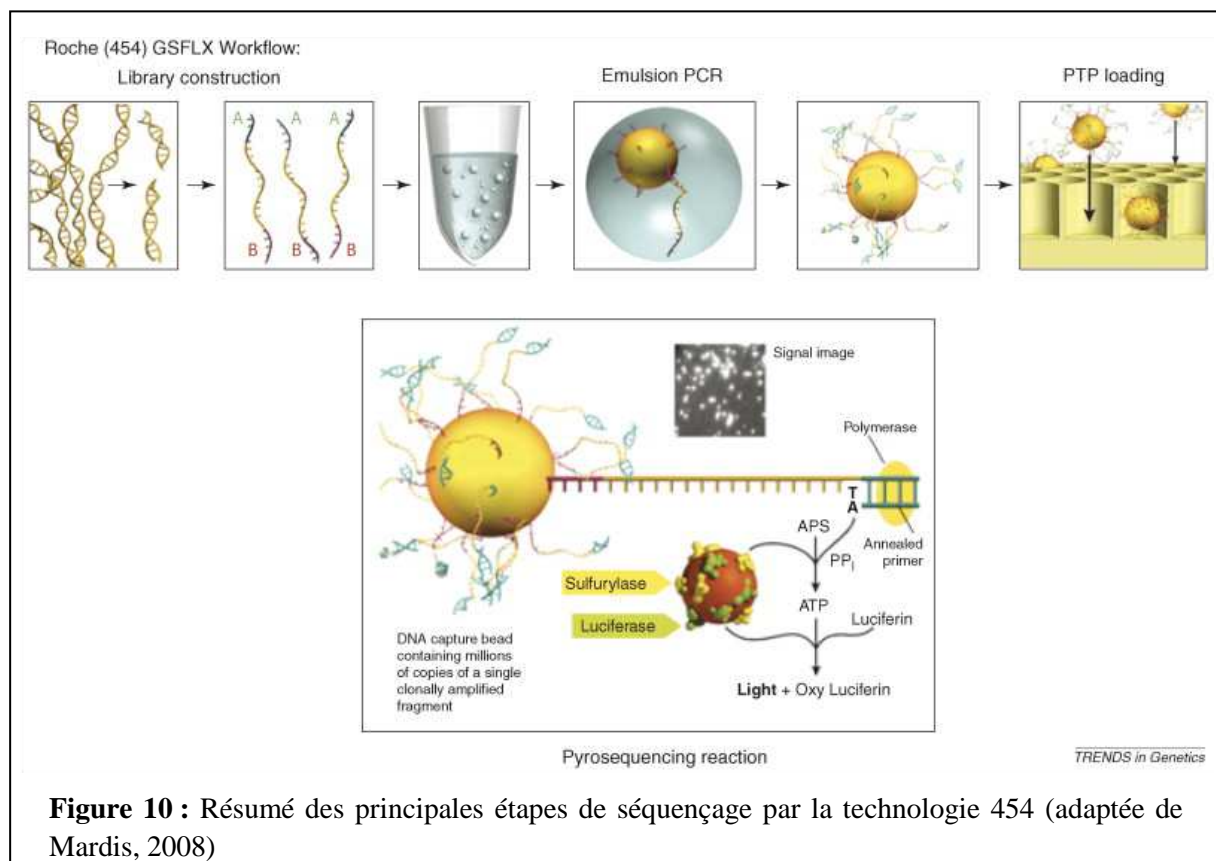
La métagénomique ou encore génomique environnementale ou génomique des communautés, repose sur l'étude des métagénomomes, c'est-à-dire l'ensemble des séquences d'ADN extraites de communautés multi-espèces prélevées dans l'environnement (Bastida *et al.*, 2009). Les premières analyses métagénomiques se sont focalisées sur la diversité des nouvelles espèces et la communauté formée par celles-ci (Gill *et al.*, 2006). La métagénomique descriptive fournit une vue relativement non biaisée non seulement de la structure des communautés (phylogénie, abondance, richesse et distribution des espèces), mais également de ses fonctions métaboliques potentielles. Par la suite, la métagénomique est devenue « intégrative » en cherchant à identifier un changement fonctionnel microbien en fonction d'une modification de l'environnement. Ainsi, en comparant et réassemblant les séquences obtenues par rapport à des génomes de référence, il est possible (i) de caractériser la structure phylogénétique des communautés microbiennes (gènes ribosomiaux), (ii) d'identifier les gènes fonctionnels par homologie aux génomes de référence, (iii) d'identifier de nouvelles fonctions et (iv) de reconstruire des génomes d'organismes non cultivables.

Cette approche est donc indissociable des méthodes de séquençage qui ont récemment connu une évolution particulièrement rapide depuis l'apparition de la technologie dite de première génération, la technologie Sanger. Cette dernière, développée au milieu des années 1970, a subi diverses améliorations incluant notamment des procédures d'automatisation permettant une réduction des coûts et une augmentation de la taille des séquences générées. Cependant cette technique repose sur des étapes préliminaires de clonage et transformation bactérienne qui sont relativement pénalisantes en termes de coût et de temps. Pour palier à ces défauts, de nouvelles techniques de séquençage à haut débit ont été mises au point ces dernières décennies et sont qualifiées de technologie de seconde génération (ou next-generation sequencing, NGS). Le principal avantage de ces méthodes à haut débit se traduit par le fait qu'elles sont hautement parallélisées, ce qui permet d'obtenir des millions de séquences différentes en une seule expérience sans clonage de fragments génomiques dans un vecteur. La longueur des fragments séquencés, aussi appelés « reads », est par contre limitée comparativement aux méthodes de séquençage classique qui peuvent séquencer jusqu'à 1000 nucléotides contigus et rend donc difficile l'obtention d'une couverture et d'un assemblage suffisant pour reconstituer indépendamment les génomes de différents microorganismes présents dans un échantillon. La taille des reads augmente néanmoins avec l'affinement des

méthodes de séquençage à haut débit, mais elle demeure encore une limitation des nouvelles méthodes. De plus, l'assemblage de ces données reste l'élément limitant, puisqu'il nécessite une capacité informatique beaucoup plus grande. Il existe actuellement trois technologies majeures de séquençage à haut débit. Les particularités de chacune d'elles sont décrites ci-dessous.

1.1. La technologie 454

La technologie 454, première technologie alternative au séquençage Sanger, a été mise sur le marché par l'entreprise 454 Life Sciences en 2005, une filiale de Roche Diagnostics. Cette technique permet un gain de temps et d'argent car elle ne nécessite pas de clonage et permet une lecture directe de la séquence obtenue après le séquençage. Actuellement cette technique permet de générer environ 500 Mb (Mega base pair) par run (*i.e.* par réaction de séquençage) avec des fragments de taille moyenne de 450 pb. Cette technique repose sur la succession de 4 étapes présentées dans la figure 10.



La première étape consiste en la préparation d'une banque d'ADN simple brin (ADNsb). Pour cela l'ADN génomique est fragmenté par nébulisation puis deux adaptateurs (A et B) sont

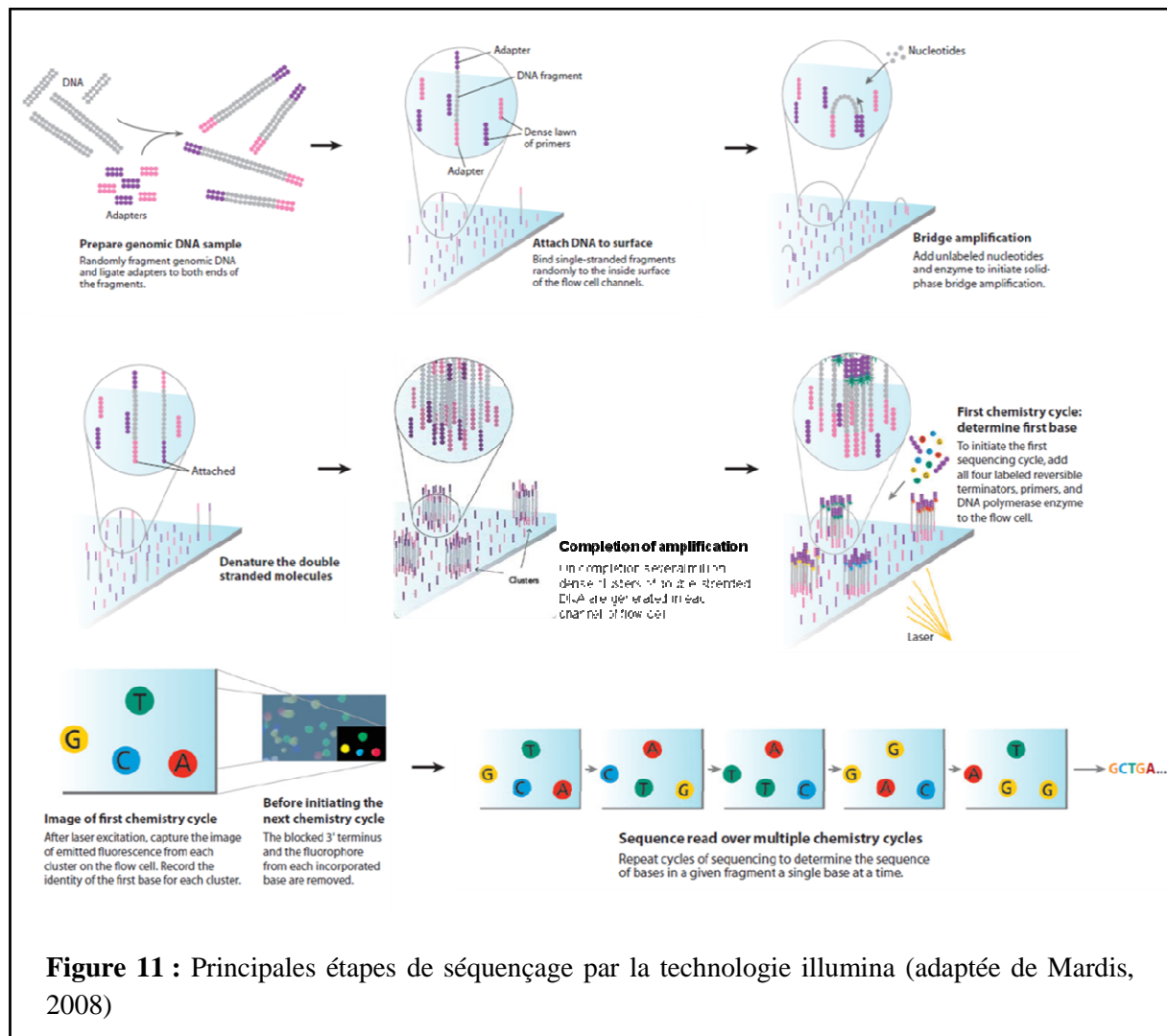
fixés par ligation aux deux extrémités. Ensuite les fragments d'ADNsb sont mis en contact avec des billes couvertes de sondes oligonucléotidiques complémentaires aux adaptateurs ne permettant de fixer qu'une seule molécule d'ADN à la fois. L'étape suivante est la PCR en émulsion (emPCR), qui consiste à ségréger les billes dans des bulles qui servent de microréacteurs. Ainsi, chaque bille sera couverte par une amplification clonale d'un seul fragment à séquencer. Après l'emPCR, des millions de séquences identiques recouvrent ainsi chaque bille. Lors de la troisième étape, les microréacteurs sont dissociés et les billes porteuses de l'ADNsb amplifié sont isolées individuellement dans les micro-puits d'une plaque de picotitration, permettant ainsi de séparer spatialement le signal de pyroséquençage de chaque fragment. La dernière étape correspond donc à la réaction de pyroséquençage. Sur chaque bille, la séquence complémentaire du fragment amplifié est synthétisée à partir de nucléotides (dATP, dGTP, dCTP et dTTP) ajoutés séquentiellement dans chaque puits dans un ordre prédéterminé. Chaque nucléotide incorporé lors de la synthèse du brin complémentaire libère un pyrophosphate, qui est ensuite dégradé par une ATP sulfurylase et une luciférase. Cette réaction enzymatique produit de la lumière qui est détectée par une caméra CCD (en anglais, Charge Coupled Device). Entre chaque ajout de nucléotide, le puits est lavé pour supprimer les nucléotides non incorporés avant injection du suivant. La détection du signal lumineux peut ensuite être associée à l'ordre d'introduction des nucléotides, ce qui permet de déterminer la séquence de chaque fragment immobilisé sur une bille.

La plus grande limite de la technologie 454 concerne la présence d'homopolymères, c'est-à-dire la succession de plusieurs nucléotides identiques, résultant de l'absence de nucléotides terminaux servant à empêcher plusieurs incorporations consécutives pour un cycle donné. Dans ce cas un signal lumineux proportionnel au nombre de nucléotides incorporé est décelé. Cependant, au-delà de 4 à 5 nucléotides, le caractère proportionnel du signal est perdu. Ces séquences homopolymériques représentent la source majeure d'erreurs de séquençage par cette méthode, comme l'ont estimé Huse *et al.* (2007), qui attribuent 39 % des erreurs de type insertions/délétions à ce contexte nucléotidique spécifique. Une autre source d'erreur possible est la présence résiduelle de nucléotides injectés au cours de cycles précédents et qui est interprétée comme étant le nucléotide courant. Cela entraîne une erreur appelée « carry-forward ». Ces erreurs limitent l'utilité de cette méthode pour le séquençage de nouveaux génomes et la détection de mutations. L'avantage clé de cette technologie est la longueur des

fragments générés qui est plus importante que celles des autres technologies et qui peut être nécessaire pour certaines applications (*e.g.* assemblage *de novo* et métagénomique).

1.2. La technologie Illumina

La société Illumina (anciennement Solexa) a commercialisé la seconde technique de séquençage haut débit en 2006. Cette technique permet le séquençage massivement parallèle de petits fragments. Cette méthode domine actuellement le marché et elle peut produire de 15 à 200 Gb (Giga base pair) par run, ce qui augmente la couverture de séquençage d'un génome et améliore ainsi la précision de la séquence finale. Les différentes étapes de cette technique sont présentées dans la figure 11.



La première étape consiste en la fragmentation de l'ADN à séquençer par nébulisation puis la ligation d'adaptateurs à chacune des extrémités. Ces molécules d'ADN sont ensuite dénaturées

sous forme simple brin puis immobilisées sur une lame de verre à l'aide d'amorces recouvrant la lame et s'hybridant avec un des adaptateurs. L'amplification des fragments immobilisés se fait alors par PCR par pont ou « bridge PCR ». Ensuite, l'extrémité opposée du brin nouvellement synthétisé est hybridée à l'amorce anti-sens, qui est aussi attachée à la surface de la lame de verre. Des cycles d'extension des amorces par une polymérase suivis d'une dénaturation au formamide permettent de créer sur la surface des clusters d'environ 1000 amplicons représentant la même séquence. Dans un second temps, une amorce de séquençage est hybridée aux fragments immobilisés et permet l'incorporation de nucléotides contenant un terminateur de chaîne réversible marqué par un fluorochrome différent selon le nucléotide. Les fluorochromes sont ensuite excités grâce à des lasers permettant l'émission d'un signal lumineux spécifique à chaque nucléotide. La fluorescence est alors détectée sur toute la lame afin d'associer les nucléotides incorporés aux différentes positions spatiales, ce qui permet de construire la séquence de chaque fragment. Les terminateurs de chaîne sont ensuite clivés afin de permettre un nouveau cycle d'ajout de nucléotides. La succession de ces étapes permet d'accéder à la séquence complète du fragment d'ADN de chaque cluster dans la limite des 100 premiers nucléotides.

Bien que cette technique de séquençage soit à l'heure actuelle l'une des plus utilisées, elle présente plusieurs biais notables. Le principal problème de cette technique est la taille relativement faible des lectures générées, entre 35 et 100 pb, comparée à celles générées par le pyroséquençage, jusqu'à 450 pb et de la méthode Sanger, jusqu'à 1 kb. L'autre biais de cette technique concerne la diminution de la fiabilité du séquençage aux extrémités des lectures. Dohm *et al.* (2008), ont émis l'hypothèse que l'accumulation d'erreurs à l'extrémité 3' des lectures était due au phénomène de déphasage, qui correspond à l'extension incomplète ou l'addition de multiples nucléotides au fragment séquencé. Ainsi, plus le nombre de cycles augmente, plus les décalages dans la séquence s'accumulent conduisant ainsi à une augmentation du bruit de fluorescence et une interprétation erronée des signaux lumineux (Erlich *et al.*, 2008). De plus, il semble que les séquences riches en AT et en GC soient sous-représentées avec cette technologie (Metzker, 2010).

1.3. Les technologies SOLiD et de troisième génération

Applied Biosystems a commercialisé en 2007 la technologie SOLiD (Sequencing by Oligonucleotide Ligation and Detection) basée sur le principe de séquençage par ligation

permettant de générer de 90 à 300 Gb par run avec des fragments d'une taille moyenne de 75 pb.

Les premières étapes de séquençage de cette technique sont similaires à celles de la technologie 454, utilisant également une PCR en émulsion sur des billes paramagnétiques. C'est après cette étape que le protocole diffère. Dans le SOLiD, les billes sont triées afin de ne conserver que celles couvertes d'amplicons. Elles sont ensuite immobilisées de façon désordonnée sur une lame de verre qui, contrairement à celle de la technique 454, n'est pas poreuse. Le séquençage proprement dit est différent car il se base sur l'utilisation d'une ADN ligase et non pas une polymérase (Figure 12a). Premièrement, une amorce universelle est hybridée à l'adaptateur qui a été ajouté à chaque fragment pendant la création de la librairie à séquencer. Ensuite, des octamères dégénérés sur toutes les bases, sauf une (la position « n »), sont hybridés et liés à l'amorce universelle par la ligase. Chacun de ces octamères est marqué par un des quatre fluorochromes dont la couleur est associée au nucléotide à la position « n ». Après la ligation il y a alors détection de la fluorescence permettant de déterminer les bases 5 et 6. L'octamère est ensuite clivé après la position « n » afin d'éliminer le fluorochrome. Plusieurs cycles similaires sont effectués, puis le fragment produit par ligation des octamères est dénaturé. Une nouvelle amorce universelle est hybridée aux fragments, suivie d'une série de cycles mettant en jeu des octamères dont une position n' (différente de n) est associée à la couleur du fluorochrome. L'utilisation d'octamères dont deux bases sont connues au lieu d'une permet de diminuer les erreurs de séquençage (Shendure & Ji, 2008) (Figure 12b). Cette méthode présente donc une meilleure spécificité et une plus grande précision que le séquençage par polymérisation, c'est pourquoi elle est principalement utilisée dans le cadre de détection de mutations nucléotidiques ponctuelles.

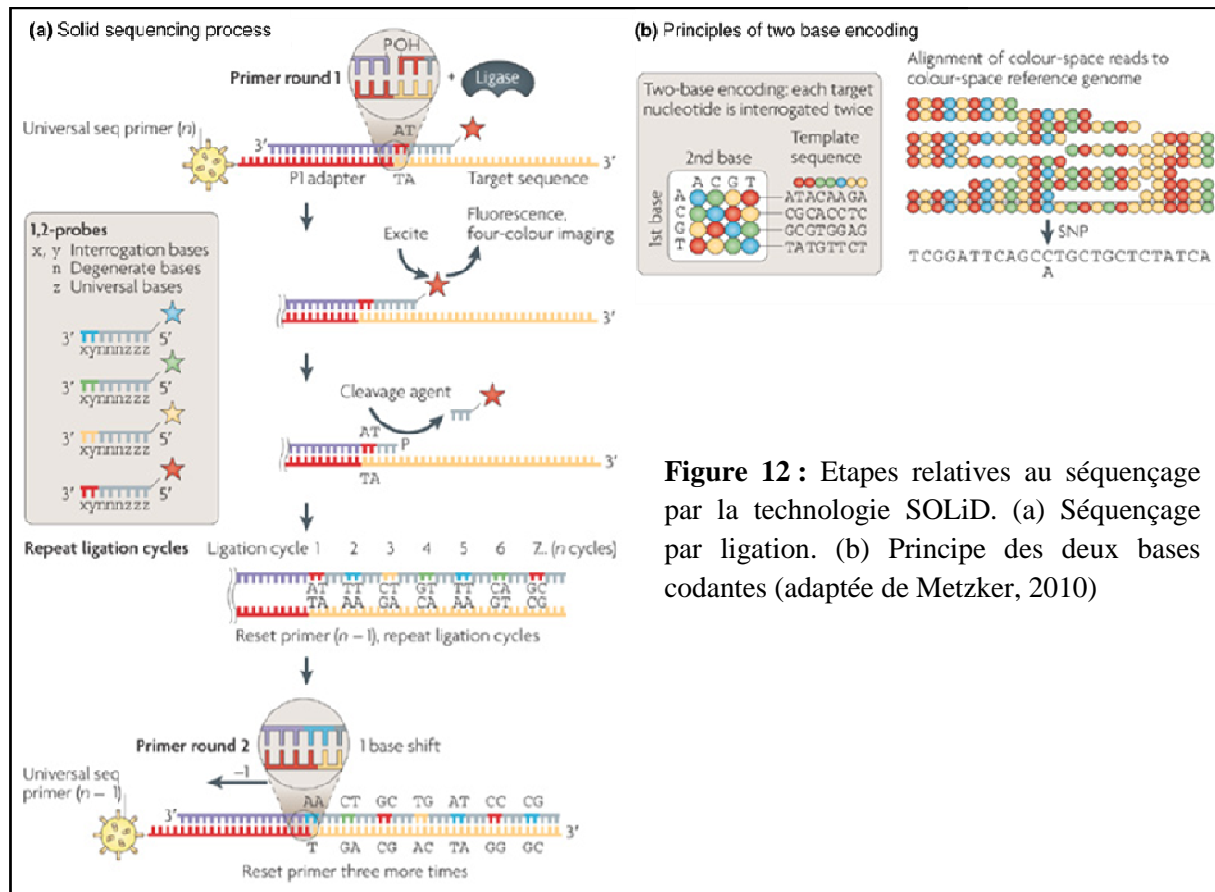


Figure 12 : Etapes relatives au séquençage par la technologie SOLiD. (a) Séquençage par ligation. (b) Principe des deux bases codantes (adaptée de Metzker, 2010)

D'autres techniques de séquençage dites de troisième génération sont en cours de développement notamment par les sociétés Helicos et Pacific Biosciences (Korlach *et al.*, 2010). Le principe de la 3^{ème} génération peut être symbolisé par le séquençage direct en temps réel d'une molécule d'ADN sans étape de pré-amplification (contrairement à la génération actuelle). Ces technologies de séquençage d'ADN simple molécule en temps réel ou SMS (Single Molecule Sequencing) peuvent être regroupées selon trois catégories : (i) technologies de séquençage en temps réel impliquant la synthèse du brin d'ADN complémentaire *via* une ADN polymérase, (ii) technologies de séquençage par détection des bases successives d'une molécule d'ADN au travers de nanopores et (iii) technologies de séquençage basées sur des techniques de microscopie (Schadt *et al.*, 2010). Dans l'avenir, ces nouvelles technologies devraient pouvoir palier à certaines limites de la génération actuelle telle que la longueur des fragments générés mais elles présentent pour l'instant des taux d'erreur encore élevés par rapport aux autres technologies (15 à 20 %).

L'ensemble des technologies présentées précédemment met en avant l'énorme potentiel en terme de quantité de données générées et souligne ainsi les nouveaux défis à prendre en compte notamment en terme d'assemblage, de stockage et de traitements des données.

2. La métatranscriptomique

Savoir quelles sont les fonctions exprimées par les communautés microbiennes est un enjeu majeur pour comprendre le fonctionnement des écosystèmes et dans quelles mesures les microorganismes peuvent y jouer un rôle. Or les analyses des génomes ne peuvent à eux seuls expliquer les phénotypes observés car d'une part, tous les gènes présents ne sont pas obligatoirement exprimés et d'autre part, on ne peut pas distinguer les organismes morts et donc non actifs, des vivants (Bomar *et al.*, 2011). C'est dans ce contexte que les approches de métatranscriptomique ont été développées, dont les objectifs sont d'analyser la présence et l'abondance relative de tous les transcrits (molécules d'ARN) exprimés par l'ensemble des microorganismes d'un écosystème à un instant donné (Joyce & Palsson, 2006). Sachant qu'à partir d'un même gène, plusieurs copies d'ARNm peuvent être produites en fonction de l'activité des cellules, le transcriptome reflétera donc le niveau d'expression de tous les gènes à un temps t pour une condition physiologique donnée. La métatranscriptomique permet donc d'avoir un aperçu de l'activité biologique des microorganismes à un moment donné, alors que la métagénomique ne fait que prédire les fonctions potentiellement exercées par ces microorganismes.

Le développement de la métatranscriptomique est limité par de nombreux facteurs impliquant notamment la demi-vie très courte des ARN, la difficulté à éliminer les acides humiques pendant les processus d'extraction, la cinétique de transcription différentielle de même gènes présents dans différentes populations et la faible corrélation entre les niveaux d'ARN et la synthèse des protéines correspondantes (Maron *et al.*, 2007). Un autre problème résulte du fait que les ARNm, beaucoup moins abondants que les ARNr, sont souvent masqués lors de l'extraction de l'ARN total. Malgré cela, différentes techniques d'analyses ont été développées basées principalement sur l'hybridation moléculaire et le séquençage d'ADN. L'apparition des technologies de séquençage à haut débit a permis par la suite une amélioration considérable de notre compréhension de l'expression des gènes dans les communautés microbiennes environnementales (DeLong, 2009; Poretsky *et al.*, 2009).

2.1. Les premières méthodologies

Parmi les premières stratégies utilisées en métatranscriptomique on trouve deux approches majeures. Dans la première approche, l'identité des gènes d'intérêt est d'abord définie puis leur niveau d'expression est évalué par la quantification des ARN en utilisant des puces à ADN. En dépit de leur capacité à mesurer l'expression de milliers de gènes simultanément, ces méthodes présentent certains désavantages car elles nécessitent de connaître au préalable les gènes ou les séquences génomiques complètes ainsi que les annotations correspondantes (Dugat-Bony *et al.*, 2012). Par conséquent, ces techniques sont plutôt inefficaces pour la découverte de nouvelles fonctions dans l'environnement. De plus, ces méthodes sont caractérisées par une sensibilité de détection et une fiabilité de quantification limitées, avec par exemple des difficultés pour détecter des ARNm de très faible abondance et pour distinguer des séquences d'ARNm très proches du fait de problèmes d'hybridations croisées et de bruit de fond élevé (Casneuf *et al.*, 2007). La seconde approche se base d'abord sur l'identification des ARNm qui diffèrent dans leur niveau d'expression sous différentes conditions expérimentales suivie de l'identification des gènes respectifs. Cette approche emploie des techniques telles que l'affichage différentiel (Differential Display, DD) ou l'analyse en série de l'expression des gènes (Serial Analysis of Gene Expression, SAGE) (Velculescu *et al.*, 2000). Les analyses SAGE offrent de nombreux avantages par rapport aux puces à ADN tel que leur capacité à détecter de nouveaux transcrits et à obtenir une mesure directe de l'abondance des transcrits permettant une meilleure comparaison entre échantillons. Cependant, ces analyses sont basées sur la technologie coûteuse de séquençage Sanger et nécessitent encore une procédure de clonage laborieuse qui est également à l'origine de biais importants. De plus, elles ne fournissent aucune information sur les isoformes issus d'épissages alternatifs ou sur la découverte de nouveaux gènes (Mortazavi *et al.*, 2008).

2.2. Le séquençage de nouvelle génération

Un progrès considérable a été réalisé pour l'analyse efficace de profils d'expression plus complexes avec le développement des technologies de séquençage de nouvelle génération (454 de Roche, Genome Analyser d'Illumina, SOLiD de ABI). Ces nouvelles technologies permettent le séquençage direct de l'ADNc (Frias-Lopez *et al.*, 2008; Gilbert *et al.*, 2008) évitant ainsi les biais inhérent à l'étape de clonage et produisant des quantités massives de données nécessaires à la réalisation d'études à grande échelle. Cette approche est appelée

RNA-Seq (Wang *et al.*, 2009). L'utilisation des NGS fournit une grande profondeur de séquençage qui est un outil extrêmement puissant pour la quantification fiable des transcrits (Mortazavi *et al.*, 2008; Sultan *et al.*, 2008). De plus, un certain nombre d'études ont démontré que cette approche est considérablement plus sensible pour la détection de transcrits rares (Cloonan *et al.*, 2008; Sultan *et al.*, 2008) et permet d'obtenir une gamme de mesure de l'expression des gènes beaucoup plus vaste que celle des puces à ADN (Marioni *et al.*, 2008; Mortazavi *et al.*, 2008; Sultan *et al.*, 2008). D'autre part, une autre approche basée sur les technologies de 3^{ème} génération (Analyseur Hélicos) a récemment été développée et implique le séquençage direct des molécules d'ARN, sans conversion préalable en ADNc, éliminant ainsi les biais introduits lors de la synthèse des ADNc (Ozsolak *et al.*, 2009). Cette technologie permet de déterminer à la fois l'identité et l'abondance des molécules d'ARNm en un seul processus analytique (Zhang *et al.*, 2010).

Ces dernières années, le séquençage de transcriptome a été utilisé pour diverses applications passant par le profilage de l'expression des gènes (annotation des génomes), la détection de réarrangements ou bien la découverte d'ARN non-codants. La métatranscriptomique représente ainsi un domaine clé d'investigations biologiques apportant des informations essentielles sur la diversité taxonomique (ARNr) et fonctionnelle (ARNm) des communautés actives et pouvant servir par exemple à identifier les conditions environnementales qui ont un impact majeur sur le fonctionnement de ces communautés.

3. La métaprotéomique

Un intérêt particulier est porté aux protéines qui sont des molécules vitales pour les organismes car elles représentent les principaux composants de la structure cellulaire, servent d'enzyme catalytique dans les voies métaboliques et de protéines de transduction du signal dans les voies de régulation (Graham *et al.*, 2007). Cet intérêt s'est traduit par l'apparition des approches de métaprotéomique qui ont pour but d'identifier et de quantifier l'ensemble des protéines exprimées par une communauté à un instant donné. L'importance de ces nouvelles approches est également soulignée par le fait que les protéines sont des marqueurs de la fonction biologique plus précis que les gènes fonctionnels et leurs ARN messagers correspondants. En effet, il faut savoir qu'il n'existe pas nécessairement de corrélation directe entre le niveau d'expression d'un gène (quantité d'ARNm) et la quantité de protéines produite (Anderson & Seilhamer, 1997; Pradet-Balade *et al.*, 2001). De même, des épissages

alternatifs au niveau des transcrits ou des modifications post-traductionnelles des protéines sont susceptibles de modifier les fonctions biologiques associées aux protéines et ne sont donc accessibles qu'en étudiant les protéines elles mêmes (Wilmes & Bond, 2006).

La protéomique a émergé dans les années 1970 avec l'analyse des profils protéiniques par électrophorèse bidimensionnelle (2-D) (O'Farrell, 1975). Cette technique a rendu possible la séparation des protéines issues de mélanges complexes d'extraits cellulaires, permettant ainsi l'analyse de la réponse bactérienne à différentes conditions de croissance (Pedersen *et al.*, 1978). Cependant, l'identification des protéines était longue et fastidieuse en raison d'un manque d'informations disponibles sur les séquences génomiques et de l'absence de technologies de séquençage précises et rapides pour l'analyse des protéines. Depuis les années 1990, la protéomique et par conséquent, la métaprotéomique ont fortement progressé du fait de trois révolutions majeures : (i) l'augmentation énorme des données génomiques et métagénomiques constituant une base solide pour l'identification des protéines, (ii) les progrès considérables dans la sensibilité et la précision des spectromètres de masse permettant une meilleure identification et quantification des protéines ainsi que la détermination des modifications post-traductionnelles et (iii) les améliorations réalisées en terme de puissance de calcul et de bioinformatique permettant le traitement et l'analyse de jeux de données substantiels (Schneider & Riedel, 2010).

Parmi les différentes techniques utilisées, la spectrométrie de masse (Mass Spectrometry, MS) est apparue comme la principale méthode permettant de caractériser la présence et l'identité des protéines dans les systèmes biologiques (Cravatt *et al.*, 2007). Cependant, l'application de cette méthode requiert une réduction préalable de la complexité du protéome sinon le signal relatif aux peptides majoritaires a tendance à masquer celui des peptides minoritaires. Ainsi la technique MS fait appel à deux stratégies majeures incluant une étape de fragmentation préalable du protéome qui peut être réalisée soit au niveau protéinique soit au niveau peptidique après digestion enzymatique des protéines. La première stratégie est basée sur la séparation des protéines ou des peptides par électrophorèse sur gel bidimensionnelle (2D-PAGE) couplée à la spectrométrie de masse et la seconde sur la séparation des protéines/peptides par chromatographie liquide (LC) uni- ou bidimensionnelle suivie par la spectrométrie de masse en tandem (MS/MS) (Keller & Hettich, 2009).

3.1. Techniques de séparation des protéines/peptides

3.1.1. Techniques basées sur l'électrophorèse

La méthode la plus utilisée est l'électrophorèse bidimensionnelle sur gel de polyacrylamide ou 2D-PAGE (Two-Dimensional PolyAcrylamide Gel Electrophoresis), combinant une première séparation des protéines en fonction de leur point isoélectrique par isoélectrofocalisation (IEF) suivie d'une deuxième séparation réalisée dans une direction perpendiculaire à la première, en fonction de leurs masses apparentes par SDS PAGE (Schneider & Riedel, 2010). Ensuite, une coloration des gels au nitrate d'argent ou au bleu de Coomassie permet la visualisation des protéines sous la forme de spots. L'abondance relative des protéines est déduite de la taille et de l'intensité des spots. En pratique, selon le type d'échantillon et la sensibilité de la méthode de détection, cette technique permet de séparer un maximum de 3000 spots (Lilley *et al.*, 2002). Les spots protéiques peuvent ensuite être excisés et digérés par une enzyme avant analyse par spectroscopie de masse. Une amélioration significative de cette technologie en terme de couverture et de sensibilité a été réalisée dans les années 1990, avec l'apparition de l'électrophorèse 2D-DIGE (2-Dimensional Differential in-Gel Electrophoresis), permettant de marquer les différents échantillons avec des colorants fluorescents.

Cependant, la technique des gels 2D présente de nombreuses limites incluant la difficulté à détecter les protéines de faible abondance, de haut poids moléculaire (> 200 kDa), très hydrophobes, très acides ($pI < 3$) ou très basiques ($pI > 10$) sans parler des protéines membranaires rarement observées (Maron *et al.*, 2007). De telles protéines peuvent représenter au total jusqu'à la moitié des protéines exprimées par un organisme (Shen & Smith, 2002). De même des problèmes de co-migration de protéines et d'isoformes protéiques empêchent d'obtenir une identification et une quantification précise. Enfin, cette méthode reste difficilement automatisable, ce qui limite toutes analyses à haut débit. Dans ce contexte, des stratégies alternatives de séparations ont été développées.

3.1.2. Techniques basées sur la chromatographie liquide

Une alternative à la technique 2D-PAGE est la chromatographie liquide (LC), une méthode de séparation uni- ou multidimensionnelle des protéines ou des peptides générés par la digestion enzymatique des protéines. Différents types de LC peuvent être distingués en fonction du

mode de séparation des protéines/peptides utilisés. Cette séparation peut se faire selon leur taille (masse), leur pI (charge) ou leur hydrophobicité, qui sont les trois caractéristiques physicochimiques principales qui définissent les protéines/peptides (Shen & Smith, 2002). Toutes les méthodes de chromatographies (gel filtration, échange d'ions, interaction hydrophobe ...) sont utilisables en analyse protéomique, mais afin d'augmenter le pouvoir de séparation et la résolution finale, ces méthodes sont employées en général sur 2 dimensions, avec le plus souvent une chromatographie d'échange d'ions (IEC) en premier où les protéines sont séparées en fonction de leur charge, suivie d'une chromatographie en phase inverse (RP-LC) où elles sont éluées en fonction de leur degré d'hydrophobicité. La chromatofocalisation qui est une variante de l'IEC est également très utilisée. Elle permet de séparer les protéines en fonction de leur pI grâce à l'établissement d'un gradient de pH dans la colonne au cours de l'éluion.

Les avantages de ces techniques de séparation reposent sur le peu de matériel biologique requis (40 µg vs 200 µg en 2D-PAGE) et surtout sur la meilleure gamme dynamique car l'on peut détecter, en fonction de la technique utilisée beaucoup plus de protéines sur une plus grande gamme de pI et d'hydrophobicité qu'en 2D-PAGE. En effet, la combinaison des divers modes chromatographiques a permis d'améliorer entre autre la détection des protéines de faible abondance et des protéines membranaires insolubles (Wolff *et al.*, 2006) d'où l'intérêt croissant de cette technique en protéomique. De plus, ces approches permettent un grand niveau d'automatisation, c'est à dire une connexion en ligne de la chromatographie et de la spectrométrie de masse permettant la génération de milliers de spectres de masse par heure et facilitant ainsi grandement les analyses à haut débit (Schneider & Riedel, 2010). Cependant, malgré ces améliorations, ni les approches 2D-PAGE ni les approches LC-MS/MS ne permettent une détection exhaustive même dans les petits protéomes bactériens car on estime que seulement 20-40 % des protéomes peuvent être détectés et ce sans fournir d'informations quantitatives. Afin d'obtenir des mesures plus précises du protéome, des méthodes de protéomiques quantitatives peuvent être employées tels que les marqueurs d'affinité contenant un isotope d'identification (Isotope Coded Affinity Tags, ICAT) ou les marqueurs isobariques (isobaric Tags for Relative and Absolute Quantitation, iTRAQ) (Haqqani *et al.*, 2008).

3.2. Ionisation et spectrométrie de masse

La spectrométrie de masse sert à établir les masses moléculaires et la structure des protéines/peptides. La caractérisation du composé se fait en analysant les ions formés à partir de la molécule étudiée. Les masses des peptides ainsi obtenus expérimentalement sont comparées aux masses théoriques obtenues par digestion *in silico* des protéines présentes dans les banques de données, ce qui permet de retrouver l'identité de la protéine d'intérêt. Un spectromètre est composé de différentes parties : une source d'ions (ESI, MALDI), un analyseur qui sépare les ions selon leur rapport masse/charge (quadripôle, TOF, FT-ICR) et un détecteur qui donne l'intensité des ions correspondant à chaque espèce (Aebersold & Mann, 2003).

Il existe deux stratégies pour séparer et identifier les protéines issues de mélanges complexes. Le protéome peut être étudié soit par des analyses « top-down » des protéines intactes ou par des analyses « bottom-up » des peptides générés par la digestion des protéines par la trypsine (Chait, 2006). L'approche bottom-up est une méthode permettant la localisation de la répartition des peptides générés par digestion enzymatique *in situ* et leur structure par analyse MS en tandem (MS/MS). Les spectres MS/MS des peptides fournissent des informations sur leurs séquences d'acides aminés. Cependant cette méthode souffre généralement d'une couverture incomplète de la séquence des protéines ayant pour conséquence une perte d'informations sur les modifications post-traductionnelles (MPT) (Wu *et al.*, 2009). Une perte d'informations concernant les peptides de faible abondance dans les spectres de masse dominés par des peptides majoritaires, a également été démontrée. Par contre, la méthode top down présente une couverture de 100 % des séquences des protéines permettant ainsi la détection des protéines isoformes, des événements protéolytiques et des MPT (Wu *et al.*, 2009). Malgré cela, cette stratégie présente des limites en termes de sensibilité et de débit avec notamment des difficultés à analyser des protéines intactes de plus de 50 kDa.

Indépendamment de l'approche utilisée, les sources d'ionisation les plus fréquentes sont les sources MALDI (Matrix Assisted Laser Desorption Ionisation) qui permettent la formation d'ions par excitation laser d'une matrice contenant l'analyte cristallisé (Hillenkamp *et al.*, 1991) ou une source ESI (ElectroSpray Ionization) qui permet la formation d'ions par électronebulisation de l'analyte (Fenn *et al.*, 1989). Ces sources d'ions peuvent ensuite être couplées à divers analyseurs dont les 3 principaux types sont : l'analyseur à temps de vol

(Time Of Flight, TOF), les trappes d'ions (Ion Trap, IT) ou les analyseurs à transformée de Fourier (FT-ICR, Orbitrap). Chacun de ces analyseurs présente une résolution, une sensibilité, une gamme de précision ainsi qu'une capacité à produire des spectres de masses qui lui sont propres. Ces analyseurs peuvent être utilisés individuellement ou en association dans le cas des analyses en tandem (Domon & Aebersold, 2006). Dans ce cas les peptides ionisés initiaux, appelés ions parents, sont fragmentés en sous-produits avant de pénétrer dans le second analyseur où le rapport masse sur charge de chaque sous-fragment est analysé. Cette analyse en tandem permet de déduire la séquence en acides aminés de chaque peptide et facilite ainsi l'identification de la protéine correspondante dans les bases de données. C'est une méthode plus sélective et plus sensible que la MS seule qui permet d'élucider des structures de composés inconnus.

La métaprotéomique est une approche qui apporte des renseignements sur la fonctionnalité réelle par rapport aux voies métaboliques par l'identification des protéines fonctionnelles présentes (Maron *et al.*, 2007; Bastida *et al.*, 2009). Cependant, étant donné le nombre très important de protéines existantes, leurs fonctions ne sont pas forcément connues et ces outils, bien que prometteurs, ne permettent d'obtenir qu'une image partielle de l'activité des communautés. De plus les méthodes actuelles ne permettent pas de réaliser des analyses exhaustives du métaprotéome car il subsiste une importante sous estimation du nombre de protéines détectées dans un échantillon du fait des limites technologiques. D'autre part la synthèse d'une protéine et donc sa présence, ne signifie pas nécessairement que l'activité biologique qui lui est associée soit également présente. L'activité biologique d'une protéine peut par exemple être régulée par des modifications post-traductionnelles. Il est donc nécessaire d'utiliser des approches alternatives ou complémentaires (mesure directe de l'activité, dosage des substrats et produits...) afin d'obtenir une vision plus précise des fonctions biologiques présentes.

4. La méta-métabolomique

La méta-métabolomique a pour objet l'étude du méta-métabolome, c'est-à-dire l'ensemble des molécules de faible poids moléculaire (métabolites) présentes dans un échantillon environnemental. Les métabolites sont les produits finaux des processus de régulation cellulaire et leur synthèse, transport, transformation et/ou dégradation sont en général régulés et catalysés par des partenaires protéiques, principales molécules effectrices de la réponse

biologique. L'étude des métabolites peut donc se révéler fondamentale pour confirmer ou non l'activité biologique des protéines qui leur sont associées. L'étude du métabolome s'inscrit au même titre que les études du transcriptome et du protéome dans un contexte post-génomique. La métabolomique peut être perçue comme le maillon final de la cascade des approches « omic » car le métabolome représente l'ultime réponse d'un système biologique à une altération génétique, une pathologie ou des changements environnementaux. Le métabolome est donc le reflet le plus proche de la fonctionnalité et du phénotype d'un système biologique (Fiehn, 2002; Weckwerth, 2003).

Le métabolome est constitué d'un grand nombre de composants de différentes natures, tels que les acides aminés, des lipides, les acides organiques, les nucléotides, etc. Ces composés sont également très divers dans leurs propriétés physiques et chimiques et présentent une grande gamme de concentrations. Par conséquent, l'étude du métabolome constitue un défi majeur pour la chimie analytique ; l'analyse quantitative de tous les métabolites, ne peut être atteinte avec les technologies actuelles (Dettmer *et al.*, 2007). Malgré cette limitation, de nombreuses méthodes analytiques ont été développées et appliquées.

Les études de métabolomique reposent généralement sur des méthodes de spectrométrie de masse couplées à une méthodes de séparation telle que la chromatographie liquide (LC/MS), la chromatographie gazeuse (GC/MS) ou l'électrophorèse capillaire (CE/MS). La spectroscopie par résonance magnétique nucléaire (RMN) est également très utilisée (Zhang *et al.*, 2010). Les stratégies analytiques employées dépendent des résultats attendus et d'éventuelles voies métaboliques d'intérêt, elles sont généralement classées en 3 catégories : l'analyse ciblée qui est une analyse quantitative centrée sur un petit nombre de métabolites produit par une protéine cible, le profilage métabolique basé sur l'analyse semi-quantitative de tous les composés appartenant à une voie ou à une famille métabolique donnée et enfin l'empreinte métabolique visant à comparer des spectres.

4.1. La spectroscopie par résonance magnétique nucléaire

La spectroscopie RMN est l'une des premières méthodes d'analyse à avoir été appliquée avec succès à l'étude du métabolome (Nicholson *et al.*, 2002; Bailey *et al.*, 2003; Mayr *et al.*, 2005). Cette méthode est basée sur les propriétés de résonance magnétique de certains noyaux atomique. Le principe est analogue à celui de l'imagerie par résonance magnétique (I.R.M.) :

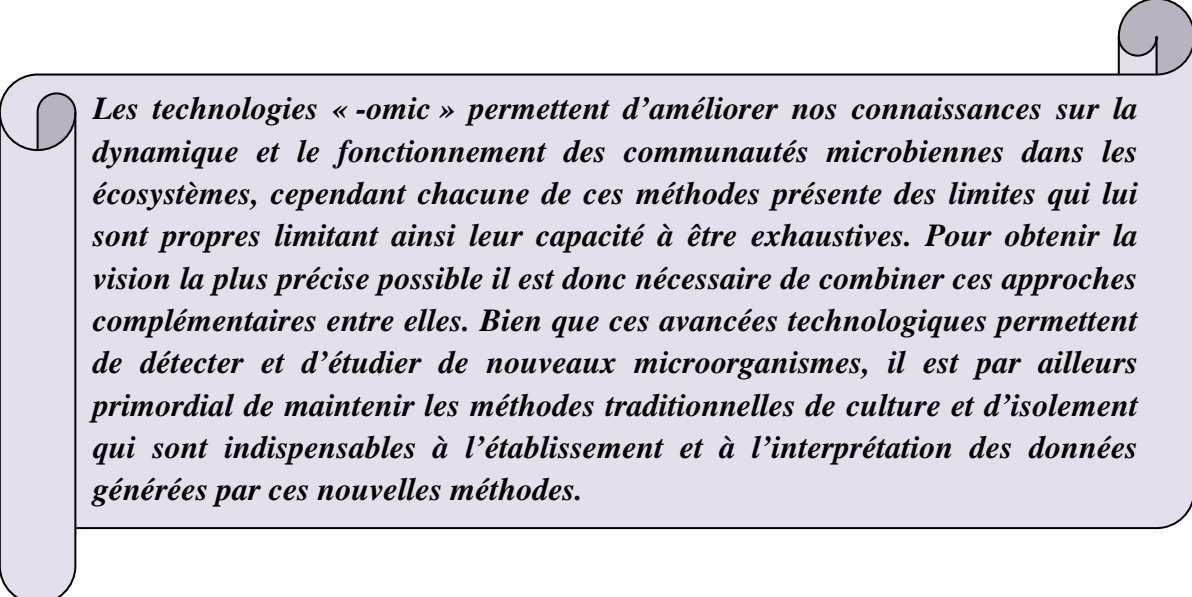
ces noyaux, placés dans un champ magnétique puissant et soumis par intermittence à une onde de radiofréquence, émettent un signal caractéristique. La RMN présente un certain nombre d'avantages : elle est non destructive, rapide, simple à mettre en œuvre, robuste et reproductible. De plus elle permet d'obtenir une image qualitative et quantitative d'une grande diversité de métabolites possédant un noyau RMN (^1H , ^{13}C , ^{19}F , ^{31}P ...). Cependant, elle reste une méthode d'analyse peu sensible et de faible résolution qui nécessite une grande quantité d'échantillon par analyse (souvent au-delà du milligramme). Il est tout de même possible de palier le problème en couplant une détection RMN avec une séparation chromatographique (LC).

4.2. La spectrométrie de masse

De part sa plus grande sensibilité par rapport à la méthode de spectrométrie RMN, la spectrométrie de masse est de plus en plus utilisée comme alternative pour l'analyse des métabolites (Mayr, 2008). La spectrométrie de masse permet la détection d'espèces au μM au cours de l'acquisition des empreintes, donnant accès à un grand nombre de classes de métabolites, détectables à leur concentration physiologique. De plus, elle autorise l'identification des métabolites *via* l'obtention de leur masse moléculaire, avec une précision dépendant du type d'analyseur et *via* l'étude de leur spectre de fragmentation qui peut être comparé avec celui d'un composé de référence qui fournit également des informations structurales. Ces propriétés permettent donc d'effectuer des identifications *de novo* de composés qui ne sont pas décrits dans les bases de données et d'atteindre une plus grande certitude dans l'identification des composés déjà décrits. Actuellement, de nombreux métabolites mis en évidence lors d'expériences de métabolomique n'ont pas encore été identifiés car les bases de données métaboliques actuelles n'ont pas de bibliothèques de spectres exhaustives. Cela constitue une limitation majeure à l'expansion de cette approche dans la mesure où l'identification de nouveaux composés et la caractérisation reproductible de ceux déjà décrits au sein de larges jeux de données est un prérequis au développement des approches métabolomiques.

Comparé à la transcriptomique et à la protéomique, les technologies utilisées pour analyser les métabolites sont moins matures, ainsi la plupart des études de métabolomique réalisées jusqu'à présent ne sont pas encore suffisamment exhaustives (en moyenne, seulement quelques dizaines de métabolites sont identifiés dans ces études) et leur précision doit

également être améliorée (Fiehn, 2001; Kell, 2004; Cascante & Marin, 2008). Cependant, ces études ont également permis de démontrer l'intérêt de telles approches pour décrypter le métabolisme des organismes et pour combler l'écart existant entre le phénotype et le génotype car la métabolomique est le domaine d'étude qui fournit la meilleure représentation du phénotype d'un organisme (Cascante & Marin, 2008).



Les technologies « -omic » permettent d'améliorer nos connaissances sur la dynamique et le fonctionnement des communautés microbiennes dans les écosystèmes, cependant chacune de ces méthodes présente des limites qui lui sont propres limitant ainsi leur capacité à être exhaustives. Pour obtenir la vision la plus précise possible il est donc nécessaire de combiner ces approches complémentaires entre elles. Bien que ces avancées technologiques permettent de détecter et d'étudier de nouveaux microorganismes, il est par ailleurs primordial de maintenir les méthodes traditionnelles de culture et d'isolement qui sont indispensables à l'établissement et à l'interprétation des données générées par ces nouvelles méthodes.

RÉSULTATS

Chapitre I – Caractérisation de la communauté bactérienne active

A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE

1. Contexte général

Le site minier de Carnoulès appartient à la commune gardoise de Saint Sébastien d'Aigrefeuille. Il est situé au sud du Massif Central et appartient à la bordure métallifère cévenole, à environ 80 km au nord-est de Montpellier et à 10 km à l'ouest d'Alès (Figure 13).

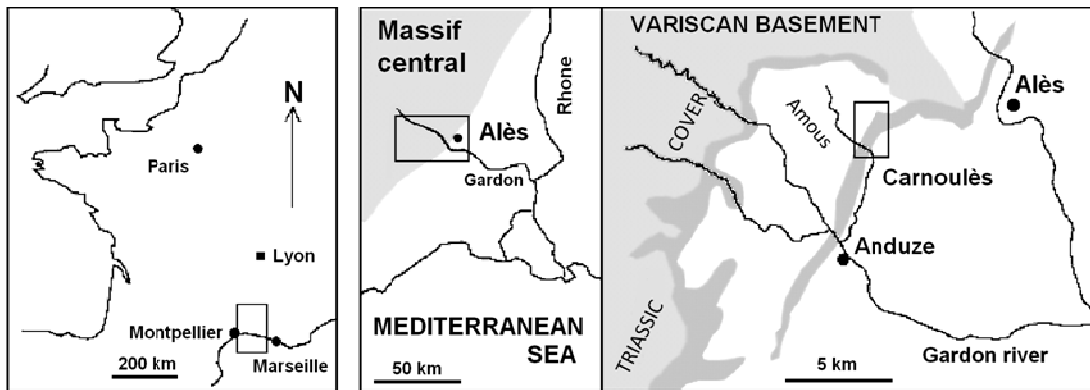


Figure 13 : Localisation géographique de la mine de Carnoulès

Le site géologique est majoritairement constitué d'arkose (résultant de la roche gneissique des Cévennes, érodée au Trias inférieur renfermant des minerais de zinc, de plomb, de fer et pouvant contenir jusqu'à 4 % d'arsenic. L'exploitation du plomb et du zinc de la mine de Carnoulès a débuté à partir de 1883 par la concession de Saint Sébastien d'Aigrefeuille puis différentes sociétés se sont succédées jusqu'en 1930. A cette époque, l'exploitation en galeries s'est arrêtée en raison des difficultés de traitement et de la chute du cours du plomb. De 1951 à 1962, le groupe Pennaroya-Métaleurope a repris l'exploitation de manière plus intensive, à ciel ouvert et en carrière avec d'importants tonnages journaliers. En 1962, la mine n'était cependant plus rentable et l'exploitation a fermé le 24 octobre 1963.

L'exploitation de cette mine a généré environ 1,2 Mt de déchets majoritairement constitué de sables à pyrite très fins (30 μm) contenant 75 % de quartz et entre 5 et 15 % de pyrite qui

contient de 1 à 4 % d'arsenic (Leblanc *et al.*, 1996). Au fond du stock, se trouve du matériel ferrugineux relativement grossier (200 µm) et fortement oxydé. Ce stock de déchet, confiné derrière une digue, a une épaisseur de 10 à 24 m et s'étend sur une surface de 5500 m². Une couche d'argile de 0,3 m a été ajoutée afin de permettre un recouvrement végétal limitant la dispersion des sables. Les eaux d'infiltration qui percolent à travers les déchets, émergent à la base de la digue et constituent la source initiale du Reigous. Ce ruisseau qui draine le site minier du nord au sud est acide (pH~3) et riche en fer dissous, sulfate et arsenic (500-1000, 2000-7700 et 100 à 350 mg L⁻¹, respectivement) (Leblanc *et al.*, 1996; Casiot *et al.*, 2003a). Le Reigous se jette à 1.5 km de là, dans la rivière Amous, elle-même affluent du Gardon d'Anduze (Figure 14). En aval de la confluence Reigous-Amous, la qualité écologique de l'Amous est grandement détériorée, il n'y a plus de traces d'organismes supérieurs et de vie piscicole sur plusieurs kilomètres.

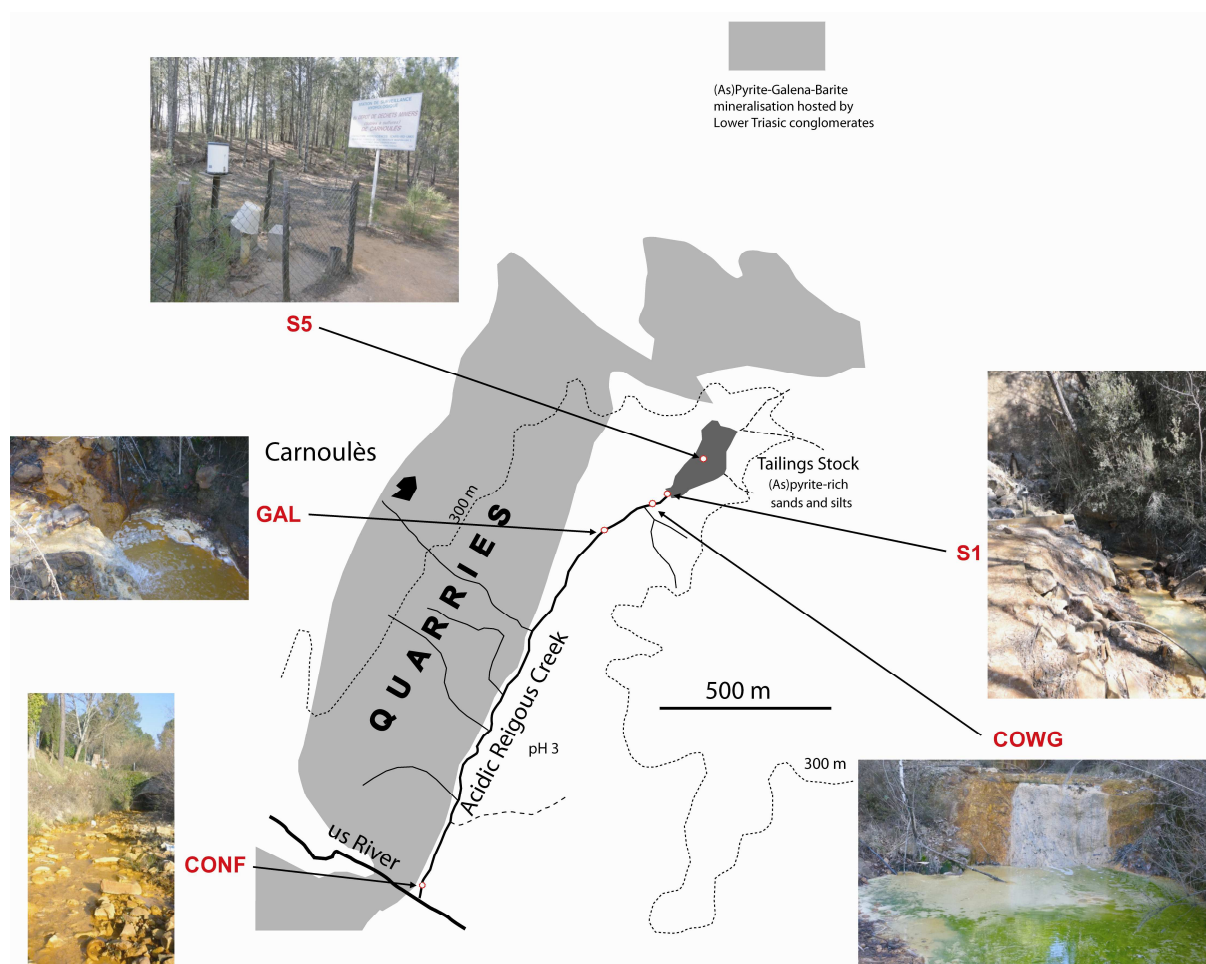


Figure 14 : Topologie du DMA de Carnoulès. En rouge sont indiqués les sites d'échantillonnages. S5 : au centre du stérile minier où a été placé un piézomètre. S1 : à la source du Reigous. COWG : à 30 m de la source. GAL : à 150 m de la source. CONF : à 1,5 km de la source, juste avant la confluence entre le Reigous et l'Amous

Ce site fournit un exceptionnel exemple d'adaptation de la vie à un environnement extrême. En effet, en plus d'avoir un pH acide, il présente des concentrations en As considérables (de 100 à 350 mg L⁻¹) dans le Reigous, qui ont même atteint des valeurs parmi les plus importantes au monde (jusqu'à 10 g L⁻¹) dans le stock de déchets miniers (Casiot *et al.*, 2003b) C'est pourquoi, ce site, considéré comme un site atelier, est étudié depuis une quinzaine d'années à la fois d'un point de vue géochimique (Leblanc *et al.*, 1996; Casiot *et al.*, 2003a, 2003b, 2009) que microbiologique (Duquesne *et al.*, 2003; Morin *et al.*, 2003; Bruneel *et al.*, 2005, 2006, 2008, 2011; Bertin *et al.*, 2011). Ce site présente également un intérêt particulier car des études ont montré la présence d'un phénomène de remédiation naturelle du milieu avec une atténuation de la contamination le long du continuum. Les concentrations en arsenic atteignent en effet jusqu'à 300 mg L⁻¹ (Bruneel *et al.*, 2006) à la source alors que 1500 mètres plus loin les valeurs sont proches de celles imposées par les normes européennes pour les eaux de consommation (0,01 mg L⁻¹ - Directive européenne n° 98/83/CE). Plus précisément, il a été montré que dans les 30 premiers mètres du Reigous, l'oxydation du Fe(II) conduit à la co-précipitation de 20 à 60 % de l'As dissous et génère ainsi des sédiments riches en fer et en sulfates, contenant jusqu'à 22 % d'As (Casiot *et al.*, 2003b; Egal *et al.*, 2010). Le Reigous présente des variations saisonnières fortement marquées : en période d'étiage, les sédiments arséniés s'accumulent mais, en période de fortes pluies (printemps, automne), les sédiments sont érodés et transportés, entraînant une forte augmentation du flux d'arsenic avec un transport essentiellement sous forme particulaire. Ainsi, les conditions météorologiques, en particulier la pluviométrie, influencent considérablement la dynamique écologique de ce type de milieu, modifiant les conditions physicochimiques dans le temps (Edwards *et al.*, 1999). Les variations des conditions physicochimiques au sein d'un même DMA montrent finalement un gradient à la fois spatial et temporel sur le continuum.

Plusieurs souches de *Thiomonas* et d'*Acidithiobacillus ferrooxidans* ont été isolées à partir des eaux du Reigous. Sur la base d'expériences de laboratoire, il a été montré que certaines de ces bactéries pourraient contribuer à l'oxydation du fer et la co-précipitation de l'As (Bruneel *et al.*, 2003; Casiot *et al.*, 2003a; Morin *et al.*, 2003; Egal *et al.*, 2010). Des analyses préliminaires ont montré que les ADN de ces deux bactéries sont présents dans le Reigous, ainsi que celui de *Gallionella* sp., de *Thiobacillus* sp. et de certaines bactéries sulfato-réductrices (Bruneel *et al.*, 2006). Il a été montré que *G. ferruginea* était capable de diminuer

efficacement les concentrations en As (III et V) dans l'eau par co-précipitation avec le fer (Katsoyiannis & Zouboulis, 2004). Cet organisme pourrait donc aussi jouer un rôle très important dans le processus de remédiation observé au sein du Reigous. La présence d'Archaea a également été détectée dans le Reigous (*Ferroplasma acidiphilum* et des séquences affiliées à des *Thermoplasmatales* non cultivés), de même que celle d'un microorganisme eucaryote, *Euglena mutabilis* (Casiot *et al.*, 2004; Bruneel *et al.*, 2006). Au niveau du piezomètre S5, situé au cœur du stérile minier (où a lieu le biolessivage), on trouve parmi les bactéries une majorité de bactéries sulfato-réductrices, comme *Desulfosarcina variabilis* mais également des *Acidithiobacillus ferrooxidans* et des bactéries du genre *Thiobacillus*. Parmi les Archaea on trouve une très large majorité de *Ferroplasma acidiphilum*, microorganisme oxydant le fer et connu pour son rôle dans la génération des DMA (Baker & Banfield, 2003).

2. Objectifs de l'étude

Un certain nombre d'études ont donc déjà été réalisées sur cet écosystème mettant en évidence la présence de diverses communautés microbiennes et ont apporté des informations indispensables à la compréhension des processus biogéochimiques intervenant dans cet environnement particulier. Cependant, il s'agissait généralement d'études sporadiques réalisées sur un nombre restreint de point d'échantillonnage et de date de prélèvement, ce qui n'a pas permis la prise en compte de l'évolution spatiotemporelle des contraintes physicochimiques *in situ*. De plus, la majorité des analyses ont été effectuées dans l'eau et non pas dans les sédiments, lieu de stockage, plus ou moins pérenne, des métaux et métalloïdes précipités. D'autre part, parmi les bactéries détectées, il faut souligner que toutes ne sont pas fonctionnellement actives et il semble donc primordial de pouvoir différencier les cellules bactériennes mortes ou inactives des cellules fonctionnelles.

Ainsi, afin d'améliorer nos connaissances sur la diversité taxonomique et d'identifier les communautés bactériennes actives dans cet environnement extrême, deux approches complémentaires ont été utilisées dans le cadre d'une première étude. Un inventaire taxonomique basé sur les gènes codant pour l'ARNr 16S a permis d'identifier les populations bactériennes présentes dans l'eau et les sédiments, ceci constituant la part expérimentale de mon travail dans cette analyse. En parallèle, une approche de métaprotéomique a été mise en place et a permis de mettre en évidence les espèces actives dans les sédiments.

B. RÉSULTATS

Les analyses taxonomiques des banques de gènes codant pour l'ARNr 16S ont permis d'identifier un total de 17 espèces bactériennes dont seulement 7 genres se sont révélés être présents à la fois dans l'eau et dans les sédiments. La plupart des ces bactéries étaient affiliées à des β -protéobactéries telles que *Gallionella* ou *Thiomonas* mais également à des γ -protéobactéries (e.g. *Acidithiobacillus ferrooxidans*), des α -protéobactéries (e.g. *Acidiphilium*), des δ -protéobactéries (e.g. *Desulfomonile limimaris*), des *Nitrospira* (e.g. *Leptospirillum ferrooxidans*), des *Actinobacteria* et des *Firmicutes*. Il s'agit majoritairement d'espèces trouvées communément dans les DMA (Baker & Banfield, 2003; Hallberg, 2010). La diversité relativement faible mise en évidence dans cet écosystème, résultant probablement de la concentration élevée en composés toxiques, s'est révélée cohérente avec des études antérieures indiquant que la biodiversité des eaux acides minières se limite principalement à des bactéries spécialisées (Rowe *et al.*, 2007).

D'autre part, dans les deux compartiments analysés, eau et sédiment, les bactéries détectées étaient principalement impliquées dans les cycles du fer, de l'arsenic et du soufre. Les bactéries impliquées dans l'oxydation du Fe(II) étaient affiliées à *Gallionella*, *A. ferrooxidans*, *Ferrimicrobium*, *Leptospirillum*, *Sideroxydans lithotrophicus*, et *Ferrovum myxofaciens* alors que la réduction du fer a été mise en évidence pour les bactéries des genres *Acidiphilium* spp., *Acidocella*, *Desulfuromonas svalbardensis*, *Rhodoferax ferrireducens*, *A. ferrooxidans* et *Ferrimicrobium acidiphilum*. Concernant le cycle du soufre, des populations capables d'oxyder des composés soufrés inorganiques ont été mise en évidence telles que *Thiobacillus*, *Thiomonas* ou *A. ferrooxidans*. Des bactéries sulfato-réductrices comme *Desulfomonile limimaris* ou *Desulfuromonas svalbardensi* ont également été identifiées et pourraient être impliquées dans la consommation du sulfate.

Au total 89 protéines ont été identifiées par les analyses de métaprotéomique, dont seulement 44 %, soit 39 protéines, étaient issues des bactéries. Cette étude a permis de caractériser plusieurs protéines de ménage, c'est-à-dire des protéines nécessaires au fonctionnement de la cellule et qui ne doivent pas, en conséquence, jouer de rôles spécifiques dans cet environnement. Des protéines impliquées dans le mécanisme de détoxification du formaldéhyde ainsi qu'une protéine externe OrpI et une protéine de fonction inconnue ont également été mises en évidence. Enfin, plus de la moitié des protéines identifiées étaient

similaires aux chaperonines GroEL et ont servi d'outil taxonomique pour identifier les bactéries fonctionnellement actives dans l'ensemble de la communauté. En effet, le gène *groEL* est conservé au sein des procaryotes et présente certaines spécificités qui en font un bon marqueur phylogénétique. Les données ainsi obtenues ont permis de révéler pour la première fois l'expression *in situ* de protéines par les bactéries appartenant aux genres *Gallionella*, *Thiomonas*, *A. ferrooxidans*, et *Acidiphilium*.

C. DISCUSSION ET CONCLUSION

Peu de protéines bactériennes ont été identifiées dans cette étude et la majorité appartenaient aux 4 genres cités précédemment. Ces résultats sont probablement dus en partie aux limites technologiques propres à l'approche de métabolomique. En effet, la technique utilisée (*i.e.* l'électrophorèse bidimensionnelle) ne permettait d'identifier que les protéines cytoplasmiques les plus abondantes. C'est pourquoi, certaines protéines susceptibles d'être exprimées dans cet écosystème, telle que la rusticyanine, une protéine membranaire issue d'*A. ferrooxidans* connue pour être impliquée dans l'oxydation du Fe(II), n'ont pas pu être détectées dans cette étude. Cependant, il est intéressant de noter qu'une étude postérieure utilisant la chromatographie liquide couplée à la spectrométrie de masse (LC-MS/MS) a palié ce problème et permis la mise en évidence de cette protéine (Bertin *et al.*, 2011). D'autre part, il faut noter que la plupart des bactéries présentes dans cet écosystème n'ont pas été cultivées ni étudiées *in vitro* jusqu'à présent. Ainsi, leurs séquences génomiques et par conséquent leurs séquences protéiniques indispensables à l'identification par spectrométrie de masse, ne sont pas forcément disponibles dans les bases de données publiques. L'ensemble de ces constatations souligne donc les faiblesses de cette approche qui ne permet à ce jour d'obtenir qu'une image partielle de l'activité des communautés comme nous l'avons vu dans la partie synthèse bibliographique (paragraphe C.4).

Cependant, la combinaison entre l'approche métabolomique et l'approche génomique a toutefois permis de mettre en évidence la présence de bactéries connues pour leurs implications dans le cycle du fer, de l'arsenic et du soufre ainsi que leur activité métabolique *via* l'expression active de protéines. Ces résultats renforcent donc l'hypothèse selon laquelle ces bactéries pourraient être fortement impliquées dans le processus de remédiation observé *in situ*. En effet, la grande quantité de sédiments constitués d'hydroxysulfate de Fe(III)-As(V) observée sur ce site serait probablement le résultat de l'activité conjointe de microorganismes

oxydant le fer tels que *A. ferrooxidans*, *Alicyclobacillus ferrooxidans*, *Ferrimicrobium* ou *Gallionella* et de microorganismes oxydant l'arsenic tel que *Thiomonas* sp.. De plus, les microorganismes autotrophes oxydant Fe, As et S seraient susceptibles de fournir les sources de carbone nécessaires à la croissance des hétérotrophes comme *Acidiphilum* présent dans cet écosystème.

En conclusion, les résultats obtenus ont souligné l'importance des approches de génomique fonctionnelle pour la description des communautés bactériennes comme celles présentes au sein de ce type d'environnement extrême et pour la détermination, du moins en partie, de leur contribution aux processus de remédiation naturel *in situ*. L'ensemble de ces données ont donc permis une meilleure compréhension du fonctionnement de cet écosystème et il semble indispensable d'utiliser des approches similaires dans les études futures afin de (i) compléter les données obtenues sur les processus fonctionnels à l'œuvre dans cet écosystème et (ii) déterminer le rôle joué par les bactéries actives les moins abondantes qui ont été identifiées dans cette étude.

D. CONTRIBUTION SCIENTIFIQUE

Ce travail a conduit à une publication scientifique dans la revue « *Microbial Ecology* » dont le manuscrit correspondant est présenté ici (Article A).

Bruneel, O., Volant, A., Gallien, S., Chaumande, B., Casiot, C., Carapito, C., Bardil, A., Morin, G., Brown, G. E., Jr., Personné, C. J., Le Paslier, D., Schaeffer, C., Van Dorsselaer, A., Bertin, P. N., Elbaz-Poulichet, F. and Arsène-Ploetze, F. (2011). Characterization of the active bacterial community involved in natural attenuation processes in arsenic-rich creek sediments. *Microb Ecol* **61**: 793-810.

Chapitre I – Article A

Characterization of the Active Bacterial Community Involved in Natural Attenuation Processes in Arsenic-Rich Creek Sediments

O. Bruneel, A. Volant, S. Gallien, B. Chaumande, C. Casiot, C. Carapito, A. Bardil, G. Morin, G. E. Brown Jr, C. J. Personné, D. Le Paslier, C. Schaeffer, A. Van Dorselaer, P. N. Bertin, F. Elbaz-Poulichet, F. Arsène-Ploetze

Microbiology Ecology (2011) 61:793–810 DOI 10.1007/s00248-011-9808-9

Characterization of the Active Bacterial Community Involved in Natural Attenuation Processes in Arsenic-Rich Creek Sediments

Odile Bruneel · Aurélie Volant · Sébastien Gallien · Bertrand Chaumande ·
Corinne Casiot · Christine Carapito · Amélie Bardil · Guillaume Morin ·
Gordon E. Brown Jr · Christian J. Personné · Denis Le Paslier · Christine Schaeffer ·
Alain Van Dorsselaer · Philippe N. Bertin · Françoise Elbaz-Poulichet ·
Florence Arsène-Ploetze

Received: 7 August 2010 / Accepted: 20 January 2011 / Published online: 12 February 2011
© Springer Science+Business Media, LLC 2011

Abstract Acid mine drainage of the Carnoulès mine (France) is characterized by acid waters containing high concentrations of arsenic and iron. In the first 30 m along the Reigous, a small creek draining the site, more than 38% of the dissolved arsenic was removed by co-precipitation

with Fe(III), in agreement with previous studies, which suggest a role of microbial activities in the co-precipitation of As(III) and As(V) with Fe(III) and sulfate. To investigate how this particular ecosystem functions, the bacterial community was characterized in water and sediments by

Electronic supplementary material The online version of this article (doi:10.1007/s00248-011-9808-9) contains supplementary material, which is available to authorized users.

O. Bruneel (✉) · A. Volant · C. Casiot · A. Bardil ·
C. J. Personné · F. Elbaz-Poulichet
Laboratoire HydroSciences Montpellier,
UMR5569 (CNRS-IRD-Universités Montpellier I et II),
Université Montpellier II, CC MSE,
Place Eugène Bataillon,
34095 Montpellier Cedex 05, France
e-mail: bruneel@msem.univ-montp2.fr

B. Chaumande · P. N. Bertin · F. Arsène-Ploetze
Génétique Moléculaire,
Génomique Microbiologie,
UMR7156, Université de Strasbourg/CNRS,
28 rue Goethe,
67083 Strasbourg Cedex, France

S. Gallien · C. Carapito · C. Schaeffer · A. Van Dorsselaer
Laboratoire de Spectrométrie de Masse Bio-organique,
Institut Pluridisciplinaire Hubert Curien,
UMR7178 (CNRS-Université de Strasbourg),
25 rue Becquerel,
67087 Strasbourg, France

G. Morin
Institut de Minéralogie et de Physique des Milieux Condensés
(IMPMC), UMR7590 (CNRS - Universités Paris 6&7 - IPGP),
140, rue de Lourmel,
75015 Paris, France

G. E. Brown Jr
Surface and Aqueous Geochemistry Group,
Department of Geological and Environmental Sciences,
Stanford University,
Stanford, CA 94305-2115, USA

G. E. Brown Jr
Stanford Synchrotron Radiation Laboratory, SLAC,
2575 23 Sand Hill Road, MS 69,
Menlo Park, CA 94025, USA

D. Le Paslier
Génomique Métabolique, UMR8030, CNRS,
2 rue Gaston Crémieux,
91057 Evry Cedex, France

D. Le Paslier
Commissariat à l'Energie Atomique (CEA),
Direction des Sciences du Vivant, Institut de Génomique,
Genoscope, Laboratoire de Génomique Comparative,
2 rue Gaston Crémieux,
91057 Evry Cedex, France

16S rRNA encoding gene library analysis. Based on the results obtained using a metaproteomic approach on sediments combined with high-sensitivity HPLC-chip spectrometry, several GroEL orthologs expressed by the community were characterized, and the active members of the prokaryotic community inhabiting the creek sediments were identified. Many of these bacteria are β -proteobacteria such as *Gallionella* and *Thiomonas*, but γ -proteobacteria such as *Acidithiobacillus ferrooxidans* and α -proteobacteria such as *Acidiphilium*, *Actinobacteria*, and *Firmicutes* were also detected.

Introduction

Acid Mine Drainage (AMD) is one of the most serious forms of water pollution in industrial and post-industrial areas worldwide [38]. AMD is generated when the wastes from the mining and processing of sulfide ores (such as pyrite or arsenopyrite) come into contact with oxygenated water [5]. AMD is often characterized by pH values of 2–4. Such waters generally contain high levels of iron, toxic metals (such as aluminum, manganese, lead, cadmium, and zinc), and metalloids (arsenic) [5, 32, 48]. AMD can still occur hundreds of years after mine closure and tens of thousands of kilometers of groundwater, streams, lakes, and estuaries throughout the world have been directly impacted [40]. In several cases of AMD, natural remediation has been observed, as for example at the Carnoulès site in France and the Rio Tinto site in Spain [18, 53]. In such AMD, toxic compounds are accumulated in sediments consisting of a variety of iron (oxyhydr)oxides and hydroxysulfates such as jarosite, schwertmannite, and ferrihydrite [48]. Natural remediation of metal pollutants is generally due to the occurrence of abiotic reactions and/or microbial activities that make these toxic compounds insoluble and lead them to accumulate in sediments [32, 40]. This toxic compound precipitation processes mainly involve the oxidation and precipitation of iron, which is often the main soluble metal present in AMD, and the adsorption of other metals and metalloids by the ferric minerals formed [32, 51]. Indeed, many elements such as Sr, Cs, Pb, U, and As show a strong affinity for solid iron oxide [18, 27, 48]. Abiotic oxidation of Fe(II) proceeds very slowly in acidic (pH 3.5) waters [51]. In contrast, iron-oxidizing bacteria catalyze the reaction and thus accelerate the formation of solid iron oxide [32, 41, 51]. In addition, several bacteria contribute to the immobilization of arsenic via their ability to oxidize this metalloid [6, 14, 18, 22, 48], arsenate (As(V)) being adsorbed more strongly than arsenite (As(III)) by Fe(III) oxides and hydroxides at acidic pH levels [10]. *Thiomonas* strains show a high arsenite oxidation capacity, and these metabolic activities have been extensively analyzed under laboratory conditions [6, 14, 22].

However, to be able to develop remediation processes and/or optimize existing processes, further knowledge is required about how these bacteria function in situ. In particular, it is of crucial importance to determine which bacteria are viable and active in such ecosystems.

The AMD of Carnoulès mine in Southern France is a highly suitable site for analyzing how microorganisms contribute to the transformation of metals and metalloids in situ, since efficient natural remediation processes are known to occur at this site [12, 18]. This former mine generated around 1.2 Mt of tailings containing 0.7% Pb, 10% FeS₂, and 0.2% As. Water percolating through the tailings forms the head of the Reigous creek. This creek is acidic (pH around 3) and highly contaminated with As (100 to 350 mg L⁻¹). The behavior of As and Fe in the Reigous creek has been intensively studied [18, 23, 48]. In the creek spring, As(III) is the main As species present and Fe occurs in the form of Fe(II) [18]. Along the first 30 m of the creek (about 1 h residence time), the oxidation of Fe(II) leads to the co-precipitation of more than 38% of the dissolved As [18, 23]. Arsenic accounts for up to 22% of the total dry weight of the sediments formed along the first 10 m along the creek. In the wet season, approximately 30 m downstream of the spring, these sediments are mainly composed of As(III)–Fe(III) oxyhydroxysulfates, whereas As(V)–Fe(III) oxyhydroxysulfates compounds predominate during the dry season [48]. Because of the very high molar As/Fe ratio (up to 0.3) existing in the dissolved phase of the Carnoulès creek, the mineralogical content of the sediments differs significantly from that classically observed at most AMD [47], especially along the first 50 m of the creek. Several strains of *Thiomonas* and *Acidithiobacillus ferrooxidans* have been isolated from Reigous creek waters, and based on the results of laboratory experiments, it has been suggested that these bacteria may contribute to the oxidation of Fe(II) and the co-precipitation of As [14, 18, 21–23, 48]. Preliminary analyses have shown that the DNAs of both bacteria are present in the Reigous creek, as well as that of *Gallionella* sp., *Thiobacillus* sp., and some sulfate-reducing bacteria [12]. Archaea have also been found to occur in the Reigous creek (*Ferroplasma acidiphilum* and sequences affiliated to uncultured *Thermoplasmatales* archaeon) as well as a eukaryotic microorganism, *Euglena mutabilis* [12, 13, 15]. However, the bacterial population inhabiting the As-rich Reigous sediments has never been characterized so far. It, therefore, seemed to be necessary not only to identify the bacteria present in this creek but also to determine which members of this community are viable and, therefore, perform metabolic activities in situ.

The aim of this study was to describe the bacterial populations occurring in both the sediments and waters at the disused Carnoulès site and to identify the bacteria at work. For this purpose, three complementary approaches

were used. First, chemical and mineralogical studies were performed in order to determine the arsenic species present. A 16S rRNA encoding gene library was then analyzed in order to identify the bacterial population present in the creek sediments and waters. Lastly, based on the findings obtained using a metaproteomic approach combined with high-sensitivity mass spectrometry methods, the active species inhabiting the sediments were identified.

Methods

Sampling and Analysis

Samples were collected from Reigous creek in April 2006 at COWG station located 30 m downstream from the spring. This sampling was part of a long-term monitoring of the physicochemistry of the Reigous Creek water [23]. The main physicochemical parameters (pH, temperature, and dissolved oxygen concentrations) were measured in situ at this sampling point. The 5-cm deep sediments on the bottom of the creek and a thin column (less than 10 cm) of running water covering the sediments were sampled. Solid samples were removed with a sterile spatula from the surface of the sediments. Water samples (300 ml) were immediately filtered through 0.22 μm Millipore membranes fitted on Sartorius polycarbonate filter holders (for water chemical analysis) or through sterile 0.22- μm Nucleopore filters that were then transferred to a collection tube (Nunc), frozen in liquid nitrogen, and stored at -80°C until DNA extraction (for 16S rRNA encoding gene analysis). Sampling was repeated three times. For total Fe and As determination, filtered water was acidified to pH=1 with HNO_3 (14.5 M) and stored at 4°C in polyethylene bottles until analysis. For As and Fe speciation, a 10 μl aliquot of filtered sample water was added to either 0.5 ml of 5% (v/v) 0.25 M EDTA solution for As speciation [7] or a mixture of 0.5 ml acetate buffer (pH 4.5) and 1 ml of 1,10-phenanthroline chloride solution for Fe speciation [50]. The vials were completed to 10 ml with deionized water. The samples used for arsenic speciation and Fe(II) and sulfate determination were stored in the dark and analyzed within 24 h.

Chemical Analysis

The determination of total dissolved As was performed by ICP-MS using Thermo X7 series with a conventional external calibration procedure. Indium was used as internal standard to correct for instrumental drift and possible matrix effects. It was not necessary to correct interference with chloride because of the extremely high As levels present. Certified reference material SLRS-4 (freshwater

samples) was used to check analytical accuracy and precision. The results showed that the recovery rate obtained was within $\pm 5\%$.

Analyses of inorganic arsenic species (As(III), As(V)) were carried out using anion-exchange chromatography (25 cm \times 4.1 mm i.d. Hamilton PRP-X100 column with Varian ProStar gradient solvent delivery system) coupled to a hydride generation (VGS 200, FISON, France) with an atomic fluorescence spectrometry detector (Excalibur, PSAnalytical, GB) [17]. The detection limit obtained was 172 ng L^{-1} for As(III) and 458 ng L^{-1} for As(V), with a precision better than 5%. Total dissolved Fe was determined by flame atomic absorption spectrometry. Fe(II) concentration was determined using colorimetry at 510 nm after complexation with 1,10-phenanthroline chloride solution in buffered samples (pH 4.5) [50] (detection limit: 11 $\mu\text{g L}^{-1}$; precision better than 5%). Sulfate concentration was determined after precipitation of BaSO_4 with BaCl_2 and spectrophotometric measurement at 650 nm [50].

Solid Sample Characterization

XAFS data were gathered on the laboratory samples and the sample taken at COWG on April 2006, at 10 K in transmission mode on a bending magnet D44 at the LURE synchrotron (Orsay, France), and in fluorescence mode on the 11-2 wiggler beamline at SSRL (Stanford, CA), respectively. Experiments and data reduction were previously reported [48, 49].

DNA Isolation, 16S rRNA Encoding Gene Cloning, Restriction Analysis, and Sequencing

Genomic DNA was extracted in triplicate from filtered water and sediments using the UltraClean Soil DNA Isolation Kit according to the manufacturer's recommendations (MoBio Laboratories Inc., USA). These triplicates were pooled before PCR amplification. All the genomic DNA samples extracted were stored at -20°C until further processing. Bacterial diversity was analyzed by cloning PCR amplified 16S rRNA encoding genes. Bacterial 16S rRNA encoding genes were amplified with 8F (5'-AGAGTTTGA TCCTGGCTCAG-3') and 1489R primers (5'-TACCTTGT TACGACTTCA-3') [43, 57], as previously described [12]. These PCR products were cloned into *Escherichia coli* TOP 10 strain using the pCR2.1 Topo TA cloning kit (Invitrogen, Inc.). Cloned 16S rRNA encoding gene fragments were re-amplified using the primers TOP1 (5'-GTGTGCT GGAATTCGCCCTT-3') and TOP2 (5'-TATCTGCAGA ATTCGCCCTT-3') that anneal to the vector and surround the inserted PCR fragment and then digested with *Hae*III or *Hinf*I enzymes. Restriction profiles were analyzed using 2.5% agarose gel electrophoresis (small fragment resolution

agarose; QA agarose, QBiogène, Inc.). Around 200 clones from each library were analyzed and grouped according to the RFLP patterns obtained. PAST (Paleontological STatistics v 1.19) software from the website <http://folk.uio.no/ohammer/past/> was used to calculate different diversity indices (rarefaction analysis, taxa, total clones, singletons, dominance, coverage, shannon, equitability, and simpson, Table 1) for each clone library. The total number of clones obtained compared with the number of clones representing each unique phylotype was used to produce the rarefaction curves. Coverage values were calculated to determine how efficiently the libraries described the complexity of a theoretical community such as an original bacterial community. The coverage [29] value is given as $C = 1 - (n_1/N)$ where n_1 is the number of clones that occurred only once in the library. Rarefaction analysis showed that the curves generated were near saturation (data not shown) and consistent with the high coverage values of the two clone libraries (97.8 for the sediment and 98.6 for the water). This indicated that the clone libraries were sufficiently sampled.

Partial sequences of the clones from dominant groups were determined by the dideoxy nucleotide chain-termination method using the BigDye 3.1 kit (Applied Biosystems) on an ABI PRISM 3730XL Genetic analyzer (Applied Biosystems). The PINTAIL program [4] was used to check the presence of chimerae. Sequences were also examined manually for chimerae, which were excluded from further analyses. These sequences were compared with known sequences (NCBI database) using the BLAST, ALIGN, and CLUSTALW programs [1, 26, 55]. All sequence data obtained were submitted to the EMBL databases under accession numbers (FR676963-FR677013; AM988784-AM988794; AM988796; AM988798; AM988801-AM988805; AM988807-AM988809).

Preparation of Protein Extracts and Gel Electrophoresis Analysis

Prokaryotes were separated from sediments and eukaryotes using a Nycodenz density gradient. It should be noted that the main population determined from the DNA directly extracted from sediments was similar to that identified after Nycodenz treatment (data not shown), which suggests that this treatment did not result in enrichment of the sample by any particular microorganisms. Ten grams of sediments

were washed in 10 mL of saline buffer and agitated overnight at 4°C. After 10 min of decantation, 7.5 mL of supernatant were added without mixing to 17.5 mL of Nycodenz solution (Axis-Shield, Dundee, Scotland), and then centrifuged for 30 min at 10,000×g. The cellular fraction (nycodenz/sample interface) was removed and washed by adding two volumes of NaCl 0.9% and centrifuged for 15 min at 10,000×g at 4°C. Proteins were extracted from this cellular fraction as previously described [58], further purified using the 2-D Clean-up kit (GE Healthcare), and resuspended in rehydration buffer (364 g L⁻¹ thiourea, 1,000 g L⁻¹ urea; 25 g L⁻¹ CHAPS, 0.6% v/v IPG buffer Pharmalyte, 10 g L⁻¹ of DTT, and 0.01% bromophenol blue). Protein concentrations were quantified using the 2-D Quant kit (GE Healthcare). These proteins were separated by 2-D gel electrophoresis as previously described [58] and finally stained with silver nitrate. Gels were analyzed using an Image Scanner, LabScan v 3.0 (GE Healthcare), and the ImageMaster 2D platinum software program (v. 6.01, GE Healthcare). The spots selected were cut out of the 2-D gels and stored at -20°C. Eighty-one spots were analyzed by performing nanoLC-Chip-MS/MS.

In-Gel Digestion, Mass Spectrometry Analysis, and Protein Identification

Unless otherwise specified, all chemicals were obtained from Sigma (St. Louis, MO, USA). In-gel digestion of gel spots was performed as previously described [58]. The resulting peptides were analyzed by performing nanoLC-MS/MS on an Agilent 1100 Series HPLC-Chip/MS system (Agilent Technologies, Palo Alto, USA) coupled to an HCT Ultra ion trap (Bruker Daltonics, Bremen, Germany). The MS/MS data were analyzed using the MASCOT 2.2.0 algorithm giving a maximum of one missed cleavage, with a mass tolerance of 0.5 Da for MS and MS/MS data and carbamidomethylation of cysteines and oxidation of methionines were specified as the variable modifications. MS/MS data searches were performed against two in-house generated databases. The first database was composed of the protein sequences of all the organisms related to the groups identified by performing 16S rRNA encoding gene analysis on the Reigous creek sediments and water (α -, β -, δ -, and γ -*proteobacteria*, *Bacilli*, *Clostridia*, *Actinobacteria*,

Table 1 Diversity indices calculated from the two clone libraries in sediment and water at the Carnoulès mine drainage creek

Clone library	Taxa	Total no. of clones	Singletons	Dominance (D)	Coverage (C)	Shannon (H)	Equitability	Simpson (1-D)
Sediments	15	229	5	0.5503	97.8	1.146	0.4231	0.4497
Water	18	221	3	0.2551	98.6	1.949	0.6744	0.7449

Nitrospira), as well as unclassified bacteria from <http://beta.uniprot.org/>, *Thiomonas* sp. from <http://www.genoscope.cns.fr/> (FP475956–FP475957), *Euglenozoa*, and *Viridiplantae*. The second database included all bacterial and archaeal GroEL chaperonins (12501 and 291 sequences, respectively) extracted from the Uniprot database (<http://www.uniprot.org/uniprot>).

To assess the false positive rate in the protein identification, a target-decoy database search was performed [25]. With this approach, peptides are matched against a database consisting of the native protein sequences detected in the database (target) and the sequence-reversed entries (decoy). Protein identification was confirmed when at least two peptides with a minimum Mascot ion score of 30 were detected. In the case of one-peptide hits, the score of the unique peptide had to be greater than the 95% significance Mascot threshold level. All the proteins identified were added to the “InPact” proteomic database developed at our laboratory (<http://inpact.u-strasbg.fr/~db/>) [8].

Phylogenetic Analyses

A search for GroEL homologs and 16S rRNA encoding sequences was carried out in the Uniprot and NCBI databases, respectively. A total number of 530 reviewed GroEL bacterial sequences were retrieved from the RefSeq database. These sequences were 500–550 amino acids in length. Only one sequence representative of each genus (259 sequences in all) was kept. GroEL and 16S rRNA encoding sequences were aligned using ClustalW [55]. Alignments were checked by hand and positions with more than 1% of gaps were automatically removed. Neighbor-joining trees were constructed with 185 amino acids in the case of the GroEL sequences and with 310 nt in that of the 16S rRNA encoding sequences. Trees were drawn up using the iTOL website (<http://itol.embl.de/>) [44].

Results and Discussion

Physical and Chemical Characteristics of Samples

The physicochemistry of the Reigous Creek water at the time of sampling was typical of that revealed during the long-term monitoring study [23]. The water sample was acid (pH=3.28) and moderately oxygenated (dissolved oxygen concentration=3.5±0.5 mg L⁻¹); it contained extremely high concentrations of SO₄²⁻ (2700±300 mg L⁻¹), Fe (620±30 mg L⁻¹), and As (140±4 mg L⁻¹), with a large predominance of Fe(II) (90±10% of total Fe concentration) and equal proportion of As(III) and As(V). The removal of As during the course of the Reigous Creek from its source to the sampling station

COWG reached 38%, corroborating the typical removal rates measured during the long-term monitoring study [23]. The pale-yellow loosely packed sediments previously observed during the dry season at the sampling point chosen in this study (COWG ~30 m downstream of the spring), consisted of an amorphous Fe(III)–As(V) hydroxysulfate mineral with an As/Fe molar ratio of 0.5 to 0.6 [48]. However, various other biominerals may be formed from dissolved Fe(II) and As(III) [47, 48]. The exact nature and structure of the sediment sample studied was, therefore, further investigated. X-ray powder diffraction (data not shown) and X-ray absorption spectroscopy data obtained at the As–K edge (Fig. 1) showed that arsenic was present in these samples in an amorphous Fe(III)–As(V) hydroxysulfate phase as previously observed [48]. These analyses showed that despite the presence of a minor As(III) impurity, the oxidized arsenic form As(V) predominates in this sediment (Fig. 1a). In a previous study, it has been shown that the catalytic oxidation of As(III) by *Thiomonas* sp. strains accelerates such As–Fe precipitation process [48]. Therefore, using extended X-ray absorption fine structure analysis at the As–K edge (Fig. 1b), the structure observed in our samples was compared with that of the minerals obtained after As(III) oxidation by the *Thiomonas* sp. strain B2 in bioassays in which sterilized Carnoulès Creek water was used [48]. These comparisons (Fig. 1b, c) showed that the molecular structure of the amorphous Fe(III)–As(V) hydroxysulfate phase observed in these sediments was similar to that of the Fe(III)–As(V) hydroxysulfate obtained in the presence of *Thiomonas*, further supporting its role in situ.

Composition of Bacterial Communities in Reigous Creek Water and Sediments

Two 16S rRNA encoding gene libraries were constructed (Table 2), containing 229 clones in the sediment library and 221 in that of the water library. The Shannon index (see “Methods”) and equitability values were greater in the water library than in the sediment library (Table 1), which suggests that the bacterial diversity was lower in the sediment than in the water samples. Eleven different species were identified in the sediments and 13 in the water (Table 2, Fig. 2).

Several of the bacteria identified in the present study in both the waters and the sediments have been previously detected in the Reigous waters [12, 14, 15, 18, 21, 22]. Classified first by abundance order, several sequences were affiliated to *Thiobacillus* sp. ML2-16. This bacterium has been frequently reported to occur in AMD [5]. The presence of several strains affiliated to this *proteobacteria* is in agreement with the results of previous studies and shows that these bacteria persist in this

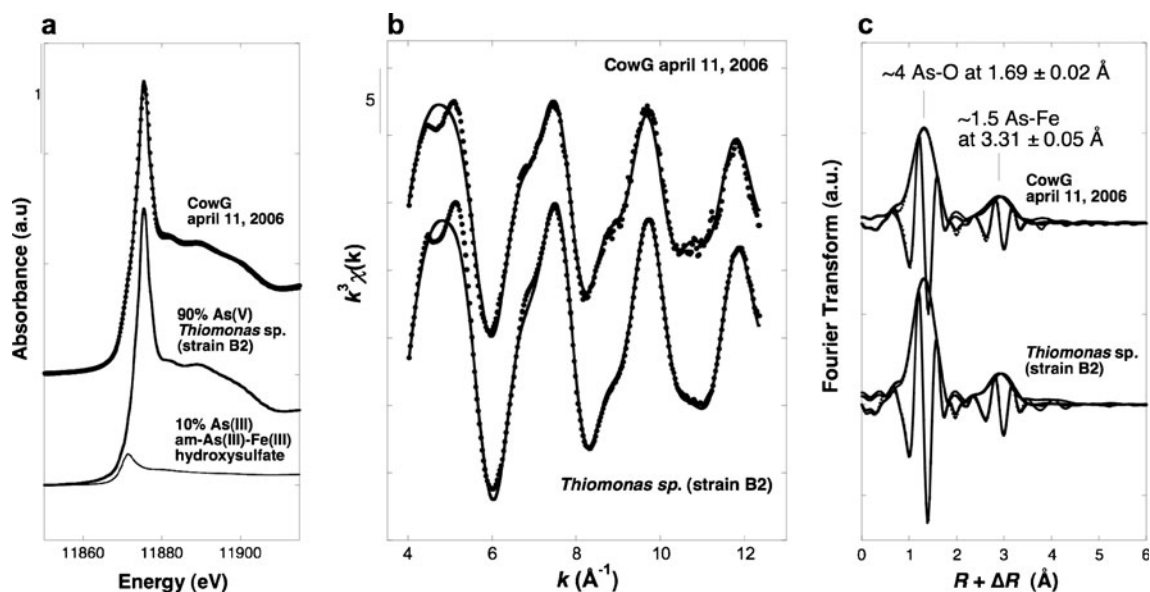


Figure 1 X-ray absorption spectra at the As-K edge of the sample, (COWG April 11, 2006,) showing similarities with the X-ray amorphous Fe(III)-As(V) hydroxysulfate phases obtained after incubating sterilized Carnoulès water with the *Thiomonas* sp. strain B2 isolated at the site [48]. These phases reached a molar As/Fe ratio of 0.8, as described in [48]. **a** Linear least-squares fitting of XANES data showed that the largest arsenic fraction ($90\pm 2\%$) was in the As(V) oxidation state. A small arsenic fraction ($10\pm 2\%$) was in the As(III) oxidation state, which resulted in a slight decrease in the amplitude in the EXAFS spectrum. **b** Shell by shell fit of the EXAFS spectra in k -

space for the COWG sediment sample and the *Thiomonas* sp. strain B2 precipitate sample. **c** Corresponding Fourier transforms of the experimental and fit curves. Dotted lines experimental; solid lines fitting curves. The local structure of both the COWG and the laboratory *Thiomonas* sp. samples includes bidentate arsenate-oxygen-iron complexes characterized by ~ 1.5 – 2.0 Fe atoms at an As-Fe distance of 3.31 ± 0.02 Å. A small arsenic fraction ($10\pm 2\%$) was in the As(III) oxidation state, which resulted in a slight decrease in the amplitude in **b** the EXAFS spectrum and in **c** the corresponding Fourier transform

ecosystem [12]. Secondly, bacteria affiliated to *Gallionella capsiferiformans* were detected in both the water and sediments. *G. capsiferiformans* is an oxygen-dependent ferrous iron-oxidizing bacterium that grows at circumneutral pH [59]. Relatives of *Gallionella*, such as *G. ferruginea* in particular, have often been detected, sometimes as the dominant group in microbial mine water communities, including Carnoulès [12, 33–35]. Bacteria related to *At. ferrooxidans* as well as *Thiomonas* strains have been previously isolated from this site [14, 18, 21, 22]. *At. ferrooxidans*, which was the first microorganism to be isolated from an acidic leaching environment, occurs ubiquitously in AMD, as does *Thiomonas* [32]. In addition to these groups, other species that have not previously been described at this site were detected in this study. Some of these species were found to occur in both sediments and water. For example, bacteria affiliated to the *Firmicutes Alicyclobacillus* sp. BRG 73 were identified. This genus, found in AMD [5], is characterized by moderately thermophilic, acidophilic, strictly aerobic, and endospore-forming bacilli [30]. Likewise, bacteria related to “*Ferrovum myxofaciens*” PSTR were detected in both the waters and sediments. “*Ferrovum myxofaciens*” is an autotrophic iron-oxidizer which predominates in some AMD and is able to grow litho-autotrophically, using ferrous iron as an electron

donor [32]. Lastly, bacteria affiliated to *Leptospirillum ferrooxidans*, an iron-oxidizing member of the *Nitrospirae* [36], were also present. This bacterium has been found to occur in several acidic environments and in biofilms originating from AMD [5, 9, 32].

Other newly characterized groups were identified only in sediments (Table 2). The presence in acidic mine waters of *Acidocella* sp., a non-iron-oxidizing heterotrophic acidophile is quite common in AMD [33, 35]. *Ferrimicrobium* is an iron-oxidizing heterotroph that can also use iron as an electron acceptor [20]. Other prokaryotes detected in sediments were affiliated to *Acidiphilium* sp. CCP3, a non-iron-oxidizing heterotrophic acidophile that is also quite common in AMD [33] and *Dokdonella koreensis*, a γ -proteobacteria.

Six newly characterized groups were identified only in water. Some of these bacteria were related to *Sideroxydans lithotrophicus* LD-1, an oxygen-dependent ferrous iron-oxidizing bacterium that grows at circumneutral pH [59]; *Rhodoferrax ferrireducens*, a psychrotolerant, facultative anaerobic bacterium which is able to oxidize acetate with the reduction of Fe(III) [28]; and an *Acidobacteriaceae* bacterium, CH1. Members of *Acidobacteria* have previously been reported in AMD [5, 32]. Finally, three sulfate-reducing bacteria were related to strain JHA1, *Desulfomonile limimaris*,

Table 2 Bacterial clones detected at the Carnoulès mine drainage with their phylogenetic group, the closest isolated relative and the relative abundance of each group versus the total number of clones (100%)

Sampling	Clones	Phylogenetic group	Closest isolated relative (accession number)	Percentage of similarity	Relative abundance of clones (%)
Sediment	CGA6Sd1a, 5a, 10b, 13c, 23c, 34c, 59b, 89c, 92c	β -Proteobacteria	<i>Thiobacillus</i> sp. ML2-16 (DQ145970)	95–96	31
	CGA6Sd4b, 36c, 48c	β -Proteobacteria	<i>Gallionella capsiferriformans</i> ES-2 (DQ386262)	96	21
	CGA6Sd13a	γ -Proteobacteria	<i>Acidithiobacillus ferrooxidans</i> DX-1 (EU084695)	99	10
	CGA6Sd10a, 37a	α -Proteobacteria	<i>Acidocella</i> sp. M21 (AY765998)	99–100	8
	CGA6Sd31c, 36a	β -Proteobacteria	<i>Thiomonas</i> sp. PK44 (AY455806)	99	6
	CGA6Sd6b, 18a, 27b	γ -Proteobacteria	<i>Dokdonella korensis</i> NML_01–0233 (EF589679)	92	5
	CGA6Sd32a	Actinobacteria	<i>Ferrinicrobium</i> sp. BGR 49 (GU167992)	99	3
	CGA6Sd38c	α -Proteobacteria	<i>Acidiphilium</i> sp. CCP3 (AY766000)	99	3
	CGA6Sd20a, 76c	Firmicutes	<i>Alicyclobacillus</i> sp. BGR 73 (GU167996)	92–99	3
	CGA6Sd58b	Nitrospirae	<i>Leptospirillum ferrooxidans</i> Sy (AF356839)	99	2
	CGA6Sd51c	β -Proteobacteria	" <i>Ferrofum</i> <i>myxofaciens</i> " PSTR (EF133508)	100	2
	CGA6Wt4c, 7a, 15c, 20a, 21b, 22b, 23a, 31b, 32a, 33c, 45b, 54c, 63b, 73c	β -Proteobacteria	<i>Thiobacillus</i> sp. ML2-16 (DQ145970)	94–96	26
	CGA6Wt7c, 17b, 21a, 23c, 27b, 35b, 36b, 67c, 80c, 86c	β -Proteobacteria	<i>Gallionella capsiferriformans</i> ES-2 (DQ386262)	89–97	18
	CGA6Wt25c, 56b	γ -Proteobacteria	<i>Acidithiobacillus ferrooxidans</i> BGR:110 (GU168011)	100	10
	CGA6Wt5a, 48c, 79b	Firmicutes	<i>Acidithiobacillus ferrooxidans</i> DSM 2392 (AJ459800)	91–92	
	CGA6Wt9a, 19a, 27a, 9b, 61c	β -Proteobacteria	<i>Alicyclobacillus</i> sp. BGR 73 (GU167996)	91–99	8
CGA6Wt11a, 29c, 78c	β -Proteobacteria	<i>Sideroxydans lithotrophicus</i> LD-1 (DQ386859)	94–97	8	
CGA6Wt3a, 8c	β -Proteobacteria	" <i>Ferrofum</i> <i>myxofaciens</i> " PSTR (EF133508)	98–99	4	
CGA6Wt86b	β -Proteobacteria	<i>Thiomonas</i> sp. PK44 (AY455806)	94	3	
CGA6Wt15a	Nitrospirae	<i>Leptospirillum ferrooxidans</i> (AB510912)	94	3	
CGA6Wt51b	β -Proteobacteria	<i>Rhodoferrax ferrireducens</i> T118 (CP000267)	99	3	
CGA6Wt61b	Actinobacteria	<i>Acidobacteriaceae</i> bacterium CHI (DQ355184)	96	3	
CGA6Wt42c		Sulfate-reducing bacterium JHAI (EF442984)	82	3	
CGA6Wt30a	δ -Proteobacteria	<i>Desulfomonile limimaris</i> (NR_025079)	87	3	
CGA6Wt10a	δ -Proteobacteria	<i>Desulfuromonas svalbardensis</i> 60 (AY835390)	82	1	

Sequences closely related to 16S rRNA genes from *Euglena* spp. chloroplast were also detected (data not shown). The 16S rRNA encoding gene of chloroplasts is closely related to the bacterial 16S rRNA encoding gene and can therefore be amplified by primers 8F and 1489R

an anaerobic dehalogenating bacterium from marine sediments [54] and *Desulfuromonas svalbardensis* 60, a psychrophilic, Fe(III)-reducing bacterium isolated from Arctic sediments [56] (Table 2).

All in all, 17 species of bacteria were identified in the water and sediments sampled at the Reigous creek. Only seven genera were found to be present in both phases, six were found only in water, and four only in the sediments (Table 2). Most of these species are common residents of AMD [5, 32]. This quite low bacterial diversity was probably due to the high concentration of toxic compounds in this AMD and was consistent with previous observations showing that the biodiversity of acidic, metal-rich mine waters is mainly restricted to specialized prokaryotes and some eukaryotes such as *Euglena* [15, 52], which has been detected in this study (data not shown). In the Reigous system, in both sediments and water, the populations observed were mainly involved in the Fe, As, and S cycles. The populations involved in Fe(II) oxidation were related to *Gallionella*, *At. ferrooxidans*, *Ferrimicrobium*, *Leptospirillum*, *Sideroxydans lithotrophicus*, or “*Ferrovum myxofaciens*” [32, 59], whereas ferric iron reduction has been described for populations like *Acidiphilium* spp., *Acidocella*, *Desulfuromonas svalbardensis*, *Rhodoferax ferrireducens*, or even *At. ferrooxidans* and *Ferrimicrobium acidiphilum* [20, 28, 56]. Since the Carnoulès creek spring contains mainly Fe(II) and As(III) in the form of dissolved species [18], the Fe(III) and As(V) may be formed as the result of microbial oxidation processes via the activity of acidophilic iron- and arsenite-oxidizing bacteria [24, 48]. In other words, the large amounts of soft pale-yellow As(V)–Fe(III) hydroxysulfate sediments analyzed here (Fig. 1) were probably formed by the joint activities of iron-oxidizing (e.g., *At. ferrooxidans* or *Gallionella*) and arsenic-oxidizing (e.g., *Thiomonas* sp.) microorganisms. Concerning S cycling, we found populations able to oxidize the reduced inorganic sulfur compounds, like *Thiobacillus*, *Thiomonas*, or *At. ferrooxidans* [32, 42, 46]. Sulfate-reducing bacteria such as *Desulfomonile limimaris* or *Desulfuromonas svalbardensis* were also found to be present in sediments and may be involved in sulfate consumption [54, 56].

Among these prokaryotes, some bacteria may be present but not functionally active, and it was, therefore, crucial to differentiate between dead or inactive cells and functional cells. To determine which organisms play a significant role in the natural remediation processes such as the Fe(II) oxidation processes observed at the study site, a metaproteomic approach was used to list the bacterial population expressing proteins, i.e., those which were active. The metaproteomic approach was possible with sediments but failed with the water samples because larger numbers of bacterial cells were recovered from sediments than from water (data not shown).

Figure 2 Phylogenetic tree based on 16S rRNA-encoding sequences. Sequences were aligned using ClustalW. Alignments were checked by hand and positions with more than 1% of gaps were automatically removed. Neighbor-joining trees were drawn up with 310 nt using ITOL (<http://itol.embl.de/>) [44]. Accession numbers: see supplementary data. In red: bacteria identified based on the metaproteomic (GroEL identifications) approach; in blue: bacteria identified using the 16S rRNA encoding gene library; in black: sequences detected in NCBI databases which are closely related to the bacteria present in the Reigous sediment

Characterization of the Main Proteins Expressed by the Sediment Community

Proteins expressed by this community were liable to correspond to orthologs originating from diverse prokaryotes and to have similar amino acid sequences. For those reasons and to improve their characterization, proteins were separated by performing 2-D gel electrophoresis (supplementary Fig. 1). A total number of 89 proteins were identified, 44% of which (39 proteins, Table 3) originated from bacteria and 39% (35 proteins) from protists, while 15 proteins originated from higher plants, probably from decomposed plant debris present at the Reigous creek. One third of the proteins originated from protists, mainly consisting of *E. mutabilis* detected in our samples (data not shown) and could not be completely removed using the Nycodenz Gradient. *E. mutabilis* is a common inhabitant of AMD [2, 11, 15, 39] and was also present at the surface of sediments. These *Euglena* proteins are involved in various metabolic processes, suggesting that this eukaryote may play a relevant role in this ecosystem, in agreement with recent results (Gouhlen-Chollet and Bertin, unpublished data). Only proteins originating from bacteria were further analyzed in this study. The 2-D gel electrophoresis approach used here allowed identifying from a single sample only the most abundant cytoplasmic proteins. Therefore, some relevant proteins that might be expressed in this environment, such as rusticyanin from *At. ferrooxidans*, which is known to be involved in Fe(II) oxidation and was previously thought to possibly play a functional role at this site [21, 24, 48], were not detected in this study. This membrane protein may not be resolved in the 2-D gel, or it may not be abundant at the sampling point. In addition, the majority of the bacteria forming the Carnoulès community have never been grown and studied in vitro so far. Their genome sequences, and hence their protein sequences, which are required for MS identification purposes, may not be available in public databases, except for those of *At. ferrooxidans*, *Thiomonas*, and *Gallionella* strains. All these hypotheses may explain why so few bacterial proteins were identified in this study.

The bacterial proteins identified originated from *Aquificae* were *Actinobacteria*, *Deinococcus*, *Synergistetes*, *Firmicutes*, *Bacteroidetes*, and α -, β -, and γ -*proteobacteria*. Among the

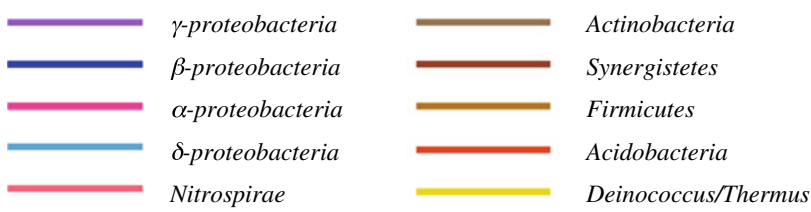
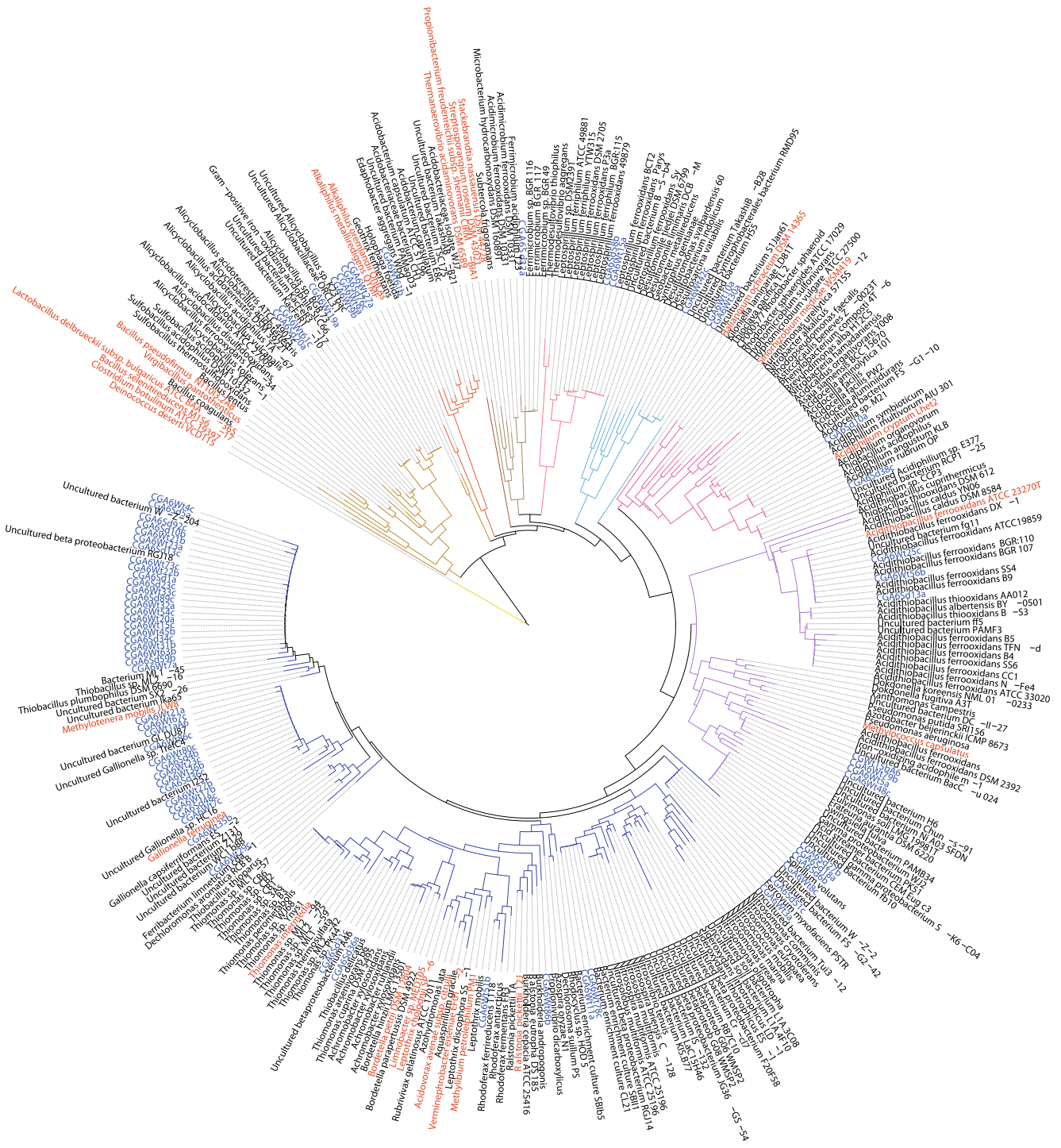


Table 3 Bacterial proteins identified in the Reigous sediment's microbial community

Phylum	Class, Family, Genus	Organism	Level of discrimination	Protein name	Protein accession numbers	Spot number ^b	Peptide sequence
<i>Aquificae</i>	<i>Aquificae</i> (class); <i>Aquificales</i> ; <i>Aquificaceae</i> ; <i>Hydrogenobaculum</i>	<i>Hydrogenobaculum</i> sp. Y04AAS1	Species	Putative uncharacterized protein	A7WFJ8	2,5	IGAAVIGR
<i>Deinococcus-Thermus</i>	<i>Deinococci</i> ; <i>Deinococcales</i> ; <i>Deinococcaceae</i> ; <i>Deinococcus</i>	<i>Deinococcus deserti</i>	Species	60 kDa chaperonins ^a	C1CZP1	4, 6	AFGFGDR QLVFDEAAR AVLVAIEEK
<i>Synergistetes</i>	<i>Deinococci</i> ; <i>Deinococcales</i> ; <i>Deinococcaceae</i> ; <i>Deinococcus</i>	<i>Deinococcus</i> sp. A62	Species	60 kDa chaperonins ^a	A1YUK7	2	IAQVASISANDK
<i>Synergistetes</i>	<i>Synergistia</i> ; <i>Synergistales</i> ; <i>Synergistaceae</i> ; <i>Thermanaerovibrio</i>	<i>Thermanaerovibrio</i> <i>acidaminovorans</i>	Genus or below	60 kDa chaperonins ^a	D1B621	4	FGSPTITNDGVTTIAK
<i>Firmicutes</i>	<i>Bacilli</i> ; <i>Lactobacillales</i> ; <i>Lactobacillaceae</i>	<i>Lactobacillus</i> subsp. <i>delbrueckii</i> subsp. <i>Indicus</i> or <i>delbrueckii</i>	Subspecies	60 kDa chaperonins ^a	Q70BV2, Q70BV5	4, 5	AFGFGDR YGAPTITNDGVTTIAK
	<i>Bacilli</i> ; <i>Bacillales</i> ; <i>Bacillaceae</i>	<i>Virgibacillus</i> <i>pantothenticus</i>	Family or below	60 kDa chaperonins ^a	A9LHR1	4, 5, 6	AVEVAVK
	<i>Bacilli</i> ; <i>Bacillales</i> ; <i>Bacillaceae</i> ; <i>Bacillus</i>	<i>Bacillus</i> <i>holodurans</i> , <i>pseudofirmus</i> or <i>selenitireducens</i>	Species	60 kDa chaperonins ^a	O50305, A8VUQ6, D3FSF9	2	NVTSGANPMVIR
	<i>Clostridia</i> ; <i>Clostridiales</i> ; <i>Clostridiaceae</i> ; <i>Alkaliphilus</i>	<i>Alkaliphilus</i> <i>metallivredigens</i> or <i>orlandii</i>	Genus or below	60 kDa chaperonins ^a	A6TLJ1, A8HJ57	6	LSGGVAVIQVGAATETELK
	<i>Clostridia</i> ; <i>Clostridiales</i> ; <i>Clostridiaceae</i> ; <i>Clostridium</i>	<i>Clostridium botulinum</i>	Species	60 kDa chaperonins ^a	A5T723, A7FYP3, A7GIN3, B1IFD4, B1LIK0, C1FLV5, C3KUC8, B1Q9U6, B1Q157 B6FW06	4	AFGFGDR LGIDIIIR
	<i>Clostridia</i> ; <i>Clostridiales</i> ; <i>Clostridiaceae</i> ; <i>Clostridium</i>	<i>Clostridium hiranonis</i>	Species	60 kDa chaperonins ^a		4, 5, 6	AFGFGDR KALEEPLR VGAATEVEMK
	<i>Clostridia</i> ; <i>Clostridiales</i> ; <i>Clostridiaceae</i> ; <i>Clostridium</i>	<i>Clostridium</i> <i>papyrosolvans</i>	Species or subspecies	60 kDa chaperonins ^a	C7IKN8	4, 6	TNDIAGDGTITATVLAQAIR AFGFGDR
<i>Bacteroidetes</i>	<i>Flavobacteria</i> ; <i>Flavobacteriales</i> ; <i>Flavobacteriaceae</i> ; <i>Gramella</i>	<i>Gramella forsetii</i> (strain KT0803)	Species	ATP synthase; beta subunit Phosphoglycerate kinase	A0M791 A0M6J2 A5JUG8	11 13, 20 4, 5, 6	FGSPTITNDGVTTIAK MPSAVGYQPTLATMGMAMQER LGDIVVNDAFGTAHR NVTAGANPIELK
<i>Actinobacteria</i>	<i>Actinobacteria</i> (class); <i>Actinobacteridae</i> ; <i>Actinomycetales</i> ; <i>Propionibacterineae</i> ; <i>Propionibacteriaceae</i>	<i>Propionibacterium</i> <i>freudenreichii</i>	Species	60 kDa chaperonins ^a		4, 6	AFGFGDR
	<i>Actinobacteria</i> (class); <i>Actinobacteridae</i> ; <i>Actinomycetales</i> ; <i>Glycomycineae</i> ; <i>Glycomycetaceae</i> ; <i>Stackebrandia</i>	<i>Stackebrandia</i> <i>nassauensis</i>	Suborder or below	60 kDa chaperonins ^a	C4DUC7	6	GMNALADAVK
	<i>Actinobacteria</i> (class); <i>Actinobacteridae</i> ; <i>Actinomycetales</i> ; <i>Streptosporangineae</i> ; <i>Streptosporangiaceae</i> ; <i>Streptosporangium</i>	<i>Streptosporangium</i> <i>roseum</i>	Family	60 kDa chaperonins ^a	D2BBD1	4	AFGFGDR GTFTSVAVK
	<i>Actinobacteria</i> (class); <i>Actinobacteridae</i> ; <i>Actinomycetales</i>	<i>Aribrobacter</i> , <i>Janibacter</i> , <i>Clavibacter</i> or <i>Kineococcus</i>	Order	ATP synthase; beta subunit	A0JY64, A1R7V3, A3TGD9, A5CQ60, A6W7G9	10, 31, 32, 33, 38	DVQNQDVLFLIDNIFR VALSALTMAYEYFR IGLFGGAGVVK
<i>Proteobacteria</i>	<i>α</i> - <i>proteobacteria</i> ; <i>Rhizobiales</i> ;	<i>Sinorhizobium medicae</i>	Genus or	60 kDa chaperonins ^a	A6UH06	2, 4	LVAAGMNPMDLK

<i>Rhizobiaceae; Sinorhizobium/Ensifer</i> group; <i>Sinorhizobium</i>									AAVEEGIVAGGGVALLR
<i>α-proteobacteria, Rhodospirillales; Acetobacteraceae; Acidiphilium</i>	<i>Acidiphilium cryptum</i>	below	60 kDa chaperonins ^a	A5G1G2		Genus or below		5, 6	APFGDR AAVEEGIVPGGVALAR AVAAAGMNPMDLK AGIIDPTK ENTTIVEGAGK APFGDR EIELADPFENMGAQLVK SVAAGMNPMDLK APFGDR VGAATEVEMK AVTALVAELKK VTLADLGQAK AAVEEGIVAGGGVALLR APFGDR AVEEPLR VGAATEVEMK EGVITVEDGK VQIEEATSDYDREK DLLPVLEQVAK VEDALHATR VQIEEATSDYDR AEILDAIAGMTVLELSELIK DLVDGAPKPKV IIAIDTNPAK TNLCVAVR
<i>α-proteobacteria, Rhodobacterales; unclassified Rhodobacterales</i>	Rhodobacterales bacterium HTCC2083	Family or below	60 kDa chaperonins ^a	B6AWC8				4, 5	ENTTIVEGAGK APFGDR EIELADPFENMGAQLVK SVAAGMNPMDLK APFGDR VGAATEVEMK AVTALVAELKK VTLADLGQAK AAVEEGIVAGGGVALLR APFGDR AVEEPLR VGAATEVEMK EGVITVEDGK VQIEEATSDYDREK DLLPVLEQVAK VEDALHATR VQIEEATSDYDR AEILDAIAGMTVLELSELIK DLVDGAPKPKV IIAIDTNPAK TNLCVAVR
<i>β-proteobacteria, Burkholderiales; Comamonadaceae; Acidovorax</i>	<i>Acidovorax avenae</i>	Species	60 kDa chaperonins ^a	A1TKQ5, D1STJ1				4	ENTTIVEGAGK APFGDR EIELADPFENMGAQLVK SVAAGMNPMDLK APFGDR VGAATEVEMK AVTALVAELKK VTLADLGQAK AAVEEGIVAGGGVALLR APFGDR AVEEPLR VGAATEVEMK EGVITVEDGK VQIEEATSDYDREK DLLPVLEQVAK VEDALHATR VQIEEATSDYDR AEILDAIAGMTVLELSELIK DLVDGAPKPKV IIAIDTNPAK TNLCVAVR
<i>β-proteobacteria, Burkholderiales; Alcaligenaceae; Bordetella</i>	<i>Bordetella petrii</i>	Species	60 kDa chaperonins ^a	A9I685				2, 4, 5	ENTTIVEGAGK APFGDR EIELADPFENMGAQLVK SVAAGMNPMDLK APFGDR VGAATEVEMK AVTALVAELKK VTLADLGQAK AAVEEGIVAGGGVALLR APFGDR AVEEPLR VGAATEVEMK EGVITVEDGK VQIEEATSDYDREK DLLPVLEQVAK VEDALHATR VQIEEATSDYDR AEILDAIAGMTVLELSELIK DLVDGAPKPKV IIAIDTNPAK TNLCVAVR
<i>β-proteobacteria</i>	<i>Bordetella avium</i> (strain 197 N)	Species	50S ribosomal protein L7/L12	Q2L2M6				NC14	ENTTIVEGAGK APFGDR EIELADPFENMGAQLVK SVAAGMNPMDLK APFGDR VGAATEVEMK AVTALVAELKK VTLADLGQAK AAVEEGIVAGGGVALLR APFGDR AVEEPLR VGAATEVEMK EGVITVEDGK VQIEEATSDYDREK DLLPVLEQVAK VEDALHATR VQIEEATSDYDR AEILDAIAGMTVLELSELIK DLVDGAPKPKV IIAIDTNPAK TNLCVAVR
<i>β-proteobacteria; Burkholderiales; Oxalobacteraceae; Herminimonas</i>	<i>Herminimonas arsenicoxydans</i>	Species	Glutathione-dependent formaldehyde dehydrogenase (alcohol dehydrogenase class III), HEAR2039 Glutathione-independent formaldehyde dehydrogenase, HEAR2048	A4G6P6 A4G6Q5				NC12 NC3, NC11	ENTTIVEGAGK APFGDR EIELADPFENMGAQLVK SVAAGMNPMDLK APFGDR VGAATEVEMK AVTALVAELKK VTLADLGQAK AAVEEGIVAGGGVALLR APFGDR AVEEPLR VGAATEVEMK EGVITVEDGK VQIEEATSDYDREK DLLPVLEQVAK VEDALHATR VQIEEATSDYDR AEILDAIAGMTVLELSELIK DLVDGAPKPKV IIAIDTNPAK TNLCVAVR
<i>β-proteobacteria, Burkholderiales; unclassified Burkholderiales; Burkholderiales; Genera incertae sedis</i>	<i>Leptothrix cholodnii</i> or <i>Thiomonas intermedia</i>	Genus	60 kDa chaperonins ^a	B1XXY9, C7HZY6				2, 3, 4, 5, 36, 60	ENTTIVEGAGK APFGDR EIELADPFENMGAQLVK SVAAGMNPMDLK APFGDR VGAATEVEMK AVTALVAELKK VTLADLGQAK AAVEEGIVAGGGVALLR APFGDR AVEEPLR VGAATEVEMK EGVITVEDGK VQIEEATSDYDREK DLLPVLEQVAK VEDALHATR VQIEEATSDYDR AEILDAIAGMTVLELSELIK DLVDGAPKPKV IIAIDTNPAK TNLCVAVR

Table 3 (continued)

Phylum	Class, Family, Genus	Organism	Level of discrimination	Protein name	Protein accession numbers	Spot number ^b	Peptide sequence
		<i>Thiomonas</i> 3As	Genus	50S ribosomal protein L1; THI3722	FP475956	38	AMLEDIALTGGK AAVEEGVAGGGVALLR VAVSSTMGIGYR VDTATVNAAVAGQ APGFGDR VGAATEVEMK GVNILANAVK SFGAPTGTK YVAAGMNPMDLKR APGFGDR VQIEEATSDYDREK LQNMGAQMK VGAATEVEMK EGVITVEDGK VIAEEVGLTLEK YVAAGMNPMDLK VTLADLQQAK VEDALHATR AAVEEGVAGGGVALLR VQIEEATSDYDR VQIEEATSDYDR AMLEDIALTGGK APGFGDR VGAATEVEMK DLLPILEQVAK AAVEEGVAGGGVALLR APGFGDR EVVFGGEAR VGAATEVEMK EGVITVEDGK VIAEEVGLTLEK VQIEEATSDYDREK AVTALVAELKK VTLADLQQAK VEDALHATR VQIEEATSDYDR AMLEDIALTGGK SVAAGMNPMDLK
	β -proteobacteria, Burkholderiales; Burkholderiaceae; <i>Limnobacter</i>	<i>Limnobacter</i> sp. MED105	Genus or below	60 kDa chaperonins ^a	A6GTE5	4	
	β -proteobacteria, Burkholderiales; unclassified, Burkholderiales; Burkholderiales Genera <i>incertae sedis</i> ; <i>Methylibium</i>	<i>Methylibium petroleiphilum</i>	Genus or below	60 kDa chaperonins ^a	A2SCV1	2, 4	
	β -proteobacteria, Burkholderiales; Burkholderiaceae; <i>Ralstonia</i>	<i>Ralstonia pickettii</i>	Subspecies	60 kDa chaperonins ^a	B2U6M6	4	
	β -proteobacteria, Burkholderiales; Comamonadaceae; <i>Verminephrobacter</i>	<i>Verminephrobacter etseniae</i>	Genus	60 kDa chaperonins ^a	A1WL05	2, 4	
	β -proteobacteria, Methylotiales; Methylotrichaceae; <i>Methylotenera</i>	<i>Methylotenera mobilis</i>	Genus or below	60 kDa chaperonins ^a	C6WTL6	5, 6	

									TNDIAGDGGTTTATVLAQAIIR VGAATEVEMK GYLSPYFINNQDR DLLPVLEQVAK VEDALHATR APFGDR HALEGFK AVIAGMNPMDLK GVNVLADAVK VVSEEIGMK VEDALHATR AMLEDMALITGGR LESTTLADLGQAK APFGDR VGAATEVEMK VEDALHATR QIVANAGDEPSVVLNK LTATEDAAAR KADEALAAQK ADEALAAQK ITATEDAAAR SVAAGMNPMDLK AQIEDTSSDYDR APFGDR VGAATEVEMK GYLSPYFVTDSE LEGNNALGAK					
								2	C5V7N1	60 kDa chaperonins ^a	Genus or below	<i>Gallionella ferruginea</i>		
								2, 3	B5EN19, B7J561	60 kDa chaperonins ^a	Species	<i>Acidithiobacillus ferrooxidans</i>		
								2, 3, 4	Q60AY0	60 kDa chaperonins ^a	Order or below	<i>Methylococcus capsulatus</i>		
								4, 15, 17, 19, 22, 26–27, 40, 44, 49, 55, 60, 62, 64–65, NC7-NC10	A2VC34, A4XXVE5, O85409-085430-085432-085437, O85439-085444, Q3K906, Q48K14, Q883S8	Outer membrane lipoprotein OprI	Genus	<i>Pseudomonas</i>		
								5	Q1YSA6	60 kDa chaperonins ^a	Group or clade	Gammaaproteobacterium HTCC2207		
								4, 6	D0LRR3	60 kDa chaperonins ^a	Order or below	<i>Haliangium ochraceum</i>		
								NC1, NC3, NC11	Q3KFU6 ^c	Succinyl-CoA synthetase; beta subunit	Phylum	Several <i>Proteobacteria</i>		

^a MS/MS data were searched against an in-house database including all the bacterial and archaeal GroEL sequences obtained from Uniprot

^b Spots from 1 to 65 originated from a pH 4 to 7 gradient gel (Supplementary Fig. 1), spots from NC1 to NC14 originated from a pH 3 to 10 gradient gel (data not shown)

^c Several proteins may correspond to this identification: Q3KFU6, A1A8Y0, A1FGM7, A1JRB6, A1KTM6, A2UEB0, A3HHN5, A3M887, A4TNT8, A4W879, A4XV90, A5W114, A5WC33, A6BTC9, A6T6F6, A6V7K5, A7FKR4, A7MQX5, A7ZJA8, A7ZXY8, A8AJ84, A8GB83, P0A836, P0A837, P0A838, P0A839, P53593, P66869, P66870, Q02K73, Q0TJW6, Q1CAG1, Q1CFM0, Q1I7L3, Q1REJ8, Q2IIW6, Q2NUM2, Q2SD35, Q324I4, Q32IK3, Q3Z476, Q48K68, Q4FVH9, Q4KFY6, Q4ZUW7, Q57RL3, Q5F878, Q5PCM7, Q66DA0, Q6D7G2, Q6F8L4, Q7N6V5, Q7NZ47, Q883Z4, Q88FB2, Q8ZH00, Q9JUT0, Q9IZP4

39 bacterial proteins detected, there were two distinct ATP synthases, two distinct 50 S ribosomal proteins, one phosphoglycerate kinase, and one succinyl-CoA synthetase, which are involved in energy metabolism, translation, glycolysis, and TCA cycle, respectively (Table 3). These proteins may not play a specific role in this environment since they are known to be housekeeping proteins in bacteria. Among the other proteins, one outer protein OrpI, one uncharacterized protein with an unknown function, one glutathione-dependent, and one glutathione-independent formaldehyde dehydrogenases were identified. The latter two proteins belong to the formaldehyde detoxification pathway. Although no correlation with the environmental conditions might explain the functional specificities of these proteins, it has been reported that the arsenite-oxidizing bacteria *H. arsenicoxydans* synthesizes alcohol dehydrogenase and glutathione-dependent formaldehyde dehydrogenase when grown in the presence of arsenic [58].

It is worth noting that more than half of the identified bacterial proteins were similar to the 60-kDa GroEL chaperonin. These data suggest that multiple chaperonins of various genetic origins are expressed by the Reigous creek community. GroEL is known to be ubiquitously present in Bacteria and Archaea. These proteins are generally abundantly expressed in bacterial cells, especially under stress conditions such as those occurring in this particularly toxic environment [3]. The *groEL* gene is conserved in prokaryotes, and has been found to be present in one copy in the majority of sequenced genomes, except in the case of some pathogens [45, 60]. Because of its conservation properties (supplementary Fig. 2), this gene is often used as a phylogenetic marker [31]. In addition, it has been previously established that some of the amino acids stretch occurring in GroEL are specific to one genus or family of bacteria. These peptide sequences can therefore be used as the signature of a specific phylogenetic group. These GroEL identifications (Table 3, Fig. 3) were, therefore, considered for use as a possible taxonomic tool in addition to the 16S rRNA-based taxonomic approach. To determine which bacteria in the whole community were active, i.e., able to express proteins, the organisms identified using the 16S rRNA encoding gene library- and GroEL-based approaches (Fig. 2) were compared.

Most of the bacteria identified based on GroEL belonged to five phyla divisions. Bacteria belonging to *Deinococcus* and *Synergistetes* did not feature among those identified based on the 16S rRNA encoding gene. One possible explanation for this discrepancy may be that a PCR or cloning bias may have prevented those bacteria from being detected with this method. These findings suggest that metaproteomic methods used as taxonomic tools can provide a useful complementary tool in addition to the 16S rRNA encoding gene approach. Interestingly, *Thiomonas*, *At. ferrooxidans*, *Acidiphilium*, and *Gallionella*, expressed proteins and were, therefore, active.

16S rRNA-encoding gene analysis showed that these bacteria abundant in this ecosystem. Bacteria affiliated to *At. ferrooxidans* and *Gallionella* are able to oxidize iron. In addition, many strains of the *Thiomonas* genus are able to oxidize As(III) into As(V) under laboratory conditions [6, 14, 22]. The fact that their proteins were detected shows that these bacteria were viable and metabolically active. This finding supports the hypothesis that the oxidation of Fe(II) to Fe(III) catalyzed by iron-oxidizing microorganism such as *At. ferrooxidans* and *Gallionella* and oxidation of As(III) into As(V) by As(III) oxidizers such as *Thiomonas*, probably leads to the precipitation of the more or less ordered iron oxy-hydroxides (Fe(III)–As(V) hydroxysulfate) detected in this study (Fig. 1) [48]. This finding is in agreement with previous data showing that *Gallionella ferruginea* efficiently remove Fe, As(III), and As(V) in water [41]. The present data show for the first time that this bacterium is active and probably plays a functional role in the sediments of the Reigous creek. Some heterotrophic bacteria such as *Acidiphilium* were also found to be active in this AMD, suggesting that they could cope with the low amount of organic carbon (dissolved organic carbon concentration 1.7 ± 0.4 mg/L [16]) of Reigous creek water. It has previously been suggested that these acidophilic heterotrophic bacteria may be involved in organic carbon turnover processes [32]. Interestingly, these four bacteria (*At. ferrooxidans*, *Thiomonas*, *Gallionella*, and *Acidiphilium*) found to be active members of this AMD community have been previously identified in AMD, but some of them were thought to have different optimum pH levels. Indeed, *G. ferruginea* is a neutrophilic bacterium which oxidizes Fe, but relatives of *Gallionella*, have often been detected in AMD [12, 33–35]. The strain occurring at Carnoulès showed less than 97% homology with *G. capsiferriformans* and its physiological characteristics are probably different. It seems probable that an acid-tolerant relative of this bacterium is able to oxidize iron under acid pH conditions.

In addition to the bacteria belonging to these four genera, other bacteria were also found to be active, but a discrepancy was again observed between the bacteria identified based on the results of 16S rRNA encoding gene analysis and the metaproteomic approach. The GroEL protein sequences of some bacteria identified using the 16S rRNA encoding gene library were not available in the Uniprot database, which might explain this discrepancy. However, phylogenetic comparisons between the 16S rRNA and metaproteomic data obtained (Fig. 2) suggested which of the bacteria present in this ecosystem may express a GroEL identified in the metaproteomic study. Based on these comparisons, it seems likely that in addition to *Thiomonas*, *At. ferrooxidans*, *Acidiphilium*, and *Gallionella*, clones related to β -, γ -, and δ -proteobacteria, such as *Limnobacter* or *Methylococcus* (*At. ferrooxidans* DSM

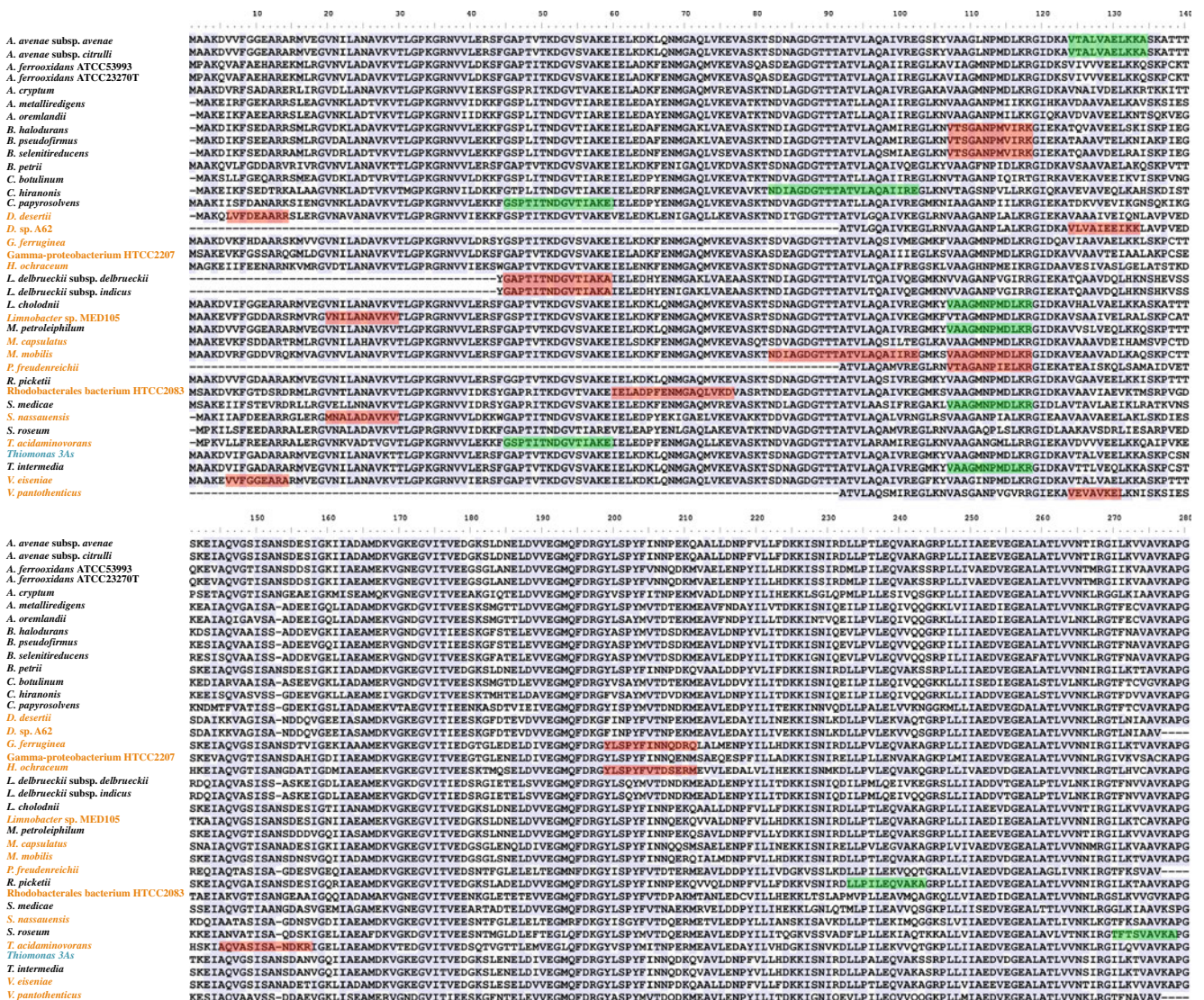


Figure 3 Alignment of the amino acid sequences of 60 kDa chaperonins. Thirty-six sequences matched the MS results: 29 Proteins were identified in Table 3: 22 of them unambiguously matched one protein (the bacterial name of which is given in orange), whereas fifteen identifications matched at least two proteins (22 bacterial names given in black). *Thiomonas* sp. 3As data were added for the sake of comparison (in blue). Thirty-seven sequences in all were aligned. These sequences were compared with ClustalW2 using the default parameters (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). Alignment was performed using the Jalview software program [19]. The blue highlighted letters correspond to identical residues among the 37 orthologs (BLOSUM62 score option). Highlighted letters corresponding to the peptide sequences identified by MS allowed distinguishing several orthologs. First, 23 proteins were identified because at least one identified peptide matched only this GroEL: these peptides (labeled in red) corresponding to signature sequences specific to one species, genus, or family of bacteria are located in the same homology regions but differed from other peptides by at least one amino acid substitution. Secondly, 13 of the proteins identified were confirmed, since several identified peptides (labeled in green) matched this GroEL. Each of these individual peptides matched several proteins; however, only one protein detected in Uniprot contained all these amino-acid sequences, which suggests that a relative of this protein was probably present in this extract. Since the full amino-acid sequences of almost all the

chaperonins expressed by the bacteria present at the study site are unknown, the possibility cannot be ruled out that two identified peptides were erroneously assigned to two distinct chaperonins, whereas these peptides may in fact have originated from one protein, the amino acid sequence of which has not yet been included in the databases. Nevertheless, because of the high level of conservation observed in GroEL proteins (Supplementary Fig. 2), the proteins identified may be expressed by a close relative of the organism identified using the Uniprot database. GroEL originated from *Acidiphilium cryptum*; *Acidithiobacillus ferrooxidans* ATCC23270; *Acidithiobacillus ferrooxidans* ATCC53993; *Acidovorax avenae* subsp. *avenae*; *Acidovorax avenae* subsp. *citrulli*; *Alkaliphilus metalliredigens*; *Alkaliphilus oremlandii*; *Bacillus halodurans*; *Bacillus pseudofirmus*; *Bacillus selenitireducens*; *Bordetella petrii*; *Clostridium botulinum*; *Clostridium hiranonis*; *Clostridium papyrosolvens*; *Deinococcus desertii*; *Deinococcus* sp. A62; *Gallionella ferruginea*; *Haliangium ochraceum*; *Lactobacillus delbrueckii* subsp. *delbrueckii*; *Lactobacillus delbrueckii* subsp. *indicus*; *Leptothrix cholodnii*; *Limnobacter* sp. MED105; *Methylidium petroleiphilium*; *Methylococcus capsulatus*; *Methylotenera mobilis*; *Protonibacterium freudenreichii*; *Ralstonia pickettii*; *Rhodobacteriales bacterium* HTCC2083; *Sinorhizobium medicae*; *Stackebrandtia nassauensis*; *Streptosporangium roseum*; *Thermanaerovibrio acidaminovorans*; *Thiomonas* 3As; *Thiomonas intermedia*; *Verminephrobacter eiseniae*; *Virgibacillus pantothenicus*; and *Gammaproteobacterium* HTCC2207

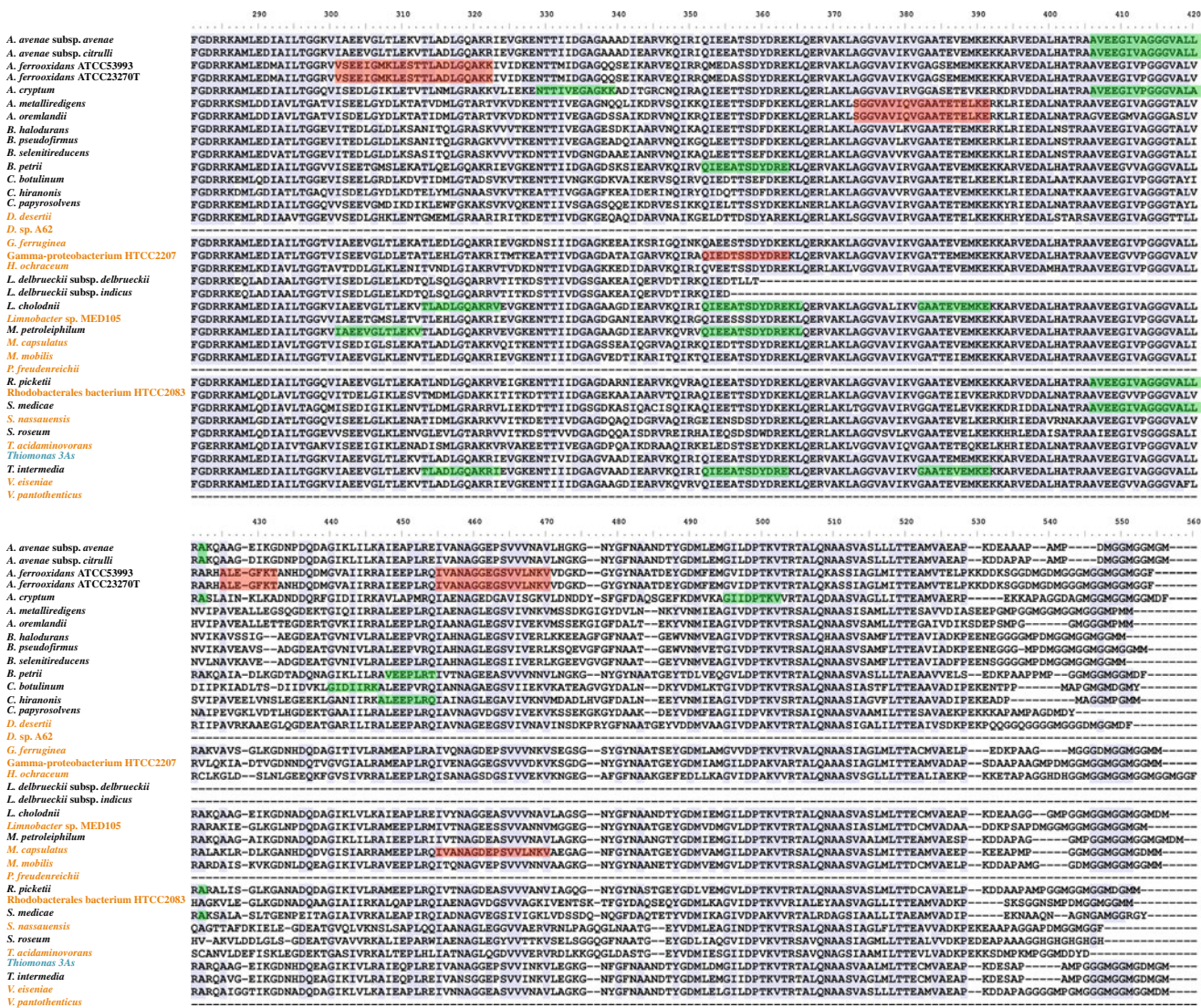


Figure 3 (continued)

2392 affiliated bacteria), *Methylothera* (*Thiobacillus*-related clones), *D. korensis* (which is affiliated to the γ -proteobacteria HTCC2207), *Ferrimicrobium*-like bacteria (which are related to the *Actinobacteria* *S. nassauensis* and *S. roseum*), and *Haliangium* (clones affiliated to sulfate-reducing bacteria JHA1, *D. limimaris* and *D. svalbardensis*) may play a role in this ecosystem. All in all, the data obtained here show that *Firmicutes*, which could be affiliated to *Alicyclobacillus ferrooxidans*, are also active. These bacteria are able to oxidize ferrous iron [37] and may, therefore, participate in the transformation of the iron present in high concentrations in these waters. The active population as a whole was not only composed of several iron-oxidizers in addition to *At. ferrooxidans* but also contained iron reducers, one known arsenite-oxidizer, sulfate-reducing, and sulfur compound oxidizers, and both autotrophic and heterotrophic bacteria. All these bacteria

may contribute importantly to the remediation process observed in situ.

In Conclusion

The active bacterial species inhabiting Carnoules AMD ecosystem were identified in this study using high-sensitivity nanoLC-chip-MS/MS methods combined with a 16S rRNA based phylogenetic approach. The meta-proteomic data obtained here show for the first time that *Gallionella*, *Thiomonas*, *At. ferrooxidans*, and *Acidiphilium* actively express proteins in situ. Previous hypotheses based on experiments performed under laboratory conditions in addition to *At. ferrooxidans* [28, 29, 32, 34, 38] suggest that microbial activity may contribute to the arsenite oxidation and As immobilization occurring in the heavily contaminated AMD at the Carnoules mine via iron oxidation processes.

Since these bacteria were found to be active and to express proteins which are among the most abundant proteins encountered at this site, it seems likely that the large amounts of pale-yellow As(V)–Fe(III) hydroxysulfate sediments forming at Carnoulès, which were characterized here, may result from the conjugate activities of iron-oxidizing microorganisms (such as *At. ferrooxidans*, *Alicyclobacillus ferrooxydans*, *Ferrimicrobium*, or *Gallionella*) and arsenic-oxidizing microorganisms (such as *Thiomonas* sp.). Several bacteria may be responsible in situ for changing the ratio between the oxidized and reduced forms of iron, arsenic, and sulfur compounds, promoting the formation of the Fe(III)–As(V) hydroxysulfate precipitates detected in this study. These bacteria are therefore of prime importance in the partial but efficient natural process of remediation undergone by the contaminated Carnoulès ecosystem. In addition, autotrophic iron, arsenic, and sulfur oxidizers may provide the organic carbon sources required by the functional heterotrophs such as *Acidiphilium* present in this ecosystem. All in all, the present findings provide evidence that the functional genomics approach provides a useful means of describing bacterial communities such as those inhabiting the Reigous creek and determining their contribution to natural attenuation processes. It is now proposed to use approaches of this kind in future studies to complete this picture of the functional processes at work in this ecosystem, as well as to investigate the role played by the less abundant active bacteria identified in this study.

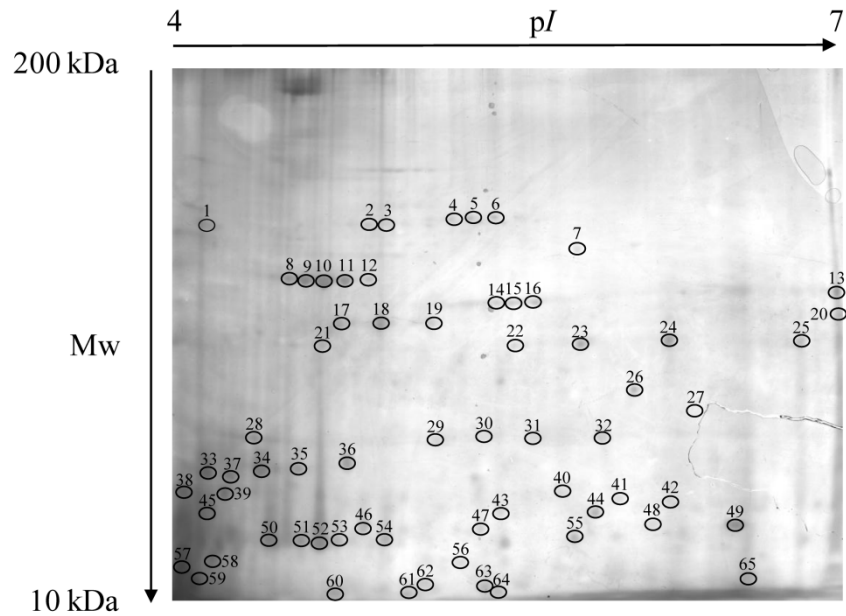
Acknowledgements The study was financed by the EC2CO program (“Institut National des Sciences de l’Univers,” CNRS), the “Observatoire de Recherche Méditerranéen en Environnement” (OSU-OREME), and by the ANR 07-BLANC-0118 project (“Agence Nationale de la Recherche”). Sébastien Gallien and Aurélie Volant were supported by a grant from the French Ministry of Education and Research. This work was performed in the framework of the “Groupement de recherche: Métabolisme de l’Arsenic chez les Microorganismes” (GDR2909-CNRS).

References

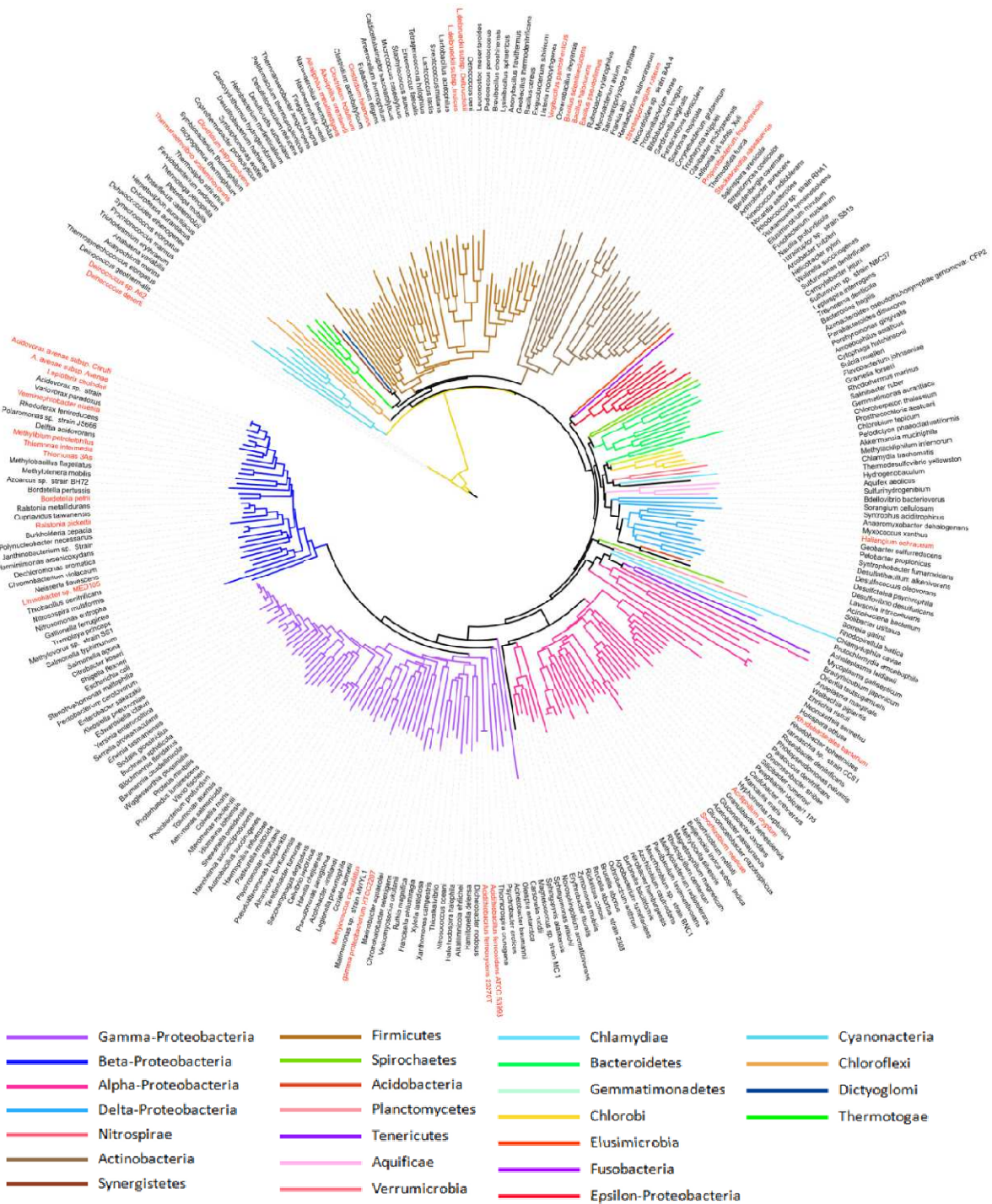
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Amaral Zettler LA, Gomez F, Zettler E, Keenan BG, Amils R, Sogin ML (2002) Microbiology: eukaryotic diversity in Spain’s River of Fire. *Nature* 417:137
- Arsène F, Tomoyasu T, Bukau B (2000) The heat shock response of *Escherichia coli*. *Int J Food Microbiol* 55:3–9
- Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Appl Environ Microbiol* 71:7724–7736
- Baker BJ, Banfield JF (2003) Microbial communities in acid mine drainage. *FEMS Microbiol Ecol* 44:139–152
- Battaglia-Brunet F, Dictor MC, 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
- Bednar AJ, Garbarino JR, Ranville JF, Wildeman TR (2002) Preserving the distribution of inorganic arsenic species in groundwater and acid mine drainage samples. *Environ Sci Technol* 36:2213–2218
- Bertin PN, Médigue C, Normand P (2008) Advances in environmental genomics: towards an integrated view of microorganisms and ecosystems. *Microbiology* 154:347–359
- Bond PL, Smriga SP, Banfield JF (2000) Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl Environ Microbiol* 66:3842–3849
- Bowell RJ (1994) Sorption of arsenic by iron oxides and oxyhydroxides in soils. *Appl Geochem* 9:279–286
- Brake SS, Hasiotis ST (2010) Eukaryote-dominated biofilms and their significance in acidic environments. *Geomicrobiol J* 27:534–558
- Bruneel O, Duran R, Casiot C, Elbaz-Poulichet F, Personné JC (2006) Diversity of microorganisms in Fe-As-rich acid mine drainage waters of Carnoules, France. *Appl Environ Microbiol* 72:551–556
- Bruneel O, Pascault N, Egal M, Bancon-Montigny C, Goni-Urriza MS, Elbaz-Poulichet F, Personné JC, Duran R (2008) Archaeal diversity in a Fe-As rich acid mine drainage at Carnoules (France). *Extremophiles* 12:563–571
- Bruneel O, Personné JC, Casiot C, Leblanc M, Elbaz-Poulichet F, Mahler BJ, Le Flèche A, Grimont PA (2003) Mediation of arsenic oxidation by *Thiomonas* sp. in acid-mine drainage (Carnoules, France). *J Appl Microbiol* 95:492–499
- Casiot C, Bruneel O, Personné JC, Leblanc M, Elbaz-Poulichet F (2004) Arsenic oxidation and bioaccumulation by the acidophilic protozoan, *Euglena mutabilis*, in acid mine drainage (Carnoules, France). *Sci Total Environ* 320:259–267
- Casiot C, Egal M, Bruneel O, Bancon-Montigny C, Cordier MA, Gomez E, Aliaume C, Elbaz-Poulichet F (2009) Hydrological and geochemical controls on metals and arsenic in a Mediterranean river contaminated by acid mine drainage (the Amous river, France); preliminary assessment of impacts on fish (*Leuciscus cephalus*). *Appl Geochem* 24:787–799
- Casiot C, Lebrun S, Morin G, Bruneel O, Personné JC, Elbaz-Poulichet F (2005) Sorption and redox processes controlling arsenic fate and transport in a stream impacted by acid mine drainage. *Sci Total Environ* 347:122–130
- Casiot C, Morin G, Juillot F, Bruneel O, Personné JC, Leblanc M, Duquesne K, Bonnefoy V, Elbaz-Poulichet F (2003) Bacterial immobilization and oxidation of arsenic in acid mine drainage (Carnoules creek, France). *Water Res* 37:2929–2936
- Clamp M, Cuff J, Searle SM, Barton GJ (2004) The Jalview Java alignment editor. *Bioinformatics* 20:426–427
- Coupland K, Johnson DB (2008) Evidence that the potential for dissimilatory ferric iron reduction is widespread among acidophilic heterotrophic bacteria. *FEMS Microbiol Lett* 279:30–35
- Duquesne K, Lebrun S, Casiot C, Bruneel O, Personné JC, 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, Muller D, Lett MC, Bonnefoy V (2008) Arsenite oxidation by a chemoautotrophic moderately acidophilic *Thiomonas* sp.: from the strain isolation to the gene study. *Environ Microbiol* 10:228–237
- Egal M, Casiot C, Morin G, Elbaz-Poulichet F, Cordier MA, Bruneel O (2010) An updated insight into the natural attenuation of As concentrations in Reigous Creek (southern France). *Appl Geochem* 25:1949–1957
- Egal M, Casiot C, Morin G, Parmentier M, Bruneel O, Lebrun S, Elbaz-Poulichet F (2009) Kinetic control on the formation of

- tooeleite, schwertmannite and jarosite by *Acidithiobacillus ferrooxidans* strains in an As(III)-rich acid mine water. *Chem Geol* 265:432–441
25. Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4:207–214
 26. Felsenstein J (1992) Estimating effective population size from samples of sequences: a bootstrap Monte Carlo integration method. *Genet Res* 60:209–220
 27. Ferris FG, Hallberg RO, Lyvén B, Pedersen K (2000) Retention of strontium, cesium, lead and uranium by bacterial iron oxides from a subterranean environment. *Appl Geochem* 15:1035–1042
 28. Finneran KT, Johnsen CV, Lovley DR (2003) *Rhodoferax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). *Int J Syst Evol Microbiol* 53:669–673
 29. Good IJ (1953) The population frequencies of species and the estimation of population parameters. *Biometrika* 40:237–264
 30. Goto K, Mochida K, Kato Y, Asahara M, Fujita R, An SY, Kasai H, Yokota A (2007) Proposal of six species of moderately thermophilic, acidophilic, endospore-forming bacteria: *Alicyclobacillus contaminans* sp. nov., *Alicyclobacillus fastidiosus* sp. nov., *Alicyclobacillus kakegawensis* sp. nov., *Alicyclobacillus macrosporangiidus* sp. nov., *Alicyclobacillus sacchari* sp. nov. and *Alicyclobacillus shizuokensis* sp. nov. *Int J Syst Evol Microbiol* 57:1276–1285
 31. Gupta RS (1998) Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeobacteria, eubacteria, and eukaryotes. *Microbiol Mol Biol Rev* 62:1435–1491
 32. Hallberg KB (2010) New perspectives in acid mine drainage microbiology. *Hydrometallurgy* 104:448–453
 33. Hallberg KB, Coupland K, Kimura S, Johnson DB (2006) Macroscopic streamer growths in acidic, metal-rich mine waters in north wales consist of novel and remarkably simple bacterial communities. *Appl Environ Microbiol* 72:2022–2030
 34. He Z, Xiao S, Xie X, Zhong H, Hu Y, Li Q, Gao F, Li G, Liu J, Qiu G (2007) Molecular diversity of microbial community in acid mine drainages of Yunfu sulfide mine. *Extremophiles* 11:305–314
 35. Heinzl E, Hedrich S, Janneck E, Glombitza F, Seifert J, Schlomann M (2009) Bacterial diversity in a mine water treatment plant. *Appl Environ Microbiol* 75:858–861
 36. Hippe H (2000) *Leptospirillum* gen. nov. (ex Markosyan 1972), nom. rev., including *Leptospirillum ferrooxidans* sp. nov. (ex Markosyan 1972), nom. rev. and *Leptospirillum thermoferrooxidans* sp. nov. (Golovacheva et al. 1992). *Int J Syst Evol Microbiol* 50(Pt 2):501–503
 37. Jiang CY, Liu Y, Liu YY, You XY, Guo X, Liu SJ (2008) *Alicyclobacillus ferrooxydans* sp. nov., a ferrous-oxidizing bacterium from solfataric soil. *Int J Syst Evol Microbiol* 58:2898–2903
 38. Johnson DB (1995) Acidophilic microbial communities: candidates for bioremediation of acidic mine effluents. *Int Biodeterior Biodegrad* 35:41–58
 39. Johnson DB, Hallberg KB (2003) The microbiology of acidic mine waters. *Res Microbiol* 154:466–473
 40. Johnson DB, Hallberg KB (2005) Acid mine drainage remediation options: a review. *Sci Total Environ* 338:3–14
 41. Katsoyiannis IA, Zouboulis AI (2004) Application of biological processes for the removal of arsenic from groundwaters. *Water Res* 38:17–26
 42. Kelly DP, Wood AP (2000) Confirmation of *Thiobacillus denitrificans* as a species of the genus *Thiobacillus*, in the beta-subclass of the *Proteobacteria*, with strain NCIMB 9548 as the type strain. *Int J Syst Evol Microbiol* 50(Pt 2):547–550
 43. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci USA* 82:6955–6959
 44. Letunic I, Bork P (2007) Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128
 45. Lund PA (2009) Multiple chaperonins in bacteria—why so many? *FEMS Microbiol Rev* 33:785–800
 46. 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
 47. Morin G, Calas G (2006) Arsenic in soils, mine tailings, and former industrial sites. *Elements* 2:97–101
 48. Morin G, Juillot F, Casiot C, Bruneel O, Personné JC, 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 Carnoules acid mine drainage, France. A XANES, XRD, and SEM study. *Environ Sci Technol* 37:1705–1712
 49. Ona-Nguema G, Morin G, Juillot F, Calas G, Brown GE (2005) EXAFS analysis of arsenite adsorption onto two-line ferrihydrite, hematite, goethite, and lepidocrocite. *Environ Sci Technol* 39:9147–9155
 50. Rodier J, Broutin JP, Chambon P, Champsaur H, Rodi L (1996) *L'Analyse des Eaux*. Dunod, Paris, p 1383
 51. Rowe OF, Johnson DB (2008) Comparison of ferric iron generation by different species of acidophilic bacteria immobilized in packed-bed reactors. *Syst Appl Microbiol* 31:68–77
 52. Rowe OF, Sanchez-Espana J, Hallberg KB, Johnson DB (2007) Microbial communities and geochemical dynamics in an extremely acidic, metal-rich stream at an abandoned sulfide mine (Huelva, Spain) underpinned by two functional primary production systems. *Environ Microbiol* 9:1761–1771
 53. Sánchez España JS, LópezPamo E, Santofimia Pastor E, Reyes Andrés J, Rubi J (2005) The natural attenuation of two acidic effluents in Tharsis and La Zarza-Perrunal mines (Iberian Pyrite Belt, Huelva, Spain). *Environ Geol* 49:253–266
 54. Sun B, Cole JR, Tiedje JM (2001) *Desulfomonile limimaris* sp. nov., an anaerobic dehalogenating bacterium from marine sediments. *Int J Syst Evol Microbiol* 51:365–371
 55. 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
 56. Vandieken V, Musmann M, Niemann H, Jorgensen BB (2006) *Desulfuromonas svalbardensis* sp. nov. and *Desulfuromonas ferrireducens* sp. nov., psychrophilic, Fe(III)-reducing bacteria isolated from Arctic sediments, Svalbard. *Int J Syst Evol Microbiol* 56:1133–1139
 57. Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697–703
 58. Weiss S, Carapito C, Cleiss J, Koechler S, Turlin E, Coppee J-Y, Heymann M, Kugler V, Stauffert M, Cruveiller S, Médigue C, Van Dorsselaer A, Bertin PN, Arsène-Ploetze F (2009) Enhanced structural and functional genome elucidation of the arsenite-oxidizing strain *Herminiimonas arsenicoxydans* by proteomics data. *Biochimie* 91:192–203
 59. Weiss JV, Rentz JA, Plaia T, Neubauer SC, Merrill-Floyd M, Lilburn T, Bradburne C, Megonigal JP, Emerson D (2007) Characterization of neutrophilic Fe(II)-oxidizing bacteria isolated from the rhizosphere of wetland plants and description of *Ferritrophicum radicolica* gen. nov. sp. nov., and *Sideroxydans paludicola* sp. nov. *Geomicrobiol J* 24:559–570
 60. Williams TA, Codoner FM, Toft C, Fares MA (2010) Two chaperonin systems in bacterial genomes with distinct ecological roles. *Trends Genet* 26:47–51

SUPPLEMENTARY MATERIAL



Supplementary Figure 1: 2DE metaproteomic map of the main cytosolic proteins expressed by the bacterial community inhabiting the Reigous creek sediments. Spots analyzed with nanoLC-MS/MS are *circled* and *numbered*. The size of protein markers (BenchMark™ Protein Ladder, Invitrogen) and the pI gradient are schematized from 10 to 220 kDa on the left and from 4 to 7 at the top, respectively.



Supplementary Figure 2: Phylogenetic tree based on GroEL amino acid sequences. A total number of 530 GroEL bacterial sequences were retrieved from the RefSeq database: these sequences were 500–550 amino acids in length. From these sequences, only one representative of each genus was kept for the comparisons, which involved 259 sequences. These GroEL sequences were aligned using ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). Accession numbers: see supplementary data. In *red*: GroEL sequences identified using the metaproteomic approach; in *black*: GroEL sequences detected in the Uniprot database.

Chapitre II – Dynamique temporelle de la communauté d'Archaea

A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE

1. Contexte général

Historiquement, on pensait que les bactéries représentaient les principaux microorganismes impliqués dans les processus de lixiviation à l'origine de la formation des DMA. Cependant, sur le site de la mine d'Iron mountain (Californie), il a été montré que la bactérie *A. ferrooxidans*, supposée être l'espèce majeure oxydant le fer, ne jouait finalement qu'un rôle mineur dans l'oxydation de la pyrite alors que la présence d'Archaea en forte proportion, dont certaines capables d'oxyder le fer et le soufre, suggérait la probable importance de ces microorganismes dans ce phénomène sur ce site (Edwards *et al.*, 2000b). Il apparaît aujourd'hui évident que les Archaea sont elles aussi susceptibles de jouer un rôle majeur dans les processus de génération et/ou de remédiation des DMA et d'autres sites pollués par des ETM (Baker & Banfield, 2003; Bini, 2010). Cependant, il est important de noter qu'en raison de difficultés d'isolement, les études réalisées jusqu'à présent n'ont permis d'acquérir des connaissances que sur un nombre restreint de souches. Ainsi, la diversité des espèces et leur rôle physiologique au sein des DMA restent encore mal connus.

Concernant le DMA de Carnoulès, une première étude a été réalisée sur les eaux et a permis de mettre en évidence la présence de *Ferroplasma acidiphilum* et de séquences affiliées à des *Thermoplasmatales* non cultivés (Bruneel *et al.*, 2008). La mise en évidence de *F. acidiphilum* au niveau du piézomètre, localisé dans le stérile minier, suggérait la possible implication de ce microorganisme dans le phénomène de lessivage intensif observé à ce niveau ainsi que la libération de concentrations très élevées d'arsenic, jusqu'à 10 g L^{-1} . En effet, ce microorganisme acidophile, sans paroi cellulaire, qui oxyde le fer est connu pour son rôle majeur dans le lessivage (Golyshina & Timmis, 2005). De plus, des études précédentes avaient souligné le fait que les bactéries oxydant le fer, *A. ferrooxidans* et *Sulfobacillus* spp., bien que présentes, ne représentaient qu'une population mineure, ce qui confortait cette

hypothèse. D'autre part, la majorité des séquences retrouvées au niveau du Reigous étaient affiliées à un clone non cultivable ce qui ne permettait pas d'émettre d'hypothèses concernant leur rôle physiologique présumé.

2. Objectif de l'étude

Pour contribuer à améliorer nos connaissances sur le fonctionnement de ce DMA, nous avons caractérisé la communauté d'Archaea présente cette fois-ci dans les sédiments ainsi que sa dynamique temporelle par l'analyse d'une banque de gènes codant pour l'ARNr 16S.

B. RÉSULTATS

L'affiliation taxonomique des Archaea a montré un faible degré de diversité composée de deux phyla : Euryarchaeota et Thaumarchaeota. La majorité des séquences étaient apparentées à l'ordre des *Thermoplasmatales* bien qu'aucun clone ne présentait de forte similarité (> 97 %) avec des espèces cultivées. Certains membres des *Thermoplasmatales* sont capables de tirer de l'énergie de l'oxydation ou de la réduction du fer et/ou du soufre. Des microorganismes affiliés à des Archaea méthanogènes, impliquées dans le cycle du carbone, telle que *Methanomassiliicoccus luminyensis*, ont également été détectées. Enfin, nous avons trouvé des microorganismes impliqués dans l'oxydation de l'ammonium, une étape clé du cycle de l'azote, avec des séquences affiliées à *Candidatus Nitrososphaera viennensis* et *Candidatus nitrosopumilus* sp..

Concernant l'aspect temporel, cette étude a montré qu'au cours de la période d'échantillonnage, la communauté d'Archaea a subi des modifications significatives. Toutes les périodes d'échantillonnage présentaient des différences en termes de structure et de composition de la communauté bien qu'avril 2006 et octobre 2008 étaient relativement similaires en termes de structure. De plus, l'analyse des coefficients de similarité a montré que janvier 2009 était très différent des autres périodes d'échantillonnage. A cette date, la communauté d'Archaea a changé et sa diversité a radicalement chuté. Presque toutes les séquences étaient affiliées à un clone non cultivé proche d'une lignée de méthanogènes (*Methanothermobacter thermautotrophicus*). Ce clone, trouvé précédemment en avril 2006 correspondait alors à seulement 2 % de la communauté. Cependant, il est devenu la population dominante en janvier 2009. Après cela, en novembre 2009, la communauté

d'Archaea s'est diversifiée, des espèces précédemment détectées ont été retrouvées et de nouvelles espèces sont également apparues.

C. DISCUSSION ET CONCLUSION

Cette étude a permis de mettre en évidence des modifications de la structure et de la composition de la communauté d'Archaea au cours du temps qui sont probablement liées à des modifications de leur environnement. Elle a également mis en évidence la probable implication des Archaea dans un métabolisme actuellement inconnu chez les bactéries, la méthanogénèse. Certaines populations semblent également impliquées dans le cycle de l'azote (*Candidatus Nitrososphaera viennensis* et *Candidatus Nitrosopumilus* sp.) et d'autres pourraient intervenir dans les cycles du fer et du soufre (*Thermoplasmatales*). L'ensemble de ces microorganismes pourrait donc contribuer conjointement avec les bactéries au processus de remédiation observé *in situ*. Ils pourraient également être impliqués dans la stabilité des sédiments à travers la modification du ratio entre les formes oxydées et réduites du fer, de l'arsenic et des composés soufrés, qui favoriserait ainsi la formation et/ou la dissolution des précipités d'hydroxysulfate de Fe(III)-As(V).

Ces résultats, bien qu'apportant des éclaircissements sur la communauté d'Archaea et son rôle potentiel dans cet écosystème, ont également mis en évidence le fait que pour progresser dans notre compréhension du fonctionnement de ces environnements extrêmes et élucider le rôle des microorganismes présents, il est encore nécessaire d'améliorer les stratégies de culture à la base de la caractérisation taxonomique, physiologique et métabolique des nouvelles espèces ainsi détectées.

D. CONTRIBUTION SCIENTIFIQUE

Ce travail a conduit à une publication scientifique dans la revue « Extremophiles » dont le manuscrit correspondant est présenté ici (Article B).

Volant, A., Desoeuvre, A., Casiot, C., Lauga, B., Delpoux, S., Morin, G., Personne, J. C., Hery, M., Elbaz-Poulichet, F., Bertin, P. N. and Bruneel, O. (2012). Archaeal diversity: temporal variation in the arsenic-rich creek sediments of Carnoulès Mine, France. *Extremophiles* **16**: 645-657.

Chapitre II – Article B

Archaeal diversity: temporal variation in the arsenic-rich creek sediments of Carnoulès Mine, France

A. Volant, A. Desoeuvre, C. Casiot, B. Lauga, S. Delpoux, G. Morin, J. C. Personné, M. Héry, F. Elbaz-Poulichet, P. N. Bertin, O. Bruneel

Extremophiles (2012) 16:645–657 DOI 10.1007/s00792-012-0466-8

Archaeal diversity: temporal variation in the arsenic-rich creek sediments of Carnoulès Mine, France

A. Volant · A. Desoeuvre · C. Casiot · B. Lauga · S. Delpoux ·
G. Morin · J. C. Personné · M. Héry · F. Elbaz-Poulichet ·
P. N. Bertin · O. Bruneel

Received: 16 February 2012 / Accepted: 3 May 2012 / Published online: 20 June 2012
© Springer 2012

Abstract The Carnoulès mine is an extreme environment located in the South of France. It is an unusual ecosystem due to its acidic pH (2–3), high concentration of heavy metals, iron, and sulfate, but mainly due to its very high concentration of arsenic (up to 10 g L⁻¹ in the tailing stock pore water, and 100–350 mg L⁻¹ in Reigous Creek, which collects the acid mine drainage). Here, we present a survey of the archaeal community in the sediment and its temporal variation using a culture-independent approach by cloning of 16S rRNA encoding genes. The taxonomic affiliation of Archaea showed

a low degree of biodiversity with two different phyla: Eurarchaeota and Thaumarchaeota. The archaeal community varied in composition and richness throughout the sampling campaigns. Many sequences were phylogenetically related to the order *Thermoplasmatales* represented by aerobic or facultatively anaerobic, thermoacidophilic autotrophic or heterotrophic organisms like the organotrophic genus *Thermogymnomonas*. Some members of *Thermoplasmatales* can also derive energy from sulfur/iron oxidation or reduction. We also found microorganisms affiliated with methanogenic Archaea (*Methanomassiliicoccus luminyensis*), which are involved in the carbon cycle. Some sequences affiliated with ammonia oxidizers, involved in the first and rate-limiting step in nitrification, a key process in the nitrogen cycle were also observed, including *Candidatus Nitrososphaera viennensis* and *Candidatus nitrosopumilus* sp. These results suggest that Archaea may be important players in the Reigous sediments through their participation in the biochemical cycles of elements, including those of carbon and nitrogen.

Communicated by F. Robb.

Electronic supplementary material The online version of this article (doi:10.1007/s00792-012-0466-8) contains supplementary material, which is available to authorized users.

A. Volant · A. Desoeuvre · C. Casiot · S. Delpoux ·
J. C. Personné · M. Héry · F. Elbaz-Poulichet · O. Bruneel (✉)
Laboratoire HydroSciences Montpellier, HSM, UMR 5569
(IRD, CNRS, Universités Montpellier 1 et 2),
Université Montpellier 2, Place E. Bataillon, CC MSE,
34095 Montpellier, France
e-mail: bruneel@msem.univ-montp2.fr

B. Lauga
Equipe Environnement et Microbiologie, EEM, UMR 5254
(IPREM, CNRS), Université de Pau et des Pays de l'Adour,
BP 1155, 64013 Pau, France

G. Morin
Institut de Minéralogie et de Physique des Milieux Condensés,
IMPMC, UMR 7590 (CNRS, Université Pierre et Marie
Curie/Paris 6), 4 place Jussieu, 75252 Paris, France

P. N. Bertin
Laboratoire de Génétique Moléculaire, Génomique,
Microbiologie, GMGM, UMR 7156 (Université de Strasbourg,
CNRS), Département Microorganismes, Génomes,
Environnement, 28 Rue Goethe, 67083 Strasbourg, France

Keywords Archaea · Diversity · Arsenic ·
Acid mine drainage · Lead and zinc mine

Introduction

Acid mine drainage (AMD) water is a worldwide environmental problem caused by active and abandoned mines (Johnson and Hallberg 2003). Mining and processing of sulfide-rich ores produce large amounts of pyrite-rich waste. In contact with meteoric water, oxidation of this material generates AMD. These effluents are generally characterized by a low pH and contain significant quantities of sulfates, metals and metalloids including arsenic. AMD generation is mainly mediated by acidophilic iron-oxidizing microorganisms

(Edwards et al. 1999). Natural remediation of AMD can be observed at the Carnoulès site (France) or at Rio Tinto (Spain) (Casiot et al. 2003; Sánchez España et al. 2005). This natural remediation of metal pollutants is generally due to the occurrence of abiotic reactions and/or microbial activities that make these toxic compounds insoluble and lead to their accumulation in sediments (Hallberg 2010). These precipitations mainly involve the oxidation and precipitation of iron and the adsorption of other metals and metalloids by the resulting ferric minerals. Sulfate-reducing bacteria also have the ability to reduce sulfate to sulfide, which then reacts with certain dissolved metals to form insoluble precipitates (Hallberg 2010). In addition, several bacteria contribute to the immobilization of arsenic, thanks to their ability to oxidize this metalloid, arsenate As(V) being less soluble than arsenite As(III) (Bowell 1994).

The microbiology of AMD streams has been the subject of numerous studies. While a large amount of information is available on acidophilic bacteria indigenous to AMD, little is known about Archaea (Hallberg 2010). Several studies evidenced the presence of archaeal communities in acidic waters (Edwards et al. 2000; Dopson et al. 2004). The Archaea reported in AMD systems include groups of sulfur and/or iron-oxidizers, such as *Sulfolobus*, *Acidianus*, *Metallorhiza*, *Sulfurisphaera*, and *Ferroplasma* (Edwards et al. 2000; Golyshina et al. 2000; Baker and Banfield 2003). It has consequently been suggested that Archaea could also play a major role in the generation and remediation of AMD via oxidation of iron (Baker and Banfield 2003). Archaea may also play a role in the biogeochemical cycling of arsenic, for example, through the presence of Archaea that respire As(V) like *Pyrobaculum aerophilum* and *Pyrobaculum arsenaticum* (Huber et al. 2000; Oremland and Stolz 2003).

In a previous study, Bruneel et al. (2008) investigated the archaeal community in water samples from Carnoulès, an AMD very rich in metallic elements and especially arsenic compared to many others AMD. This study reported the presence of *Ferroplasma acidiphilum* and sequences affiliated to uncultured *Thermoplasmatales* archaeon. However, the archaeal population that inhabits the arsenic-rich Reigous sediments has never been characterized. Thus, to improve our understanding of AMD functioning, we characterized the archaeal communities present in sediment samples from the arsenic-rich AMD of the Carnoulès mine (France) and their temporal variations using a 16S rRNA encoding gene library.

Materials and methods

Description of the study site

The Pb-Zn Carnoulès mine, located in southern France, was closed in 1962. The mining extraction left about 1.2 Mt of

solid sulfidic wastes containing 0.7 wt% lead, 10 wt% iron and 0.2 wt% arsenic, which are stored behind a 6 m high dam on the uppermost course of Reigous Creek. The seepage water, which percolates through the wastes, emerges at the base of the tailings dam, and is the initial source of Reigous Creek. The water is acidic ($2 < \text{pH} < 3$) and rich in dissolved sulfate, iron and arsenic (2000–7700, 500–1000 and 50–350 mg L⁻¹, respectively) the later being predominantly in reduced forms: Fe(II) and As(III) (Casiot et al. 2003). The arsenic concentration decreases within the first 30 m of the creek mainly due to bacterial iron oxidation which leads to the coprecipitation of 20–60 % of dissolved arsenic (Casiot et al. 2003). Although the arsenic level remains high, its concentration subsequently decreases by around 95 % between the source of Reigous Creek and its confluence with the Amous River, 1.5 km downstream. The precipitates, which form around stromatolitic-like bacterial structures, are mainly composed of Fe(III)–As(III) in winter in the first 10 m and of amorphous Fe(III)–As(V) during the rest of the year (Casiot et al. 2003; Morin et al. 2003). Many studies (including culture-dependent and independent) have been conducted on the bacterial communities inhabiting the Carnoulès mine. Two of them focused specifically on sediment. The active bacterial species were identified in the sediments in the April 2006 library using high sensitivity nanoLC-chip-MS/MS methods combined with a 16S rRNA based phylogenetic approach (Bruneel et al. 2011). This study showed that *Gallionella*, *Thiomonas*, *Acidithiobacillus ferrooxidans*, and *Acidiphilium* actively expressed proteins in situ. Meta- and proteo-genomics approaches were also used on sediments in the May 2007 library and allowed reconstruction of seven bacterial strains (Bertin et al. 2011). These studies and previous results (Casiot et al. 2003; Morin et al. 2003) suggest that the large amounts of As(V)–Fe(III) hydroxysulfate sediments forming at Carnoulès may result from the combined activities of iron-oxidizing microorganisms (such as *At. ferrooxidans*, *Alicyclobacillus ferrooxidans*, *Ferrimicrobium*, or *Gallionella*) and arsenic-oxidizing microorganisms (such as *Thiomonas* sp.).

Sampling procedure and measurement of physicochemical properties

Four sampling campaigns were carried out in April 2006, October 2008, January 2009 and November 2009. Samples were collected at the station called COWG (Carnoulès Oxidizing Wetland, point G) located 30 m downstream of the spring (Bruneel et al. 2003). 5 cm deep pale yellow loosely packed sediments were collected at the bottom of the creek using a sterile spatula and pooled [global positioning system (GPS) coordinates: 44107001.8000N/4100006.9000E]. This sampling was done in three replicates. Solid phases were harvested by centrifugation and

dried under vacuum before mineralogical and spectroscopic analyses. The main physicochemical parameters (pH, T °C, and dissolved oxygen concentration) of the running water at the sampling point were measured in the field. pH and water temperature were measured with an Ultrameter Model 6P (Myron L 125 Company, Camlab, Cambridge). Water samples (500 ml) were immediately filtered through 0.22 µm Millipore membranes fitted on Sartorius polycarbonate filter holders. For total iron and arsenic determination, the filtered water was acidified to pH 1 with HNO₃ (14.5 M) and stored at 4 °C in polyethylene bottles until analysis. For arsenic and iron speciation, a 20 µl aliquot of filtered water sample was added to either a mixture of acetic acid and EDTA (Samanta and Clifford 2005) for arsenic speciation or to a mixture of 0.5 ml acetate buffer (pH 4.5) and 1 ml of 1,10-phenanthroline chloride solution (Rodier et al. 1996) for Fe(II) determination. The vials were completed to 10 ml with deionized water. The samples for iron and arsenic speciation and sulfate determination were stored in the dark and analyzed within 24 h. Chemical analysis was carried out as previously described (Bruneel et al. 2011).

Solid sample characterization

The mineralogical composition of the solid samples collected at COWG was qualitatively determined using powder X-ray diffraction analysis (XRD). Data were collected with Co K-alpha radiation on an X'Pert PRO P analytical diffractometer equipped with an X'Celerator detector, in continuous mode and a counting time of 4 h per sample. X-ray absorption spectroscopy data were collected on the solid phases sampled at COWG in October 2008, January 2009, and November 2009. X-ray absorption near edge structure (XANES) and extended X-ray absorption fine structure (EXAFS) spectra were recorded at a temperature 10–15 K in fluorescence mode on the FAME BM30B bending magnet beamline at ESRF (Grenoble, France). Data for the April 2006 COWG sample were previously collected at the 11-2 wiggler beamline at SSRL (Stanford, CA) and analyzed in Bruneel et al. (2011). Experimental details and data reduction procedures are reported in previous studies (Morin et al. 2003; Ona-Nguema et al. 2005; Hohmann et al. 2011). XANES and EXAFS data were interpreted by linear combination fitting using a set of model compound spectra. This set includes As(V)- and As(III)–Fe(III) oxyhydroxides and oxyhydroxysulfates synthesized via biotic and abiotic pathways (Morin et al. 2003; Maillot 2011).

DNA isolation

Triplicate genomic DNA was extracted from sediments using the UltraClean Soil DNA Isolation Kit according to

the manufacturer's recommendations (MoBio Laboratories Inc., Carlsbad, CA, USA). These triplicates were pooled before PCR amplification. All extracted genomic DNA samples were stored at –20 °C until further analysis.

PCR amplification

Amplification of archaeal 16S rRNA genes was obtained by PCR using primers Arch21F (5'-TTCCGGTTGATCC YGCCGGA-3') and Arch958R (5'-YCCGGCGTTGAMTC CAATT-3') (Delong 1992). Two PCR protocols were used due to major amplification difficulties. The first PCR amplification mixture contained 2 µl of DNA template, 2 µl of both primers (20 µM), 25 µl of PCR Master Mix Ampli Taq Gold 360 (Applied Biosystems, Foster City, CA, USA). Sterile distilled water was added to reach a final volume of 50 µl. The PCR conditions were as follows, an initial denaturation step of 95 °C for 7 min followed by 35 denaturation cycles at 95 °C for 1 min, an annealing cycle at 55 °C for 45 s and an extension cycle at 72 °C for 1 min. Final extension was at 72 °C for 10 min. As amplification of the January 2009 sample failed with this protocol, another enzyme was used, the PCR Extender Polymerase Mix (5Prime, Hamburg, Deutschland) as well as for a part of the November 2009 sample, which was also very difficult to amplify. The second PCR amplification mixture contained 2 µl of DNA template, 2 µl of both primers (20 µM), 2.5 µl of dNTPs 10 mM, 5 µl reaction Tuning buffer ×10 and 0.5 µl of PCR Extender Polymerase Mix (5Prime, Hamburg, Deutschland). Sterile distilled water was added to reach a final volume of 50 µl. The PCR conditions were as follows: initial denaturation step at 94 °C for 3 min followed by 35 denaturation cycles at 94 °C for 1 min, an annealing cycle at 55 °C for 1 min, and an extension cycle at 72 °C for 1.5 min. Final extension was at 72 °C for 10 min. PCR products were purified with the GFX PCR DNA purification kit (Amersham-Pharmacia). The PCR Extender polymerase mix creates blunt ended products. For TA Cloning®, 3' A-overhangs are needed on these PCR products, which are obtained with a different Taq polymerase. To 25 µl of purified PCR product, 2.5 µl of buffer 10×, 0.5 µl of dATPs 10 mM, and 0.5 µl of Taq DNA polymerase (Eurobiotaq®, Eurobio, France) were added. The PCR amplification mixture was then incubated at 72 °C for 20 min.

Cloning and 16S rRNA gene sequencing

The PCR products were cloned in *E. coli* TOP 10 strain using the pCR2.1 Topo TA cloning kit (Invitrogen, Inc., Carlsbad, CA, USA). Cloned 16S rRNA gene fragments were re-amplified using the primers TOP1 (5'-GTGTGCT GGAATTCGCCCTT-3') and TOP2 (5'-TATCTGCAGAA

TTCGCCCTT-3') that anneal to the vector and flank the inserted PCR fragment. A total of 340 clones from the four libraries were sequenced. Partial sequences of the clones were determined by the dideoxy nucleotide chain-termination method using the BigDye 3.1 kit (Applied Biosystems) on an ABI PRISM 3730XL Genetic analyzer (Applied Biosystems). The MALLARD program (Ashelford et al. 2006) was used to detect and then remove chimera. Sequences were also examined manually for chimera, which were excluded from further analyses. Sequences were then aligned in Mothur (<http://www.mothur.org>) (Schloss et al. 2009) using the SILVA archaeal database as reference alignment. The same program was used to calculate a neighbor-joining (NJ) (Saitou and Nei 1987) distance matrix using the Jukes-Cantor (JC) correction. The matrix was then used to assign sequences to operational taxonomic units (OTUs) defined at 97 (species level) and 85 % (class level) cutoff using the furthest-neighbor algorithm. Sequences were compared with the available databases NCBI and Greengenes (<http://greengenes.lbl.gov>) by BLAST online searches (Altschul et al. 1990) and Mothur to identify their taxonomic identities. Representative sequences for each OTU defined at 97 % cutoff were identified using the tool implemented in Mothur and were submitted to the EMBL databases under accession numbers (HE653775–HE653816).

Phylogenetic analysis

Archaeal 16S rRNA gene homologs were collected from the database at NCBI using the BLAST program with default parameters; one representative of each OTU was selected, giving a dataset of 99 sequences for final analysis. Multiple sequence alignment of partial prokaryotic sequences was performed using Clustal W (Thompson et al. 2000). A maximum likelihood phylogenetic reconstruction was obtained using the PhyML program (Guindon and Gascuel 2003) with the GTR model, four evolutionary rates, a calculated proportion of invariant sites and calculated nucleotide frequencies (default parameters). Statistical likelihood at nodes was calculated via a likelihood-ratio test (Anisimova and Gascuel 2006).

Statistical analysis of diversity and comparison of archaeal libraries

The Mothur software package was also used to generate diversity indices and statistics (OTUs, total clones, singletons, Chao1, Shannon, evenness, coverage) for each clone library as sequence similarity with a 97 % cutoff. The total number of clones obtained compared with the number of clones representing each unique phylotype was used to produce the rarefaction curves at the 85 % level. Coverage

values were calculated to determine how efficiently the libraries described the complexity of a theoretical community like an original archaeal community. The coverage (Good 1953) value is given as $C = 1 - (n_1/N)$ where n_1 is the number of clones that occurred only once in the library. To determine the significance of differences between archaeal libraries, a LIBSHUFF statistical analysis was performed in Mothur following Singleton et al.'s (2001) method. A LIBSHUFF comparison of libraries yielded the following formula using the Bonferroni correction: $0.05 = 1 - (1 - a)k(k - 1)$, where a is the critical P value and k is the number of libraries. The critical P value is 0.0042 when four libraries are compared. If any comparison of two libraries has a P value below or equal to 0.0042, then there is 95 % confidence to believe that the two libraries concerned differ significantly in community composition. Jaccard and Yue & Clayton theta tree clustering analysis (Yue and Clayton 2005) were also performed in Mothur to identify community membership and structure relationships between the libraries.

Results

Physical and chemical characteristics of samples

The physicochemical characteristics of the waters are listed in Table 1. The physicochemistry of Reigous Creek water at these sampling periods was typical of that observed during a previous long-term monitoring study (Egal et al. 2010). The water samples were acid (pH = 2.91–3.28) and very rich in sulfate (1830–3400 mg L⁻¹), iron (510–1735 mg L⁻¹), and arsenic (70–194 mg L⁻¹), with predominance of the reduced forms Fe(II) and As(III). Dissolved oxygen concentrations ranged from 3.5 to 7.86 mg L⁻¹. The January 2009 sample showed the lowest iron, arsenic, and sulfate concentrations.

The nature and structure of the sediment samples were investigated using mineralogical and spectroscopic methods. XANES analyses at the arsenic K-edge showed that, despite the presence of an As(III) component equal to $12\text{--}34 \pm 5\%$ of total arsenic, the oxidized arsenic form As(V) predominated in all the sediments (Fig. S1). EXAFS data (Fig. S2) showed that As(V) was mainly present in the samples in an amorphous Fe(III)–As(V) hydroxysulfate phase, as previously observed (Morin et al. 2003; Bruneel et al. 2011), As(III) being likely sorbed to poorly ordered schwertmannite. For January 2009, there was not enough time exposure to X-ray beam in EXAFS analysis to record this sample, however, based on the XANES data we can assume that this sample should be similar to the others. XRD analyses (Fig. S3) showed that these arsenic-bearing phases were mixed with sandy components (quartz,

Table 1 Physico-chemical characteristics of the water (mg L^{-1}) during sampling at COWG

Sampling period	pH (\pm SD)	T ($^{\circ}\text{C}$)	DO (\pm SD)	As(III) (\pm SD)	As(V) (\pm SD)	Fe (total) (\pm SD)	Fe(II) (\pm SD)	SO_4^{2-} (\pm SD)
April 2006	3.28 (\pm 0.05)	12.9	3.5 (\pm 0.5)	69 (\pm 3)	71 (\pm 4)	620 (\pm 30)	560 (\pm 28)	2700 (\pm 300)
October 2008	3.13 (\pm 0.05)	14.3	5.7 (\pm 0.1)	133 (\pm 7)	20 (\pm 1)	1250 (\pm 62)	1220 (\pm 61)	3400 (\pm 340)
January 2009	2.91 (\pm 0.05)	9.4	5.5 (\pm 0.5)	43 (\pm 2)	27 (\pm 1)	510 (\pm 25)	540 (\pm 27)	1830 (\pm 183)
November 2009	3.26 (\pm 0.05)	13.1	7.9 (\pm 0.1)	161 (\pm 8)	33 (\pm 2)	1735 (\pm 87)	1440 (\pm 72)	3300 (\pm 330)

SD standard deviation

K-feldspar and micas) and that pyrite was only detected after October 2008.

Diversity analysis

A total of 340 clones obtained from the four independent 16S rRNA gene libraries were fully sequenced and phylogenetically analyzed. Thirteen sequences were identified as likely chimeras and excluded from further analyses. Sequencing and phylogenetic analysis of the 327 remaining cloned sequences led to the identification of 9 and 42 OTUs defined at two different levels of identity (85 and 97 %, respectively). Rarefaction curves calculated at the class level (85 % identity, the rank usually used for representing the microbial community) were near saturation (Fig. 1). Table 2 shows the Shannon, evenness, and Chao1 indices and the coverage values calculated for each library at 97 % identity. The coverage values of the four clone libraries (90, 88, 96 and 92, respectively, for April 2006, October 2008, January 2009, and November 2009) indicate that the clone libraries were sufficiently sampled. The estimations of the diversity indices show that the structure and membership composition of the archaeal community changed over the sampling period. The Shannon diversity (H) and Chao1 richness indices ranged from 1.37 to 2.57 and 6.5 to 30.5, respectively. The diversity ($H = 1.37$) and richness (Chao1 = 6.5) were significantly lower in January 2009 whereas November 2009 displayed the highest values ($H = 2.57$; Chao1 = 30.5), which is consistent with the rarefaction curves.

Comparison of archaeal community

The overall community structure was analyzed for each sample using the Mothur software package. LIBSHUFF analysis was performed to compare the OTU compositions of each clone library revealing a high degree of variation between individuals and showing that with Bonferroni correction, each library differed significantly from all others (Table 3). The resulting dendrograms of Jaccard and Yue & Clayton theta similarity coefficient analysis (Fig. 2) identified one major cluster and one outlier (January 2009). The similarity in community membership (Jaccard index,

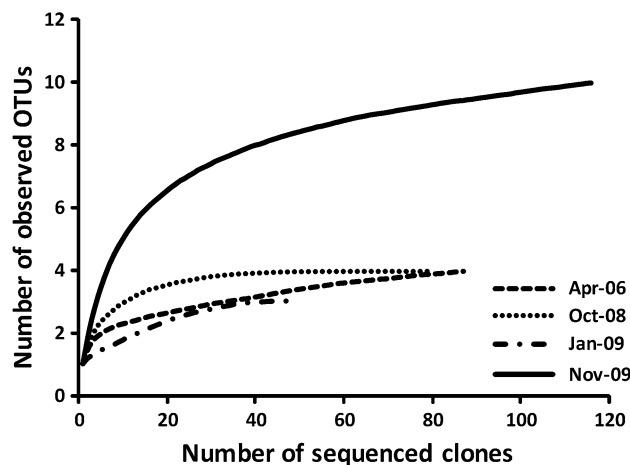


Fig. 1 Rarefaction curves of the archaeal 16S rRNA sequences from Carnoulès mine sediments at 85 % identity. The total number of sequenced clones is plotted against the number of OTUs observed in the same library

Fig. 2a) showed that April 2006 and November 2009 were more related to each other in this respect, whereas April 2006 and October 2008 were more related to each other in terms of community structure (Yue & Clayton index, Fig. 2b).

Phylogenetic analysis of archaeal community

Four 16S rRNA encoding gene libraries were constructed each containing a distinct archaeal community, which varied in composition and richness throughout the sampling campaigns.

In April 2006, the 16S rRNA phylogenetic reconstruction (Fig. 3) showed that all the sequences corresponding to 17 OTUs (OTUs 1–17) were affiliated to the phylum Euryarchaeota, as previously observed in the water samples from Carnoulès (Bruneel et al. 2008). The most abundant OTU (OTU 1, 53 clones representing around 61 % of the sample) was affiliated to the order *Thermoplasmatales* which contained 97 % of the sequences grouped in 15 OTUs. Within this order, the majority of the OTUs were closely related to uncultured clones from an acidic environment such as acidic mine water and sediments (Fig. 3).

Table 2 Diversity indices and statistics calculated for the four clone libraries from COWG station at different sampling periods

Clone library	No. of sequences	No. of OTUs ^a	Singletons	Good's coverage ^b	Shannon diversity ^c	Chao1 richness
April 2006	87	17	9	90	1.63	24.2
October 2008	80	20	10	88	1.96	27.5
January 2009	47	6	2	96	1.37	6.5
November 2009	113	25	9	92	2.57	30.5

^a OTUs were defined at 97 % cutoff

^b Coverage: sum of probabilities of observed classes calculated as $(1 - (n/N))$, where n is the number of singleton sequences and N is the total number of sequences

^c Takes into account the number and evenness of species

Table 3 Community comparison using LIBSHUFF

	Y library			
	Apr-06	Oct-08	Jan-09	Nov-09
X library				
Apr-06	–	0.0260	<0.0001*	0.1529
Oct-08	0.0001*	–	<0.0001*	0.0001*
Jan-09	<0.0001*	<0.0001*	–	0.2222
Nov-09	<0.0001*	<0.0001*	<0.0001*	–

* Significant difference. Bonferroni correction P value = 0.0042

– Not compared

The BLAST affiliation (Table 4) showed that some of these OTUs displayed 89–94 % similarity with *Thermogymnomonas acidicola*, a moderately thermophilic, acidophilic, strictly aerobic heterotroph that uses yeast extract, as well as glucose and mannose (in the presence of yeast extract) as carbon and energy sources (Itoh et al. 2007). Additionally, OTU 15 related to the uncultured clone ORCMO 26 retrieved from a copper mine drainage (Rowe et al. 2007, Fig. 3) was found. This OTU was assigned to methanogenic lineage (*Methanomicrobia*, Fig. 3) with the closest relative *Methanomassiliicoccus luminyensis*, a methanogenic Archaea recently isolated from human faeces (Dridi et al. 2011). However, the Greengenes classification (Table 4) assigned this OTU to the order *Thermoplasmatales*. Lastly, an unknown group belonging to the Euryarchaeota and represented by OTU 8 was detected. This group formed an independent branch that was distantly related to the identified groups and showed low similarity with the uncultured archaeon clone hfm29 isolated from an iron-rich microbial mat (Kato et al. in press).

Twenty OTUs were retrieved from the October 2008 library, 18 of which belonged to the Euryarchaeota and two to the Thaumarchaeota (Fig. 3). Like in April 2006, most of the Euryarchaeota sequences were affiliated with the *Thermoplasmatales* (OTUs 1, 6, 9, 11, 12, 17, 19, 20,

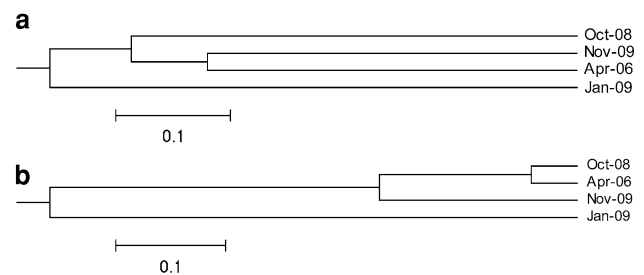


Fig. 2 Similarities in archaeal community membership (Jaccard **a**) and in community structure (Yue & Clayton **b**) between samples. Values are based on 0.03 distances

22–27, and 31) which accounted for 86 % of the total archaeal clones including the same most abundant OTU (OTU1; around 54 %, Fig. 3). The BLAST affiliation (Table 4) also revealed similarity of some OTUs with *Thermogymnomonas acidicola*. Three OTUs (18, 21, and 28) affiliated with uncultured clones isolated from acidic environments (clone SALE1B1 and clone anta6) and from a forested wetland impacted by reject coal (clone ARCP2-12) (Broffft et al. 2002; García-Moyano et al. 2007, Fig. 3), respectively, were assigned to *Methanomicrobia*. The remaining OTUs (OTUs 29 and 30) were affiliated with environmental sequences originating from acidic soil and acidic hot springs, which likely represent uncultured lineages of Thaumarchaeota.

A significant change in the archaeal community appeared in the January 2009 library, when diversity decreased and no cultured species were identified. Indeed, almost all the sequences (96 %) clustered in five OTUs (OTUs 8, 32, 33, 34, and 35, Fig. 3) were related to the uncultured archaeon clone hfmA029 previously found in April 2006. This group formed an independent branch that was far away from the remaining groups. This clone displayed around 97 % similarity with *Methanothermobacter thermautotrophicus*, an autotrophic thermophilic methanogen recovered from an anaerobic sewage sludge digester (Zeikus and Wolee 1972). Remaining sequences grouped



Fig. 3 Maximum likelihood tree of 16S rRNA gene homologs from the archaeal clones (in *bold*) along with a selection of representatives of archaeal diversity. Numbers at nodes indicate a LTR (approximate likelihood ratio test) branch support as computed by PhyML. The

scale bar gives the average number of substitutions per site. The number in parenthesis indicates the number of clones for the sampling period which is represented by a symbol (*star* April 2006, *square* October 2008, *circle* January 2009 and *diamond* November 2009)

Table 4 Identification number of the OTUs retrieved from the Reigous Creek sediment of Carnoulès mine, taxonomic affiliation and representative sequence for each OTU

OTU ID	Number of sequences	Representative sequence	Taxonomic affiliation			Closest relative (% of identity)
			Phylum	Class	Order	
1	132	ArCMSdO8D35	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Aciduliprofundum</i> sp. EPR07-39 (85 %)
2	15	ArCMSdA6A12	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (89 %)
3	2	ArCMSdA6A86	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Aciduliprofundum</i> sp. EPR07-39 (85 %)
4	4	ArCMSdA6A46	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (90 %)
5	1	ArCMSdA6A17	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (91 %)
6	16	ArCMSdA6A67	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (92 %)
7	3	ArCoSdN9H63	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (93 %)
8	24	ArCMSdJ9B78	Euryarchaeota	–	–	Clone hfmA029 (86 %)
9	4	ArCMSdA6A30	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (88 %)
10	3	ArCoSdN9D80	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (94 %)
11	7	ArCMSdO8B50	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (94 %)
12	4	ArCMSdO8B53	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (90 %)
13	1	ArCMSdA6A52	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermoplasma volcanium</i> (84 %)
14	2	ArCMSdA6A84	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (90 %)
15	3	ArCoSdN9H35	Euryarchaeota	<i>Thermoplasmata</i>	–	<i>Methanomassiliicoccus luminyensis</i> B10 (80 %)
16	1	ArCMSdA6A92	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (91 %)
17	4	ArCoSdN9H43	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (88 %)
18	5	ArCMSdO8A3	Euryarchaeota	<i>Thermoplasmata</i>	–	<i>Methanomassiliicoccus luminyensis</i> B10 (82 %)
19	1	ArCMSdO8A13	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (89 %)
20	1	ArCMSdO8A16	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (87 %)
21	2	ArCoSdN9H67	Euryarchaeota	<i>Thermoplasmata</i>	–	<i>Methanomassiliicoccus luminyensis</i> B10 (83 %)
22	1	ArCMSdO8A24	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (88 %)
23	2	ArCMSdO8A56	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (89 %)
24	4	ArCMSdO8A49	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermoplasma volcanium</i> GSS1 (89 %)
25	1	ArCMSdO8A54	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (90 %)
26	3	ArCMSdO8A74	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (99 %)
27	1	ArCMSdO8A85	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (93 %)

Table 4 continued

OTU ID	Number of sequences	Representative sequence	Taxonomic affiliation			Closest relative (% of identity)
			Phylum	Class	Order	
28	1	ArCMSdO8A89	Euryarchaeota	<i>Thermoplasmata</i>	–	<i>Methanomassiliicoccus luminyensis</i> B10 (85 %)
29	5	ArCMSdJ9A29	–	–	–	<i>Candidatus Nitrosocaldus yellowstonii</i> HL72 (84 %)
30	2	ArCMSdO8C25	–	–	–	<i>Candidatus Nitrososphaera gargensis</i> (83 %)
31	1	ArCMSdO8E23	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (92 %)
32	15	ArCMSdJ9C75	Euryarchaeota	–	–	Clone SVB_Fis_02_p137c09 (86 %)
33	16	ArCMSdJ9C68	Euryarchaeota	–	–	Clone hfmA029 (85 %)
34	1	ArCMSdJ9C55	Euryarchaeota	<i>Methanomicrobia</i>	<i>Methanomicrobiales</i>	Clone hfmA029 (84 %)
35	2	ArCoSdN9A45	Euryarchaeota	–	–	Clone hfmA029 (85 %)
36	15	ArCoSdN9B9	Thaumarchaeota	No class	<i>Nitrososphaerales</i>	<i>Candidatus Nitrososphaera</i> sp. EN76 (96 %)
37	6	ArCoSdN9F14	Thaumarchaeota	No class	<i>Cenarchaeales</i>	<i>Candidatus Nitrosopumilus</i> sp. NM25 (93 %)
38	11	ArCoSdN9D53	Euryarchaeota	–	–	Clone TG_FD0.2_SA043 (100 %)
39	2	ArCoSdN9H79	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (90 %)
40	1	ArCoSdN9E14	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (88 %)
41	1	ArCoSdN9G7	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (93 %)
42	1	ArCoSdN9H80	Euryarchaeota	<i>Thermoplasmata</i>	<i>Thermoplasmatales</i>	<i>Thermogymnomonas acidicola</i> JCM 13583 (92 %)

OTU definition and taxonomic identification of representative sequences were done using mothur (Schloss et al. 2009; see “Materials and methods” for details). Only taxonomic affiliations with 100 % similarity are shown. The closest relative was obtained by BLAST search on NCBI nr database

in a single OTU (OTU 29) belonged to an unknown group of Thaumarchaeota previously found in October 2008 and were affiliated with clone GBX-A-COQ1-158 isolated from an acidic hot spring (Fig. 3).

An increase in archaeal diversity was observed in the November 2009 library, with 25 OTUs belonging to the Euryarchaeota (69 clones corresponding to 61 % of the sample) and to the Thaumarchaeota (22 clones corresponding to 19 % of the sample, Fig. 3). The sequences from the Euryarchaeota were distributed in 22 OTUs. Fifteen were related to the order *Thermoplasmatales*, 11 of which (OTUs 1, 2, 4, 6, 7, 9, 10, 11, 17, 24, 26) were previously found in the April 2006 and October 2008 libraries (Fig. 3). As in the results observed in these two sampling periods, OTU 1 was also the most abundant group in the sample in November 2009 (32 %). Additionally, OTUs 15 and 21, also found in the two first libraries, were assigned to the order *Methanomicrobia*. The remaining five OTUs (8, 32, 33, 35, and 38), were not shown to be related to any known species and formed

unknown groups of the Euryarchaeota (Fig. 3). Among these, the most abundant sequences belonging to OTU 38, displayed a strong similarity (99 %) with an uncultured archaeon clone LC15_L00B08 isolated from the monimolimnion of a stratified lake (Gregersen et al. 2009). The four other OTUs (8, 32, 33, and 35) were affiliated with the uncultured archaeon clones hfmA029 mainly present in the January 2009 library. The Thaumarchaeota detected in this study fell into different lineages clustered in three OTUs. The first (OTU 36) belonged to Thaumarchaeota group I.1b and the 15 sequences within this OTU showed from 95 to 96 % similarity with *Candidatus Nitrososphaera viennensis* a chemolithoautotrophic ammonia-oxidizing archaeon (Tourna et al. 2011). The second (OTU 37) was assigned to Thaumarchaeota group I.1a and the sequences displayed 92–93 % similarity with *Candidatus Nitrosopumilus* sp., another ammonia-oxidizing prokaryote (Matsutani et al. 2011). The last OTU (OTU 29), previously found in October 2008 could not be related to any known species.

Discussion

Archaeal 16S rRNA gene analysis of the sediment sampled at the Reigous Creek showed that the Carnoulès archaeal community includes the phylum Euryarchaeota and Thaumarchaeota. The relatively low archaeal diversity revealed by molecular-based methods is consistent with the results of studies in similar environments (Bond et al. 2000; Baker and Banfield 2003; Bruneel et al. 2008; Sánchez-Andrea et al. 2011). This may reflect the limited number of different electron donors and acceptors available in this AMD and the high concentration of toxic compounds along with the low pH. Most of the phylotypes identified in this study were related to genera and species usually found in extreme environments (hot springs, acidic springs, hydrothermal vents, etc.) and showed similarities with sequences obtained in previous studies of Tinto River and other AMD (Sánchez-Andrea et al. 2011; García-Moyano et al. 2007; Rowe et al. 2007, Fig. 3).

Regarding the dynamics of the archaeal community, our study showed that significant modifications in this community occurred throughout the sampling period. All the sampling periods showed differences in community structure and membership although April 2006 and October 2008 were more similar in terms of community structure. Similarity coefficient analysis showed that January 2009 was very different from all the other sampling periods. In January 2009, the archaeal community changed and diversity decreased. Almost all the sequences were related to an uncultured archaeon clone hfmA029 affiliated with methanogenic lineage (*Methanothermobacter thermautotrophicus*). This clone, hfmA029, previously found in April 2006 (OTU 8) in only 2 % of the sample became the dominant population in January 2009. The differences in the archaeal community observed in January 2009 may result from a modification in the composition of the sediment, although the physicochemical analysis of the sediments appeared to be similar throughout the sampling period, and consisted mainly of an amorphous Fe(III)–As(V) hydroxysulfate mineral. Indeed, XRD analyses revealed that pyrite first appeared in October 2008. Likewise, since late 2007, a leakage of fine grey sulfide-rich sands out of the tailings pile has been observed after the rainfall events that generally occur in September and October. This is probably due to the corrosion of the drains at the bottom of the tailing stock that are responsible for the water discharge inside the mine tailing. In January 2009, the sulfide sands, originated from the tailings stock, formed a very thick layer (around 3 cm deep) in the bottom of the creek which could explain the change in the archaeal community.

In the Reigous sediment, most of the sequences were phylogenetically related to the order *Thermoplasmatales*,

although none of the clones could be identified with high similarity (>97 %) as belonging to any cultured species. This order is represented by thermoacidophilic organisms (Reysenbach 2001), which often derive energy from sulfur oxidation or reduction. So far, the order contains three families, each represented by one genus: the *Thermoplasmataceae*, the *Picrophilaceae* and the *Ferroplasmaceae* (Itoh et al. 2007). The *Thermoplasmataceae* comprises species like *Tp. acidophilum* that couple the oxidation of organic carbon with reduction of elemental sulfur, whereas members of the *Ferroplasmaceae* are strict iron-oxidizing chemolithotrophs such as *Ferroplasma acidiphilum* (Itoh et al. 2007). Microorganisms affiliated with methanogenic Archaea such as *Methanomassiliococcus luminyensis* were also identified. Methanogenic communities play an important role in the global carbon cycle, completing the conversion of organic carbon into methane gas by utilizing the metabolic products of bacteria (CO₂, H₂, acetate, and formate) and other simple methyl compounds available in the environment (Sanz et al. 2011). Lastly, we found microorganisms involved in ammonia oxidation, a key step in the nitrogen cycle (Brochier-Armanet et al. 2011), with presence of sequences affiliated to *Candidatus Nitrososphaera viennensis* and *Candidatus Nitrosopumilus* sp. Until recently, ammonia oxidation, the first nitrification step of the nitrogen cycle was thought to be carried out only by autotrophic ammonia-oxidizing bacteria (AOB) belonging to the Beta- and Gammaproteobacteria lineages (Purkhold et al. 2000) occasionally supported by heterotrophic nitrifiers in soil environments (De Boer and Kowalchuk 2001). Ammonia-oxidizing Archaea (AOA) are members of the proposed novel Phylum Thaumarchaeota, and are currently being identified in almost all environments (Brochier-Armanet et al. 2008). These Archaea may thus play a major role in the nitrogen cycle in the Carnoulès sediments.

Previous studies focused on the bacterial communities inhabiting the Carnoulès AMD sediment. These studies showed that the active population of bacteria also contained iron reducers, sulfate-reducing, and sulfur compound oxidizers, and both autotrophic and heterotrophic bacteria (Bruneel et al. 2011). Statistical analysis of genomic and proteomic data demonstrated that both metabolic specificity and partnerships can co-exist in this arsenic-rich sediment (Bertin et al. 2011). These processes include the fixation of inorganic carbon and nitrogen by several strains, in particular those belonging to the *Thiomonas*, *Acidithiobacillus*, and *Gallionella* related genera. However, this study did not find evidence for the presence of archaeal species among the dominant organisms, suggesting that they may represent a small proportion of the microbial community in the sediment. Despite the fact that we cannot really infer the implication of the Archaea

detected in most of these metabolic pathways because most of them could not be affiliated to cultured species, we can point to their probable implication in a specific metabolism currently unknown in bacteria (Forterre et al. 2002), methanogenesis. Archaea are also involved in the nitrogen cycle (*Candidatus Nitrososphaera viennensis* and *Candidatus Nitrosopumilus* sp.) and some of them may also be involved in the sulfur and iron cycles (*Thermoplasmatales*). All these microorganisms may contribute to the remediation process observed in situ and could also be involved in the stability of this sediment by changing the ratio between oxidized and reduced forms of iron, arsenic, and sulfur compounds, promoting the formation and/or dissolution of the Fe(III)–As(V) hydroxysulfate precipitates.

Because isolation and phenotypic characterization of many environmental Archaea are currently not possible, the physiological features and ecological significance of some Archaea detected in this AMD remain difficult to assess. Moreover, the fact that most of the archaeal sequences were only distantly related (<94 % similarity) to known archaeal species suggests that other taxa may exist. Additionally, the contradictions observed in the taxonomic affiliation resulting from the 16S rRNA phylogenetic reconstruction (Fig. 3) and the Greengenes classification (Table 4) suggest that there is still a lack of information making the taxonomic identification difficult to assess. Indeed, almost half of the 16S rRNA gene sequences archived in GenBank database lacks clear taxonomic information (DeSantis et al. 2006). As a consequence, different authors use different names for uncultured clusters which lead to conflicting nomenclatures. Recently developed high-throughput techniques (metagenomics, metaproteomics, and microarrays) may help link the identity of AMD-promoting prokaryotes to their function in mining environments (Mohapatra et al. 2011; Bertin et al. 2011) in the absence of laboratory culture. In the future, these new genomic tools should provide a more precise assessment of the archaeal diversity that will probably lead to substantial changes in current archaeal phylogeny and taxonomy (Brochier-Armanet et al. 2008; Schleper et al. 2005) and to a better understanding of the evolution and metabolic capacities of uncultured Archaea. In conclusion, to increase our insight into the functioning of these highly acidic environments and to elucidate the role of these microorganisms, both improving culture strategies for further physiological and metabolic characterization of newly detected species and a greater sequencing effort are still needed.

Acknowledgments The French CRG is gratefully acknowledged for provision of beamtime on the FAME BM30B beamline. This work was supported by EC2CO CNRS/INSU program, by ACI/FNS Grant

#3033 and by SESAME IdF Grant #1775. Part of the field chemical data was acquired through the OSU OREME.

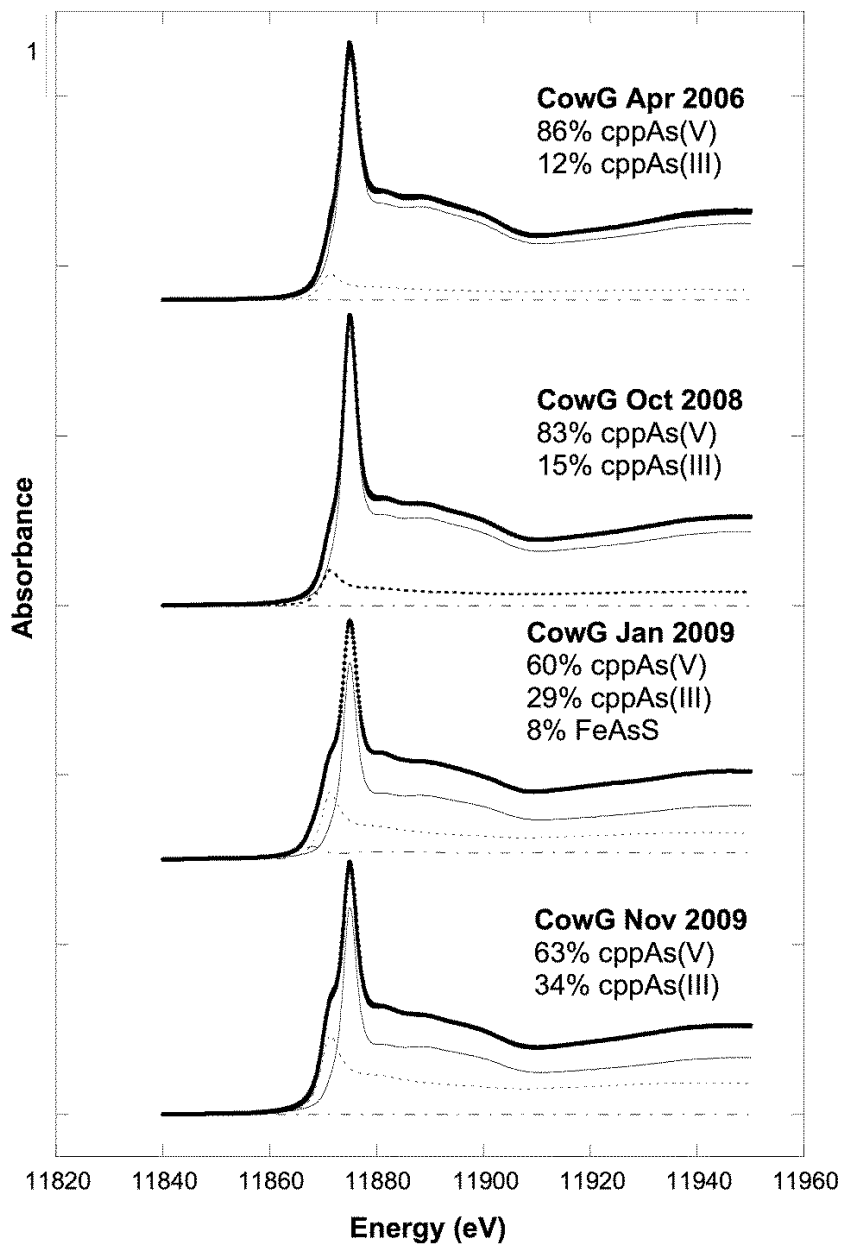
References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
- Anisimova M, Gascuel O (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst Biol* 55(4):539–552
- Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ (2006) New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl Environ Microbiol* 72(9):5734–5741
- Baker BJ, Banfield JF (2003) Microbial communities in acid mine drainage. *FEMS Microbiol Ecol* 44(2):139–152
- Bertin PN, Heinrich-Salmeron A, Pelletier E, Goulhen-Chollet F, Arsène-Ploetze F, Gallien S, Lauga B, Casiot C, Calteau A, Vallenet D, Bonnefoy V, Bruneel O, Chane-Woon-Ming B, Cleiss-Arnold J, Duran R, Elbaz-Poulichet F, Fonknechten N, Giloteaux L, Halter D, Koechler S, Marchal M, Mornico D, Schaeffer C, Smith AAT, Van Dorsselaer A, Weissenbach J, Medigue C, Le Paslier D (2011) Metabolic diversity among main microorganisms inside an arsenic-rich ecosystem revealed by meta- and proteo-genomics. *ISME J* 5(11):1735–1747
- Bond PL, Smriga SP, Banfield JF (2000) Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl Environ Microbiol* 66(9):3842–3849
- Bowell RJ (1994) Sorption of arsenic by iron oxides and oxyhydroxides in soils. *Appl Geochem* 9(3):279–286
- Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* 6(3):245–252
- Brochier-Armanet C, Forterre P, Gribaldo S (2011) Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr Opin Microbiol* 14(3):274–281
- Broffitt JE, McArthur JV, Shimkets LJ (2002) Recovery of novel bacterial diversity from a forested wetland impacted by reject coal. *Environ Microbiol* 4(11):764–769
- Bruneel O, Personne JC, Casiot C, Leblanc M, Elbaz-Poulichet F, Mahler BJ, Le Fleche A, Grimont PA (2003) Mediation of arsenic oxidation by *Thiomonas* sp. in acid-mine drainage (Carnoules, France). *J Appl Microbiol* 95(3):492–499
- Bruneel O, Pascault N, Egal M, Bancon-Montigny C, Goñi-Urriza M, Elbaz-Poulichet F, Personne JC, Duran R (2008) Archaeal diversity in a Fe–As rich acid mine drainage at Carnoules (France). *Extremophiles* 12(4):563–571
- Bruneel O, Volant A, Gallien S, Chaumande B, Casiot C, Carapito C, Bardil A, Morin G, Brown GE, Personne JC, Le Paslier D, Schaeffer C, Van Dorsselaer A, Bertin PN, Elbaz-Poulichet F, Arsène-Ploetze F (2011) Characterization of the active bacterial community involved in natural attenuation processes in arsenic-rich creek sediments. *Microb Ecol* 61(4):793–810
- Casiot C, Morin G, Juillot F, Bruneel O, Personne JC, Leblanc M, Duquesne K, Bonnefoy V, Elbaz-Poulichet F (2003) Bacterial immobilization and oxidation of arsenic in acid mine drainage (Carnoules creek, France). *Water Res* 37(12):2929–2936
- De Boer W, Kowalchuk GA (2001) Nitrification in acid soils: microorganisms and mechanisms. *Soil Biol Biochem* 33(7–8):853–866
- DeLong EF (1992) Archaea in coastal marine environments. *Proc Natl Acad Sci USA* 89(12):5685–5689

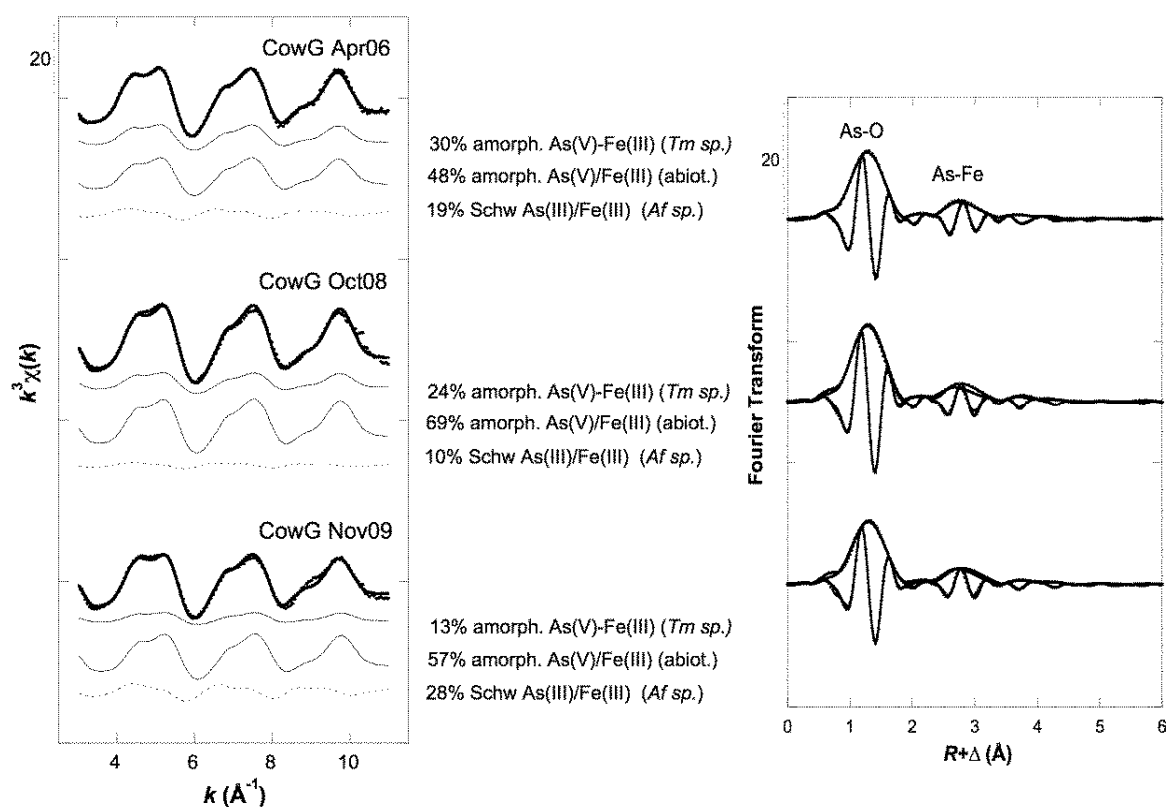
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72(7):5069–5072
- Dopson M, Baker-Austin C, Hind A, Bowman JP, Bond PL (2004) Characterization of *Ferroplasma* isolates and *Ferroplasma acidarmanus* sp. nov., extreme acidophiles from acid mine drainage and industrial bioleaching environments. *Appl Environ Microbiol* 70(4):2079–2088
- Dridi B, Fardeau ML, Ollivier B, Raoult D, Drancourt M (2011) The antimicrobial resistance pattern of cultured human methanogens reflects the unique phylogenetic position of archaea. *J Antimicrob Chemother* 66(9):2038–2044
- Edwards KJ, Goebel BM, Rodgers TM, Schrenk MO, Gihring TM, Cardona MM, Hu B, McGuire MM, Hamers RJ, Pace NR, Banfield JF (1999) Geomicrobiology of pyrite (FeS₂) dissolution: Case study at Iron Mountain, California. *Geomicrobiol J* 16(2):155–179
- Edwards KJ, Bond PL, Gihring TM, Banfield JF (2000) An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science* 287(5459):1796–1799
- Egal M, Casiot C, Morin G, Elbaz-Poulichet F, Cordier MA, Bruneel O (2010) An updated insight into the natural attenuation of As concentrations in Reigous Creek (southern France). *Appl Geochem* 25(12):1949–1957
- Fortere P, Brochier C, Philippe H (2002) Evolution of the Archaea. *Theor Popul Biol* 61(4):409–422
- García-Moyano A, González-Toril E, Aguilera A, Amils R (2007) Prokaryotic community composition and ecology of floating macroscopic filaments from an extreme acidic environment, Río Tinto (SW, Spain). *Syst Appl Microbiol* 30(8):601–614
- Golyshina OV, Pivovarova TA, Karavaiko GI, Kondrateva TF, Moore ER, Abraham WR, Lunsdorf H, Timmis KN, Yakimov MM, Golyshin PN (2000) *Ferroplasma acidiphilum* gen. nov., sp. nov., an acidophilic, autotrophic, ferrous-iron-oxidizing, cell-wall-lacking, mesophilic member of the Ferroplasmaceae fam. nov., comprising a distinct lineage of the Archaea. *Int J Syst Evol Microbiol* 50 Pt 3:997–1006
- Good IJ (1953) The population frequencies of species and the estimation of population parameters. *Biometrika* 40(3/4):237–264
- Gregersen LH, Habicht KS, Peduzzi S, Tonolla M, Canfield DE, Miller M, Cox RP, Frigaard NU (2009) Dominance of a clonal green sulfur bacterial population in a stratified lake. *FEMS Microbiol Ecol* 70(1):30–41
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52(5):696–704
- Hallberg KB (2010) New perspectives in acid mine drainage microbiology. *Hydrometallurgy* 104(3–4):448–453
- Hohmann C, Morin G, Ona-Nguema G, Guigner JM, Brown GE Jr, Kappler A (2011) Molecular-level modes of As binding to Fe(III) (oxyhydr)oxides precipitated by the anaerobic nitrate-reducing Fe(II)-oxidizing *Acidovorax* sp. strain BoFeN1. *Geochim Cosmochim Acta* 75(17):4699–4712
- Huber R, Sacher M, Vollmann A, Huber H, Rose D (2000) Respiration of arsenate and selenate by hyperthermophilic archaea. *Syst Appl Microbiol* 23(3):305–314
- Itoh T, Yoshikawa N, Takashina T (2007) *Thermogymnomonas acidicola* gen. nov., sp. nov., a novel thermoacidophilic, cell wall-less archaeon in the order *Thermoplasmatales*, isolated from a solfataric soil in Hakone, Japan. *Int J Syst Evol Microbiol* 57(11):2557–2561
- Johnson DB, Hallberg KB (2003) The microbiology of acidic mine waters. *Res Microbiol* 154(7):466–473
- Maillot F (2011) Structure locale des nano-oxyhydroxydes de fer(III) de type ferrihydrite et schwertmannite. PhD thesis, UPMC, Paris
- Matsutani N, Nakagawa T, Nakamura K, Takahashi R, Yoshihara K, Tokuyama T (2011) Enrichment of a novel marine ammonia-oxidizing archaeon obtained from sand of an eelgrass zone. *Microbes Environ* 26(1):23–29
- Mohapatra BR, Douglas Gould W, Dinardo O, Koren DW (2011) Tracking the prokaryotic diversity in acid mine drainage-contaminated environments: a review of molecular methods. *Miner Eng* 24(8):709–718
- Morin G, Juillot F, Casiot C, Bruneel O, Personne JC, Elbaz-Poulichet F, Leblanc M, Ildefonse P, Calas G (2003) Bacterial formation of tooleite and mixed arsenic(III) or arsenic(V)-iron(III) gels in the Carnoules acid mine drainage, France. A XANES, XRD, and SEM study. *Environ Sci Technol* 37(9):1705–1712
- Ona-Nguema G, Morin G, Juillot F, Calas G, Brown GE Jr (2005) EXAFS analysis of arsenite adsorption onto two-line ferrihydrite, hematite, goethite, and lepidocrocite. *Environ Sci Technol* 39(23):9147–9155
- Oremland RS, Stolz JF (2003) The ecology of arsenic. *Science* 300(5621):939–944
- Purkhold U, Pommerening-Roser A, Juretschko S, Schmid MC, Koops HP, Wagner M (2000) Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and amoA sequence analysis: implications for molecular diversity surveys. *Appl Environ Microbiol* 66(12):5368–5382
- Reysenbach AL (2001) Order I. *Thermoplasmatales* ord. nov. In: Boone DR, Castenholz RW, Garrity GM (eds) *Bergey's manual of systematic bacteriology*, vol 1, 2nd edn. Springer, New York, p 335
- Rodier J, Broutin JP, Chambon P, Champsaur H, Rodi L (1996) *L'Analyse des Eaux*. Dunod, Paris, p 1383
- Rowe OF, Sánchez España J, Hallberg KB, Johnson DB (2007) Microbial communities and geochemical dynamics in an extremely acidic, metal-rich stream at an abandoned sulfide mine (Huelva, Spain) underpinned by two functional primary production systems. *Environ Microbiol* 9(7):1761–1771
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4):406–425
- Samanta G, Clifford DA (2005) Preservation of inorganic arsenic species in groundwater. *Environ Sci Technol* 39(22):8877–8882
- Sánchez España J, Pamo EL, Pastor ES, Andrés JR, Rubí JAM (2005) The natural attenuation of two acidic effluents in Tharsis and La Zarza-Perrunal mines (Iberian Pyrite Belt, Huelva, Spain). *Environ Geol* 49(2):253–266
- Sánchez-Andrea I, Rodríguez N, Amils R, Sanz JL (2011) Microbial diversity in anaerobic sediments at Río Tinto, a naturally acidic environment with a high heavy metal content. *Appl Environ Microbiol* 77(17):6085–6093
- Sanz JL, Rodríguez N, Diaz EE, Amils R (2011) Methanogenesis in the sediments of Río Tinto, an extreme acidic river. *Environ Microbiol* 13(8):2336–2341
- Schleper C, Jurgens G, Jonuscheit M (2005) Genomic studies of uncultivated archaea. *Nat Rev Microbiol* 3(6):479–488
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: open-Source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75(23):7537–7541
- Singleton DR, Furlong MA, Rathbun SL, Whitman WB (2001) Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. *Appl Environ Microbiol* 67(9):4374–4376

- 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(15):2919–2926
- Tourna M, Stieglmeier M, Spang A, Könneke M, Schintlmeister A, Urich T, Engel M, Schloter M, Wagner M, Richter A, Schleper C (2011) *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci USA* 108(20):8420–8425
- Yue JC, Clayton MK (2005) A similarity measure based on species proportions. *Commun Stat Theory Methods* 34(11):2123–2131
- Zeikus JG, Woolee RS (1972) *Methanobacterium thermoautotrophicus* sp. n., an anaerobic, autotrophic, extreme thermophile. *J Bacteriol* 109(2):707–713

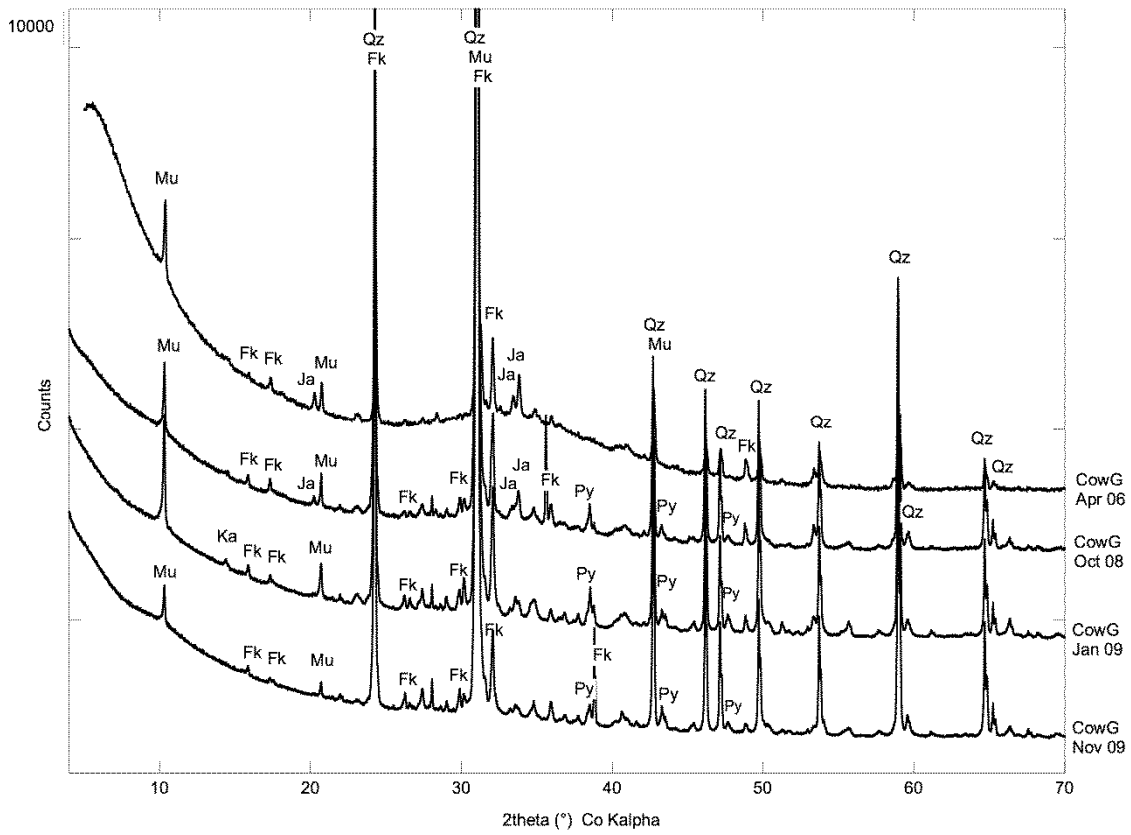
SUPPLEMENTARY MATERIAL



Supplementary Figure S1: Results of As K-edge XANES data fit using linear combination of two model compound spectra representative of arsenate and arsenite. These model compounds are amorphous As(V)- and As(III)-Fe(III) oxyhydroxide phases (referred to as cppAs(V) and cppAs(III), respectively) and characterized in detail in Morin et al. (2003). Accuracy of individual components is estimated to $\pm 5\%$ and components lower than 10% are not significant. A minor component of arsenopyrite was suspected in sample COWG Jan09, while being close to the limit of detection.



Supplementary Figure S2: Results of As K-edge EXAFS data fit using linear combination of three model compound spectra representative of arsenate and arsenite species in the samples studied. These model compounds include biogenic amorphous As(V)-Fe(III) oxyhydroxysulfate synthesized using *Thiomonas sp.* strain B2, abiotic amorphous As(V)-Fe(III) oxyhydroxysulfate with As/Fe = 0.8 mol/mol (Maillot *et al.*, 2011), and biogenic As(III) rich schwertmannite synthesized using *Acidithiobacillus ferrooxidans* strain CC1. Accuracy of individual components is estimated to $\pm 20\%$ and components lower than 15% are not significant.



Supplementary Figure S3: XRD powder data of the solid COWG samples studied. Qz: quartz, Fk: K-feldspar, Mu: muscovite; Ja: jarosite; Py: pyrite.

Chapitre III – Variabilité spatiale de la communauté eucaryote

A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE

1. Contexte général

De nombreuses études se sont intéressées aux communautés procaryotes présentes dans les DMA, avec un intérêt majeur porté aux bactéries. Ces communautés ont été intensivement caractérisées par des méthodes de culture (Edwards *et al.*, 2000b; Bruneel *et al.*, 2003; Anderson & Cook, 2004; Golyshina & Timmis, 2005; Guo *et al.*, 2009; Johnson *et al.*, 2009; Delavat *et al.*, 2012), de génomique basées sur les gènes codant pour l'ARNr 16S (Bond *et al.*, 2000; Baker & Banfield, 2003; Hallberg *et al.*, 2006; Bruneel *et al.*, 2008., 2011; Sánchez-Andrea *et al.*, 2011), de séquençage aléatoire global (Tyson *et al.*, 2004), de métagénomique (Bertin *et al.*, 2011) ou de métaprotéomique (Ram *et al.*, 2005; Bertin *et al.*, 2011; Bruneel *et al.*, 2011). Paradoxalement, les communautés eucaryotes de ces environnements ont été très peu étudiées bien qu'un intérêt croissant leur soit porté de part leur rôle écologique potentiellement important dans ces écosystèmes. En effet, les microorganismes eucaryotes sont susceptibles de modifier l'abondance, la composition et l'activité des communautés microbiennes (bactéries et Archaea) par de nombreux mécanismes comme la prédation, etc. (Baker *et al.*, 2004; Gadanho *et al.*, 2006). Or, les procaryotes sont directement responsables de la génération des DMA *via* l'oxydation du fer et du soufre (Hallberg, 2010). De plus, certains eucaryotes comme les protozoaires sont responsables de la dégradation de la matière organique et jouent donc un rôle important dans le recyclage des nutriments (Zirnstern *et al.*, 2012). D'autre part, alors que les études réalisées sur les communautés procaryotes ont révélé une diversité phylogénétique relativement faible des bactéries et des Archaea, les communautés eucaryotes semblent présenter des niveaux de diversité beaucoup plus variables dépendant des caractéristiques intrinsèques du DMA qu'elles colonisent. Ainsi, le site du Río Tinto (pH ~2) en Espagne, présente une grande diversité de microorganismes eucaryotes parmi lesquels des bacillariophytes (diatomées), des chlorophytes, des euglénides, des rhodophytes, des ciliés, des cercomonades, des amibes, des straménopiles, des champignons et des levures (Amils *et al.*, 2007). Par contre le site de

Richmond aux États-Unis (pH < 0,9), est caractérisé par une faible diversité essentiellement dominée par des champignons et dans une moindre mesure par des protistes affiliés au clade des protistes acidophiles (APC) ainsi que des percolozoaires de la classe des *Heterolobosea* (Baker *et al.*, 2004, 2009). Chaque DMA est donc unique et il apparaît essentiel de continuer les recherches sur ces communautés eucaryotes et leurs relations avec leur environnement afin de comprendre leur physiologie et leur rôle écologique sous ces conditions extrêmes.

Concernant notre site d'étude (*i.e.* la mine de Carnoulès), aucune analyse de diversité concernant la communauté eucaryote n'a encore été réalisée bien que des études précédentes se soient intéressées au protozoaire photosynthétique *Euglena mutabilis* mis en évidence sur ce site (Casiot *et al.*, 2004; Bruneel *et al.*, 2006; Halter *et al.*, 2012a). Ces études ont révélé la capacité de cet organisme à tolérer de fortes concentrations en arsenic (Casiot *et al.*, 2004) tandis que son mécanisme d'adaptation à l'arsenic a été décrit par Halter *et al.* (2012b). Ce protozoaire contribue également à l'élimination du fer *in situ* par séquestration intracellulaire et/ou par la génération d'oxygène par photosynthèse ce qui facilite l'oxydation et la précipitation du Fe(II) (Brake *et al.*, 2002). Les activités biologiques de cet organisme pourraient donc jouer un rôle important dans le processus naturel de remédiation observé sur ce site. D'autre part, une étude récente combinant des approches de protéomique et de métabolomique a permis de caractériser les types de métabolites sécrétés par cet organisme (acides aminés, polyamines, urée, sucres) et de révéler l'absence de production d'acides gras, suggérant ainsi un apport de matière organique sélectif dans cet écosystème (Halter *et al.*, 2012a). Or, une telle production peut avoir un impact crucial sur l'activité métabolique de la communauté bactérienne présente sur le site d'étude, car les bactéries transportent et recyclent *in situ* la plupart des métabolites sécrétés par *Euglena mutabilis* (Bertin *et al.*, 2011). Par conséquent, ce protiste peut exercer un rôle indirect mais néanmoins important, dans les DMA par la production de sources carbonées simples et facilement assimilables par les microorganismes procaryotes (Halter *et al.*, 2012a).

2. Objectif de l'étude

L'objectif de ce travail a été d'identifier les communautés eucaryotes présentes dans les sédiments du Reigous et d'étudier leur distribution spatiale en fonction d'un gradient de pollution. Ces analyses ont été réalisées par pyroséquençage 454 des gènes codant pour l'ARNr 18S. L'originalité de ce travail repose sur le fait que cette identification n'a encore

jamais été faite au niveau de ce site qui présente les caractéristiques uniques citées précédemment. Ce travail avait pour but d'améliorer les connaissances concernant la diversité des microorganismes eucaryotes pour mieux appréhender le rôle qu'ils pourraient avoir dans ce DMA ainsi que les paramètres physicochimiques contrôlant leur répartition spatiale.

B. RÉSULTATS

Les analyses réalisées dans cette étude portaient sur les communautés microbiennes eucaryotes présentes au niveau de cinq sédiments prélevés le long du Reigous. Ces sédiments ont été échantillonnés aux stations S1, COWG, GAL et CONF. A noter que deux échantillons ont été prélevés au niveau de la station GAL car ils présentaient des textures très différentes, GAL de type stromatolithique présentant une structure laminée rigide et GALm, sédiment mou et crémeux de couleur beige.

Diversité et composition des communautés eucaryotes

Les analyses phylogénétiques ont révélé la présence de 14 taxons essentiellement dominés par six groupes (91 % des séquences totales) affiliés aux kingdom/phyla des Ascomycètes, Basidiomycètes, Alveolates, Stramenopiles, Streptophytes et Chlorophytes. Parmi ces groupes, les champignons, constituaient à eux seuls près de 60 % des séquences et sont apparus être le groupe majoritaire sur l'ensemble des sédiments prélevés sur le site de Carnoulès, suivis dans une moindre mesure par les Alveolates et les Stramenopiles. Une diversité relativement élevée de ces champignons a également été mise en évidence par l'identification de 16 classes différentes. Les Alveolates, caractérisés par 8 classes appartenaient à 3 groupes (dinoflagellés, ciliés et apicomplexes) tandis que les Straménopiles étaient apparentés à 4 lignées photosynthétiques (*Bacillariophyceae*, *Chrysophyceae*, *Eustigmatophyceae* et *Synurophyceae*) et deux lignées hétérotrophes (*Bicosoecida* et *Oomycetes*).

Dynamique spatiale et relation avec les paramètres physicochimiques

Les paramètres physicochimiques de l'eau prélevée aux différentes stations le long du Reigous présentaient une variation spatiale traduisant un gradient de toxicité. En effet, une diminution significative des concentrations en arsenic, fer et sulfates associée avec une augmentation de la concentration en oxygène et du pH ont été mises en évidence entre la source (S1, le site le plus extrême) et la station CONF située quelques mètres en amont de la

confluence avec l'Amous (site le moins extrême). De même chacun des sédiments prélevés présentait une composition minéralogique différente. Cependant, bien qu'une variation significative de la structure et de la composition des communautés eucaryotes présentes au niveau de ces stations ait été mise en évidence, aucune relation spécifique ne semblait lier les niveaux de diversité observés et le gradient de pollution. En effet, les communautés les moins diversifiées se trouvaient au niveau des stations S1, GAL et GALm présentant des indices de diversité similaires, tandis que la communauté la plus diversifiée se trouvait au niveau de la station COWG, la station CONF représentant une situation intermédiaire. Néanmoins, la structure de ces communautés eucaryotes s'est révélée être, du moins partiellement, liée à la physicochimie de l'eau. En effet, les concentrations élevées en As(III) et en Fe sont apparues significativement liées aux communautés présentes aux stations S1 et COWG. Parallèlement, le potentiel redox serait également impliqué dans la structuration de la communauté au niveau de CONF tandis qu'aucun paramètre physicochimique particulier influant sur les stations GAL et GALm n'a pu être mis en évidence.

C. DISCUSSION ET CONCLUSION

La majorité des séquences obtenues dans cette étude se sont révélées être apparentées à des taxons trouvés précédemment dans d'autres DMA tels que les Chlorophytes, les Streptophytes, les Bacillariophytes, les Rhizaires, les Alveolates, les Amibes, les Straménoplies et les Champignons (Amils *et al.*, 2007). Ceci implique donc une certaine adaptation de ces organismes à des conditions extrêmes. Le pyroséquençage a également permis de mettre en évidence de nouveaux taxons non détectés auparavant dans ce type de milieu tels que les Apusozoaires, les Centroheliozoaires et les Jakobides. D'autre part, en accord avec des études menées précédemment par Baker *et al.* sur le DMA de la mine de Richmond (2009), les champignons se sont également révélés être le groupe dominant sur l'ensemble des échantillons prélevés bien qu'ils présentaient une diversité relativement plus importante avec un total de 16 classes identifiées. Il est également important de noter que cette étude n'a pas permis de mettre en évidence la présence d'euglènes, connues pour être largement présentes sur ce site. Elles forment en effet des biofilms très épais, visibles presque toute l'année et particulièrement abondant à COWG (Casiot *et al.*, 2004; Halter *et al.*, 2011). Cette absence est probablement due au fait qu'un seul type d'amorce ne peut raisonnablement pas amplifier l'ADN de l'ensemble des eucaryotes (Caron *et al.*, 2004), suggérant ainsi un biais probable dans l'estimation de la diversité eucaryote sur le site de Carnoulès.

Ces travaux ont également permis de mettre en évidence une variation spatiale des communautés eucaryotes avec la présence de communautés distinctes au sein des sédiments prélevés le long du Reigous. Cette structuration spatiale des communautés s'est révélée être, du moins partiellement, liée à la physicochimie de l'eau. En effet, les analyses ont montré que l'arsénite, le fer et le potentiel redox pourraient être des facteurs significativement responsables de la structuration des communautés eucaryotes de Carnoulès. Cependant, contrairement à ce qui a été observé au niveau du Río Tinto par Amaral-Zettler *et al.* (2011), les sites les plus extrêmes de Carnoulès ne présentaient pas la diversité eucaryote la plus importante. En effet, le niveau de diversité des communautés eucaryotes ne semblait pas suivre le gradient de pollution (*i.e.* diversité la plus importante à COWG), suggérant ainsi la présence de facteurs supplémentaires contribuant à structurer ces communautés.

Les résultats obtenus dans cette étude ont apporté des connaissances nouvelles sur le DMA de Carnoulès en nous éclairant pour la première fois sur les microorganismes eucaryotes présents sur ce site. Ces microorganismes peuvent potentiellement jouer un rôle important dans les DMA. Ainsi, la caractérisation de leur diversité ne représentait qu'une première étape pour nous aider à améliorer nos connaissances sur le fonctionnement du DMA de Carnoulès et notamment les interactions existant entre les communautés eucaryotes et les cycles biogéochimiques. Il sera en effet nécessaire de compléter ces analyses par le développement des techniques de culture et d'isolement de ces organismes eucaryotes afin de mieux caractériser les espèces présentes, leur mode de nutrition (autotrophie, hétérotrophie, mixotrophie) ainsi que leurs interactions entre elles et avec leur environnement.

D. CONTRIBUTION SCIENTIFIQUE

Ce travail, finalisé sous forme d'article sera prochainement soumis dans la revue « Environmental Microbiology » (Article C).

Volant, A., Héry, M., Desoeuvre, A., Casiot, C., Morin, G., Delpoux, S., Bertin, P. N., Elbaz-Poulichet, F., and Bruneel, O. Spatial distribution of eukaryotic communities along a pollution gradient in the arsenic-rich creek sediments of Carnoulès Mine, France.

Chapitre III – Article C

Spatial distribution of eukaryotic communities along a pollution gradient in the arsenic-rich creek sediments of Carnoulès Mine, France

A. Volant, M. Héry, A. Desoeuvre, C. Casiot, G. Morin, S. Delpoux, P. N. Bertin, F. Elbaz-Poulichet, O. Bruneel

1 **Spatial distribution of eukaryotic communities along a pollution gradient in**
2 **the arsenic-rich creek sediments of Carnoulès Mine, France**

3
4
5 A. Volant¹, M. Héry¹, A. Desoeuvre¹, C. Casiot¹, G. Morin², S. Delpoux¹, P. N. Bertin³,
6 F. Elbaz-Poulichet¹ and O. Bruneel¹

7
8
9 ¹. Laboratoire HydroSciences Montpellier, HSM, UMR 5569 (IRD, CNRS, Universités
10 Montpellier 1 et 2), Université Montpellier 2, Place E. Bataillon, CC MSE, 34095
11 Montpellier, France

12
13 ². Institut de Minéralogie et de Physique des Milieux Condensés, IMPMC, UMR 7590
14 (CNRS, Université Pierre et marie curie/Paris 6), 4 place Jussieu, 75252 Paris, France

15
16 ³. Laboratoire de Génétique Moléculaire, Génomique, Microbiologie, GMGM, UMR 7156
17 (Université de Strasbourg, CNRS), Département Microorganismes, Génomes,
18 Environnement, 28 Rue Goethe, 67083 Strasbourg, France

19
20
21 *Corresponding author. Aurélie VOLANT. Laboratoire HydroSciences Montpellier,
22 UMR5569, Université Montpellier 2, Place E. Bataillon, CC MSE, 34095 Montpellier,
23 France. Phone: (+33)4 67 14 90 19; Fax: (+33)4 67 14 47 74. Email:
24 aurelie.volant@msem.univ-montp2.fr

25 **Summary**

26 Microscopic eukaryotes are key factors in ecosystem functioning, but their diversity in the
27 majority of environments remains widely unexplored. The Carnoulès mine (France) is
28 characterized by acidic waters containing very high concentrations of arsenic (80 to 350 mg
29 L⁻¹), iron (750 to 2700 mg L⁻¹), and sulfate (2000 to 7500 mg L⁻¹). During the first 30 m of
30 downflow in Reigous Creek, emerging from the mine tailings, 20 to 60% of the dissolved
31 arsenic is removed by coprecipitation with Fe(III). Previous studies, mainly focusing on
32 bacteria and archaea suggested a role of prokaryotic activities in the coprecipitation of As(III)
33 and As(V) with Fe(III) and sulfate. This study analyzed for the first time the eukaryotic
34 diversity present in the sediment along the 1500 m of the Reigous Creek, draining the site. A
35 bar-coded pyrosequencing approach targeting the hypervariable V2-V3 regions of the
36 eukaryotic small-subunit 18S rRNA gene was used to characterize the diversity of eukaryotic
37 communities along the pollution gradient of arsenic-rich Creek sediment and to address the
38 factors that structure these communities. The pyrosequencing approach yielded a total of
39 8688 reads leading to the identification of 865 OTUs. Phylogenetic analyses revealed that
40 around 60% of the sequences belonged to Fungi. The majority of sequences (91%) clustered
41 into six main groups falling into the kingdom/phyla Ascomycota, Basidiomycota, Alveolata,
42 Stramenopiles, Streptophyta and Chlorophyta. As(III) concentration was highlighted as
43 significant factor driving the eukaryotic communities structure, followed by Fe concentration
44 and redox potential.

45 Introduction

46 Acid Mine Drainages (AMD) are extreme aquatic environments generated from the
47 dissolution of pyrite-rich mining wastes. Because of their acidic pH and their high
48 concentrations of toxic metals and metalloids, they represent a threat to public health and the
49 environment (Motsi *et al.*, 2009). In these ecosystems, adapted microorganisms are involved
50 in biogeochemical mechanisms controlling metal and metalloids behavior at the
51 water/sediment interface. Indeed, microorganisms colonizing AMD are responsible for metals
52 and metalloids release and acidification consecutive to pyrite-rich waste oxidation. Microbial
53 metabolisms can also lead to the pollution mitigation, for example by coprecipitation of
54 arsenic with iron (Morin *et al.*, 2003). Prokaryotic communities colonizing AMD, and
55 particularly bacteria, have been extensively studied and are generally characterized by a low
56 diversity including organisms involved in iron, sulfur or arsenic cycling (Johnson and
57 Hallberg, 2003; Bruneel *et al.*, 2011; Johnson, 2012; Volant *et al.*, 2012). However,
58 investigations about eukaryotic members of the microbial communities remain relatively
59 scarce and their potential involvement in biogeochemical cycling in these extremes
60 environments is poorly understood (Aguilera *et al.*, 2007; Amils *et al.*, 2007; Baker *et al.*,
61 2009, Amaral-Zettler *et al.*, 2011). Culture-independent gene surveys of AMD-impacted
62 phototrophic communities (pH 2) from the Río Tinto River in Spain have revealed high
63 eukaryotic richness including members of the Bacillariophyta, Chlorophyta, Rhodophyta and
64 Euglenophyta phyla as well as Ciliates, Cercomonads, Amoebae, Stramenopiles, Fungi,
65 Heliozoans and Rotifers (Aguilera *et al.*, 2007; Amils *et al.*, 2007). In contrast, some studies
66 carried out on the Richmond mine, an extremely acidic (pH < 1), subsurface AMD site at Iron
67 Mountain in California, have evidenced a low eukaryotic diversity mainly dominated by
68 Fungi associated with the presence, in small proportion, of *Heterolobosea* and protists
69 affiliated with the deeply branching lineage named the acidophilic protist clade (APC) (Baker
70 *et al.*, 2009). Some protists isolated from AMD graze on specific acidophilic bacteria
71 (McGinness and Johnson, 1992; Baker *et al.*, 2004), suggesting that they may impact the
72 abundance and the community composition of prokaryotes known to be directly involved in
73 acid generation rates *via* the oxidation of iron and sulfur compounds. Various species of algae
74 and Fungi encountered in these extreme environments effectively enrich the carbon sources
75 that help to maintain some populations like the Sulfate Reducing Bacteria (SRB) in
76 predominantly anaerobic environments (Das *et al.*, 2009). Fungi also show capacity to absorb
77 significant amount of metals in their cell wall, or by extracellular polysaccharide slime

78 limiting their biodisponibility in the environment (Das *et al.*, 2009). In spite of their potential
79 importance, the full diversity and the ecological role of eukaryotes in AMD remain poorly
80 understood as well as the underlying mechanisms that structure their communities.

81 In Carnoulès (France), the level of pollution is particularly high (up to 10 g L⁻¹ of As
82 in the tailing stock pore water) and from 80 to 350 g L⁻¹ at the spring of the Reigous Creek,
83 draining the site (Casiot *et al.*, 2003). Previous studies indicate that microbial activity is
84 involved in the natural attenuation process occurring at this site, leading to a pollution
85 gradient along the 1500 meters of the creek corresponding to the AMD. Indeed, the arsenic
86 concentration decreases within the first 30 m of the Reigous Creek due to microbial iron
87 oxidation which leads to the removal of 20–60% of dissolved arsenic by coprecipitation with
88 Fe(III) and formation of As(III)-Fe(III) and As(V)-Fe(III) phases associated with sulfate
89 (Casiot *et al.*, 2003; Morin *et al.*, 2003). Although the arsenic level remains high, its
90 concentration subsequently decreases by around 95% between the source of Reigous Creek
91 and its confluence with the Amous River, 1500 meters downstream. If bacterial and archaeal
92 communities have been characterized, to date, the eukaryotic community thriving at this site
93 remains unexplored although some studies revealed the presence of the acidophilic protozoan
94 *Euglena mutabilis* (Casiot *et al.*, 2004). This photosynthetic organism may play an important
95 role in the natural recovery of contaminants, by increasing the concentration of dissolved
96 oxygen in the water, which enhances the rate of inorganic precipitation of reduced metal
97 species, and by sequestering and storing iron intracellularly (Brake *et al.*, 2001, Casiot *et al.*,
98 2004).

99 The first steps to better assess the potential influence of eukaryotes on metals and
100 metalloids behavior and on Carnoulès ecosystem functioning is the determination of the
101 eukaryotic diversity present. For this purpose, eukaryotic diversity in Carnoulès sediments
102 was investigated using a bar-coded pyrosequencing approach targeting the hypervariable V2-
103 V3 regions of the eukaryotic small-subunit 18S rRNA gene. In order to help understanding
104 how environmental stressors shape microbial community structure, the eukaryotic spatial
105 pattern was characterized in relation with the distinct physicochemical conditions observed *in*
106 *situ* along the pollution gradient.

107 **Results**

108 ***Physicochemical characteristics of the samples***

109 The main physicochemical characteristics of the water samples collected from the
110 different stations are listed in Table 1. The Reigous was characterized by low pH (around 3),
111 remaining rather constant along the creek. Redox potential (Eh) showed an increasing trend
112 with values ranging from 476 to 715 mV. Extremely high concentrations of As (222 mg L^{-1})
113 and iron (1904 mg L^{-1}) were recorded in the upstream station S1 in relation with the oxidation
114 of As-rich pyrite within the tailings impoundment (Casiot *et al.* 2003). This station was
115 almost anoxic and displayed high concentration of sulfate (2226 mg L^{-1}). However, natural
116 remediation occurred along the creek characterized by a clear decreasing gradient of arsenic
117 ($222\text{-}0.32 \text{ mg L}^{-1}$), iron ($1904\text{-}23 \text{ mg L}^{-1}$), and sulfate ($2226\text{-}853 \text{ mg L}^{-1}$) associated with a
118 general increase of dissolved oxygen concentration ($0.5\text{-}9.4 \text{ mg L}^{-1}$). The reduced forms
119 Fe(II) and As(III) were dominant at all stations except at the downstream station CONF which
120 was the most oxygenated and showed the lowest concentrations of iron, sulfate and arsenic.

121 The nature and structure of the sediment samples were investigated using
122 mineralogical and spectroscopic methods. XRD analyses (Fig. S1) showed that the
123 mineralogical composition evolves from mainly tooeleite (To: $\text{Fe}_6(\text{AsO}_3)_4\text{SO}_4(\text{OH})_4 \cdot 4\text{H}_2\text{O}$)
124 upstream (sample S1: spring), to amorphous ferric arsenate hydroxysulfate (samples GAL and
125 GALm), and to schwertmannite (Schw: $\text{Fe}_8\text{O}_8(\text{OH})_6\text{SO}_4$) (sample CONF). Sample COWG is
126 sandy with quartz (Qz) and micas (Mi). XANES analyses for COWG, GAL and CONF at the
127 arsenic K-edge, showed that despite the presence of an As(III) component accounting for 23–
128 $34 \pm 2\%$ of total arsenic, the oxidized arsenic form As(V) predominated in these samples (Fig.
129 S2). Similar observations would be expected for sample GALm. In contrast, the proportion of
130 As(III) is expected to be much larger in sample S1, which essentially consists of tooeleite, a
131 rare iron-arsenite mineral found in large amounts in the Carnoulès sediments (Morin *et al.*,
132 2003). Recording of XANES and EXAFS data for these latter samples is under progress.
133 EXAFS data (Fig. S3) showed that As(V) was mainly present in the samples in an amorphous
134 Fe(III)–As(V) hydroxysulfate phase, as previously observed (Morin *et al.*, 2003; Bruneel *et*
135 *al.*, 2011), As(III) being likely sorbed to poorly ordered schwertmannite.

136

137 ***Diversity and species richness estimators of eukaryotic communities***

138 A total of 16 510 sequence reads with an average length of 504 bp were generated in a
139 single run of 454 pyrosequencing from the five independent 18S rRNA libraries. After

140 trimming and processing with Mothur, 15 061 reads remained (495 bp average length), 42%
141 of which were sequences classified as Metazoa and Coniferospida (Table 2). This study
142 focusing on microorganisms, these sequences were removed from the dataset. However, they
143 accounted for highly variable percentages of eukaryotic sequences from the different locations
144 (75%, 3%, 74%, 73% and 2%, respectively, for S1, COWG, GAL, GALm and CONF)
145 leading to a significant decrease in the number of sequences in some samples. Clustering of
146 the 8688 remaining sequences led to the identification of 865 OTUs (including 413
147 singletons) defined at 97% of identity.

148 Rarefaction curves based on the full dataset did not reach an asymptote (Fig. 1),
149 especially for COWG, suggesting that a large number of unseen OTUs still existed in the
150 original samples and that more sequencing efforts may be required to detect additional
151 phylotypes. However, Good's coverage (Table 2), which provides an estimate of sampling
152 completeness using a probability calculation with randomly selected amplicon sequences, was
153 high ranging from 92 to 98%, suggesting that the majority of eukaryotic phylotypes present in
154 each sample were identified. The total number of OTUs for each sample, as estimated by the
155 Chao1 estimator for the full dataset, was considerably higher than the observed number of
156 OTUs, covering an average of 55% of the estimated richness (Table 2). The diversity indices
157 showed that the structure and membership composition of the eukaryotic community varied
158 strongly across the stream. The Shannon diversity (H) and Chao1 richness indices ranged
159 from 2.13 to 5.18 and 65 to 1124, respectively. While S1, GAL and GALm displayed a
160 similar level of diversity, COWG displayed the highest diversity (H = 5.18; Chao1 = 1124),
161 CONF representing an intermediate situation, which is consistent with the rarefaction curves.
162 Thus, there seemed to be no clear relationship between the eukaryotic diversity and the
163 pollution gradient observed along the Reigous creek. Rank-abundance curves, calculated for
164 the 50 most abundant OTUs in the full dataset, indicated that a high majority of the OTUs
165 belonged to rare organisms represented by only a few sequences, whereas a relatively low
166 proportion of OTUs were highly abundant (Fig. 2). The same patterns of diversity occurred
167 for both full and resampled data sets (Table 2).

168

169 *Taxonomic assignment of eukaryotic pyrosequencing reads*

170 Table 3 shows the relative abundance of taxa in each sample. About 5.6% of the total
171 sequences could not be assigned to a taxonomic group. The highest proportion of unassigned
172 sequences occurred in COWG (7.8%) (*i.e.* the more diverse sample), with the values ranging

173 from 0 to 5.9% for the other samples. Across all the samples, 14 high-level eukaryotic
174 lineages were identified. However the majority of sequences (91%) clustered into six main
175 groups falling into the kingdom/phyla Ascomycota (39.7%), Basidiomycota (17.4%),
176 Alveolata (11.6%), Stramenopiles (10%), Streptophyta (7.7%), and Chlorophyta (4.6%)
177 (Table 3). The eight additional groups were each represented by less than 1% of the sequences
178 and belonged to Chytridiomycota, Blastocladiomycota, Choanoflagellida, Rhizaria,
179 Amoebozoa, Centrohelioczoa, Apusozoa and Jakobida. Out of these 14 lineages collectively
180 identified, five were detected in all five samples (*i. e.* Ascomycota, Basidiomycota, Alveolata,
181 Stramenopiles, and Streptophyta), one in only four samples, one in only three samples, two in
182 only two samples, and five in only one sample. Fungi were clearly dominant in all the
183 samples comprising about 84%, 49%, 83%, 67% and 59% of the sequences, respectively, for
184 S1, COWG, GAL, GALm and CONF. The Basidiomycota were dominant in S1 (53.9%),
185 GAL (66%) and GALm (50.8%) whereas both COWG (40.6%) and CONF (55.5%) showed a
186 major proportion of Ascomycota. A high diversity of Fungi was observed with 16 classes
187 encountered within this kingdom (Fig. 3A). The Ascomycota fell into six classes mainly
188 represented by *Sordariomycetes* and *Leotiomycetes* whereas eight classes were identified for
189 the Basidiomycota with the predominance of *Microbotryomycetes* and *Agaricomycetes*. Only
190 one class was identified within the Blastocladiomycota and Chytridiomycota. Within the
191 Alveolata we retrieved eight classes falling into three groups (Dinoflagellida, Coliophora and
192 Apicomplexa) (Fig. 3B). The Stramenopiles were related to four photosynthetic lineages
193 (*Bacillariophyceae*, *Chrysophyceae*, *Eustigmatophyceae*, and *Synurophyceae*) and two
194 heterotrophic lineages (*Bicosoecida* and *Oomycetes*) (Fig. 3C).

195

196 ***Community membership and spatial distribution***

197 To explore the extent of spatial variation of community patterns we compared
198 eukaryotic communities based on community membership. Fig. 4 depicts a Venn diagram
199 representing the unique and overlapping OTUs at the five sites calculated from the
200 normalized dataset. Across the 301 OTUs retrieved from all the reduced samples, a relatively
201 high percentage of OTUs in each sample were unique (38%, 83%, 51%, 32% and 67%,
202 respectively, for S1, COWG, GAL, GALm and CONF) and only one OTU corresponding to
203 47 sequences was shared between the five samples, indicating that an important turnover of
204 the eukaryotic community occurred along the stream (Table 2, Fig. 4). This OTU showed
205 99% of similarity with *Escovopsis* sp. Esc19, an Ascomycota belonging to the

206 Sordariomycetes class, known to be involved in symbiosis with fungus-growing ants (Currie
207 *et al.*, 2003). The two samples GAL and GALm, retrieved from distinct sediments located at
208 the same site in the creek, shared only 10 OTUs corresponding to 27% and 40% of their
209 respective taxa highlighting the probable influence of the sediment characteristics on
210 eukaryotic membership patterns. Moreover, while the percentage of shared OTUs between the
211 five sites was very low (15% shared between two or more site), the majority of the total
212 sequences (68%) were shared between at least two sites (Fig. 4). This disparity between the
213 percentage of shared sequences and the percentage of shared OTUs between the five sites
214 suggest that members with low abundance could play a disproportionate role in the level of
215 beta diversity.

216

217 ***Community structure and relationship with environmental parameters***

218 Across all the samples, the Bray-Curtis dissimilarity index ranged from 0.63 to 0.94
219 indicating a high level of betadiversity (data not shown). According to the Mantel test, the
220 distance matrices corresponding to eukaryotic community structure (based on the normalized
221 abundance of all OTUs) and environmental factors were significantly correlated ($r = 0.59$, $P <$
222 0.05). To disentangle the environmental factors shaping the eukaryotic community,
223 Nonmetric MultiDimensional Scaling (NMDS) based on normalized abundance of OTUs was
224 used to show the relationship between diversity patterns of microbial communities and
225 environmental parameters. The NMDS ordination (Fig. 5) showed a clear separation of the
226 eukaryotic communities by the pollution gradient along the second ordination axis, with the
227 most polluted sites (S1 and COWG) opposed to the less polluted site CONF. Communities
228 from sites GAL and GALm showed similar structures by clustering together between these
229 two extremes. Vector fitting of possible explanatory environmental variables suggested
230 As(III) concentration as most powerful predictor separating eukaryotic communities at
231 different sites ($r^2=0.99$, $p=0.025$), followed by redox potential ($r^2=0.96$, $p=0.017$) and total Fe
232 concentration ($r^2=0.94$, $p=0.041$) (Table 1 and Fig. 5). Thus, the structure of the eukaryotic
233 community found in CONF was correlated with high values of redox potential whereas high
234 concentration of As(III) and Fe mainly shaped the communities in S1 and COWG. However,
235 no particular parameters seemed to influence the communities occurring in the stations GAL
236 and GALm as shown by their central position along the second axis.

237 Discussion

238 *Diversity patterns and methodological considerations*

239 The aim of this study was to characterize the diversity of microbial eukaryotes and
240 their spatial variations within five sediment stations along the Reigous Creek of the Carnoulès
241 mine (France), an AMD heavily contaminated with As. This mine drainage is characterized
242 by a gradient of pollution, with a decrease of As concentration of around 95% between the
243 Reigous Spring and its confluence with the Amous River, 1.5 km downstream. For this
244 purpose, pyrosequencing of 18S rRNA gene hypervariable regions was performed using
245 universal eukaryote set of primers (Casamayor *et al.*, 2002; López-García *et al.*, 2003;
246 Monchy *et al.*, 2011). Although rarefaction curves did not reach an asymptote, the coverage
247 values for each sediment, exceeded 90%, suggesting a comprehensive sampling of the
248 dominant microbial groups. Clustering of the 8688 sequences found in this study led to the
249 identification of 865 OTUs. Most of the sequences retrieved from this extreme environment
250 corresponded to lineages previously found in others AMD, including members of the
251 Chlorophyta, Streptophyta and Bacillariophyta phyla as well as Rhizaria, Alveolata,
252 Amoebozoa, Stramenopiles and Fungi (Amils *et al.*, 2007). However, some taxa detected in
253 Carnoulès had not been previously found in any AMD system such as Apusozoa,
254 Centroheliozoa, Jakobida and Choanoflagellida. In biofilms from the AMD of Richmond
255 mine (California) a low eukaryotic diversity was highlighted, mainly dominated by two
256 classes of Fungi (the *Dothideomycetes* and *Eurotiomycetes*), some Heteroblosea and APC
257 (Baker *et al.*, 2009). In accordance with this study, Fungi also represented the most dominant
258 group in all the Carnoulès samples but they exhibit a higher diversity with 16 different classes
259 identified. In addition we found a more diverse eukaryotic diversity characterized by a total of
260 12 lineages. Amaral-Zettler *et al.*, (2002, 2003) also showed a high eukaryotic species
261 richness in the Río Tinto River in Spain. This high diversity observed in both Carnoulès and
262 the Río Tinto may be, at least partially, explained by the presence of photoautotrophic
263 eukaryotes that may sustain the development of diverse heterotrophic lineages. However, it
264 should be noted that pyrosequencing of rRNA gene hypervariable region allowed to reveal
265 much greater eukaryote diversity than classical approach developed previously for the study
266 of Río Tinto (Amaral-Zettler *et al.*, 2003; Aguilera *et al.*, 2007) and Richmond AMD (Baker
267 *et al.*, 2009) which tend to unveil the dominant rarest ones. On the 865 OTUs identified in our
268 study, 413 were singletons. Rare species correspond to low abundance organisms of the “rare
269 biosphere”. The rare microbial eukaryotes are largely underexplored but may play an

270 important role in AMD by becoming dominant in response to environmental changes and
271 buffering the effects of dramatic environmental shifts and could also open great perspectives
272 for future identification of novel “species”, genes and metabolic pathways (Sogin *et al.*, 2006;
273 Monchy *et al.*, 2011). However, the deep sequencing raised also questions about the accuracy
274 of OTU richness estimates, as sequencing errors coupled with inadequate choice of clustering
275 algorithms can lead to artificially elevated estimates of community richness (Huse *et al.*,
276 2010).

277 In this study, we did not encounter *Euglena mutabilis*, a photosynthetic acidophilic
278 protist known to be widely present in this ecosystem, forming continuous, green, macroscopic
279 biofilms, and particularly abundant at the sampling station COWG (Casiot *et al.*, 2004, Halter
280 *et al.*, 2011). We probably captured only a portion of the entire eukaryotic diversity present in
281 Carnoulès AMD, due to the fact that a single PCR primer set does not amplify the 18S rRNA
282 from all Eukaryotes (Caron *et al.*, 2004). This bias has been highlighted before and to
283 accurately determine environmental diversity, next generation sequencing may need to be
284 completed by a multiple PCR primers set approach (Stoeck *et al.*, 2006). Furthermore,
285 different targeted SSU rRNA gene sequences might potentially not be PCR-amplified with
286 equal efficiency by the different primers, and variability within the amplified region may have
287 a great effect on the apparent species richness (Casamayor *et al.*, 2002). In addition, richness
288 should also be temperate by some technological limits like differential cell lysis, alignments
289 of large sets of sequences, pyrosequencing errors, statistical tools etc. (Huse *et al.*, 2011) and
290 by the unknown variability of the multiple copy of rDNA operon from the same organism
291 (Monchy *et al.*, 2011). Pyrosequencing can thereby only be considered as a semiquantitative
292 measurement.

293

294 ***Ecological role***

295 The different eukaryotic populations found in the Reigous Creek can be grouped
296 according to their ecological role as primary producers, decomposers, and consumers.
297 Photosynthetic eukaryotes act as the primary producers of biomass in this ecosystem and
298 mainly belong to the Chlorophyta (green algae), Streptophyta and some autotrophic classes of
299 the Stramenopiles such as the *Bacillariophyceae* (diatoms), *Chrysophyceae* (golden algae),
300 *Eustigmatophyceae* and *Synurophyceae*. In addition to the diffusion of atmospheric oxygen
301 along the stream, the oxygen produced by the photosynthetic organisms must favor the
302 aerobic oxidation of iron and sulfur, and the subsequent microbial transition described above.

303 Fungi showed a high abundance and diversity in all the sampling stations, were they are
304 decomposers and main producers of inorganic substances. In waters affected by AMD, the
305 content of easily degradable organic carbon is often very low, and then Fungi may play an
306 important role as primary degraders of complex organic matter. Furthermore, it was suggested
307 that fungal filaments belonging to the *Eurotiomycetes* and *Dothideomycetes*, could provide
308 rigidity and organization to the biofilms and anchor them to pyrite sediment in flowing AMD
309 solutions (Baker *et al.*, 2004). In addition, they keep organic carbon levels low and produce
310 dissolved carbonate ions, which are likely important for the growth of chemolithoautotrophic
311 acidophilic prokaryotes (Baker *et al.*, 2004). Fungi can also adsorb metals in their cell wall or
312 adsorb in extracellular polysaccharide slime (Zafar *et al.*, 2007). This capacity enables them
313 to grow in the presence of high amounts of heavy metals. Such sequestration of potentially
314 toxic metals and metalloids could allow less tolerant species to exist where they might not
315 have otherwise, possibly enhancing biodiversity in these areas (Amaral Zettler *et al.*, 2003).
316 The mixotrophic community is dominated by Cercozoa and Stramenopiles related to the
317 *Chlamydomyridae*, *Cryomonadida*, *Oomycetes* and *Bicosoecida*. The consumer community
318 is colonized by heterotrophic protozoa characterized by representative of the Dinoflagellida,
319 Choanoflagellida, Ciliophora, Apicomplexa, Amebozoa, Apusozoa and Centrohelioczoa. In
320 AMD, protozoa can graze on acidophilic bacteria, suggesting that they may control bacterial
321 populations and biomass. Moreover, grazing by protozoa stimulates the rate of decomposition
322 of organic matter (Finlay and Esteban, 1998). Our results confirm that diverse eukaryotes
323 could colonize extreme environments such as AMD were they probably play a substantial role
324 particularly in carbon cycling (Baker *et al.*, 2009).

325

326 ***Spatial distribution***

327 The spatial distribution of taxa within the five sediments along the creek should be
328 interpreted with care since no biological repeats were performed. The sediment samples were
329 retrieved from different locations along the AMD reflecting distinct environmental
330 conditions. Indeed, the water chemistry varied significantly, with a clear decreasing gradient
331 of As, Fe and sulfates associated with an increasing gradient of oxygen content and pH,
332 between the spring (S1, the most extreme site) and the confluence with the Amous (CONF,
333 the less extreme site). Moreover, the sediment sample compositions were different according
334 to the DRX analyses, including the two samples collected in the same area (GAL and
335 GALm). As highlighted by the Bray-curtis dissimilarity index, eukaryotic communities along

336 the Reigous were also very different and distinct eukaryotic communities (at the OTU level)
337 were observed. Although the phylogenetic distribution of sequences varied among samples
338 and sites, some taxa were clearly dominant in all or most of them. The Reigous sediments
339 appeared to form a succession of distinct ecosystems housing distinct eukaryotic communities
340 whose structure was at least partially linked to water chemistry. Indeed, arsenite concentration
341 in water was identified to be significantly correlated to the community composition. Redox
342 potential and iron content also revealed plausible predictive power as further driving factors.
343 Contrary to what was observed in the Río Tinto by Amaral-Zettler *et al.*, (2011), the
344 eukaryotic richness was not highest at the most extreme sites of Carnoulès. Indeed, the level
345 of eukaryotic diversity did not follow the water pollution gradient along the creek, suggesting
346 that additional factors than water chemistry may contribute to shape the eukaryotic
347 communities. The eukaryotic complexity encountered at COWG could be, for example,
348 indicative of mature and well-established communities able to sustain more or less complex
349 trophic interactions. This sampling station is localized in a small bowl and is protected by a
350 low wall. During intensive precipitations, characterizing the Mediterranean climate of this
351 region, important flooding was observed, able to take the sediments away and thus to
352 completely wash the bottom of the creek, except for this point, at least in part.

353 Moreover, Metazoan sequences represented a relative high proportion of
354 pyrosequencing reads. As metazoans contribute many cells and, consequently, more DNA per
355 individual, these sequences may represent an important bias in gene libraries (Ragon *et al.*,
356 2012). They accounted for highly variable percentages of eukaryotic sequences and they are
357 particularly abundant from the sampling point S1 (75%), GAL (74%) and GALm (73%). This
358 important number of sequences can be explained in part by the presence of many trees in the
359 area around S1 station and the presence of a very small creek, crossing the Reigous Creek at
360 the station GAL and characterized by less acidic conditions (pH 5-6) and less extreme
361 concentrations of arsenic and toxic metalloids (Casiot, personal communication). Physical
362 and chemical factors are known to influence the species community structure and their spatial
363 variability. The characterization of the biodiversity is only a first step to understand the
364 functioning of an ecosystem. To gain some insights into interactions between eukaryotic
365 communities and the biogeochemical cycling in Carnoulès, this study needs to be completed
366 by further characterizations of the species, their mode of nutrition (autotrophy, heterotrophy,
367 mixotrophy), their interactions between each other (food web, saprotrophism, parasitism,

368 commensalism etc.) and their interactions with metal and metalloid cycling. For this purpose,
369 improving culture isolation would be necessary.

370

371 **Experimental procedures**

372 *Description of the study site*

373 The Pb-Zn Carnoulès mine is located in southern France in the drainage basin of the
374 Rhône River. Mining activity stopped in 1962 and left about 1.2 Mt of sulfidic wastes
375 containing 0.7% Pb, 10% Fe, and 0.2% As which are stored behind a dam. Water percolating
376 through the wastes emerges at the base of the tailings dam and forms the head of the Reigous
377 Creek. This water presents oxygen concentration below 1 mg L⁻¹, acidic pH (around 3), and is
378 rich in dissolved sulfate, iron and arsenic (2000-7700, 500-1000 and 50-350 mg L⁻¹,
379 respectively) the later being predominantly in reduced forms: Fe(II) and As(III) (Casiot *et al.*,
380 2003). The mean annual rainfall is 1320 mm of rain with a distribution typical of the
381 Mediterranean climate characterized by long drought periods and intense rainfall events
382 occurring mainly in autumn and spring.

383

384 *Sampling procedure and measurement of physicochemical properties*

385 A sampling campaign was carried out in November 2009 at five sampling stations
386 along the Reigous Creek (Fig. 6). Samples were collected at the spring (S1), 30 m
387 downstream from the spring (COWG), 150 m downstream at two sampling points
388 characterized by a different sediment texture (stromatolitic like structure with laminated hard
389 concretions called GAL and soft and creamy sediment called GALm), and 1500 m
390 downstream (CONF) just before the confluence between the Reigous and the Amous River.
391 Sediments were collected in triplicates at the bottom of the creek using a sterile spatula. The
392 triplicates were pooled and homogenized. Solid phases were harvested by centrifugation and
393 dried under vacuum prior to mineralogical and spectroscopic analyses. Measurements of the
394 main physicochemical parameters of the water (pH, T°C, redox potential and dissolved
395 oxygen concentration) and determination of SO₄²⁻, Fe and As concentrations were done as
396 previously described (Bruneel *et al.* 2011).

397 *Solid sample characterization*

398 The mineralogical composition of the collected sediments samples was qualitatively
399 determined using powder X-ray diffraction analyses (XRD). Data were collected with Co K-
400 alpha radiation on an X'Pert PRO P analytical diffractometer equipped with an X'Celerator

401 detector, in continuous mode and a 4 h counting time of per sample. X-ray absorption
402 spectroscopy data were collected on the sediment phases sampled at CONF, GAL, GALm and
403 S1. X-ray absorption near edge structure (XANES) and extended X-ray absorption fine
404 structure (EXAFS) spectra were recorded at a temperature of 10–15 K in fluorescence mode
405 on the FAME BM30B bending magnet beamline at ESRF (Grenoble, France). Data for the
406 COWG sample were previously collected and analyzed in Volant *et al.*, (2012). Experimental
407 details and data reduction procedures are reported in previous studies (Morin *et al.*, 2003;
408 Ona-Nguema *et al.*, 2005; Hohmann *et al.*, 2011). XANES and EXAFS data were interpreted
409 by linear combination fitting using a set of model compound spectra. This set includes As(V)-
410 and As(III)–Fe(III) oxyhydroxides and oxyhydroxysulfates synthesized via biotic and abiotic
411 pathways (Morin *et al.*, 2003; Maillot *et al.*, submitted).

412

413 ***DNA isolation***

414 Triplicates genomic DNA were extracted from sediments using the UltraClean Soil
415 DNA Isolation Kit according to the manufacturer's recommendations (MoBio Laboratories
416 Inc., Carlsbad, CA, USA). These triplicates were pooled before PCR amplification. All
417 extracted genomic DNA samples were stored at -20 °C until further analysis.

418

419 ***Pyrosequencing of eukaryotic 18S rRNA fragments***

420 The 18S rRNA genes were amplified by PCR for multiplexed pyrosequencing using
421 bar-coded primers. The set of primers, 18S-82F (5'-AxxxGAAACTGCGAATGGCTC-3')
422 and Ek-516r Ek-516r (5'-BxxxACCAGACTTGCC CTCC-3') was designed to amplify a 480
423 bp region containing the complete V2 and V3 domains of all eukaryotes SSU rDNA genes
424 (Casamayor *et al.*, 2002; López-García *et al.*, 2003; Monchy *et al.*, 2011). A and B represent
425 the two FLX Titanium adapters (A adapter sequence: 5'-CGTATCGCCTCCCTCGCGCC
426 ATCAG-3'; B adapter sequence: 5'-CTATGCGCCTTGCCAGCCCGCTCAG-3') and xxx
427 represent the sample-specific bar-code sequence. Triplicate PCR amplifications with the same
428 barcode were performed for each sample. The reaction mixture contained 1 µl of DNA
429 template, 1 µl of both primers (10 µM) and 12.5 µl of PCR Master Mix Ampli Taq Gold 360
430 (Applied Biosystems, Foster City, CA, USA). Sterile distilled water was added to reach a
431 final volume of 25 µl. The PCR conditions used were 95 °C for 10 min, 35 cycles at 95 °C for
432 45 s (denaturation), 56 °C for 45 s (annealing) and 72 °C for 45 s (extension), followed by 10
433 min at 72 °C. The PCR products with an approximate length of 480 bp were excised from a

434 1% agarose gel and purified with QIAquick gel extraction kit (QIAGEN Inc., Valencia, CA,
435 USA). Triplicates were pooled prior to pyrosequencing to minimize random PCR bias. DNA
436 concentrations of the purified PCR products were measured using a Qubit fluorometer
437 (Invitrogen, Carlsbad, CA, USA). The products were mixed at equimolar concentrations to
438 obtain a total of 500 ng of DNA per sample. Pyrosequencing of the five amplicon libraries
439 was performed by a Roche 454 Life Sciences Genome Sequencer FLX Titanium sequencer at
440 GATC Biotech (Konstanz, Germany).

441

442 ***Processing pyrosequencing data and taxonomic classification***

443 Preliminary quality checks, sorting, and trimming of the 454-reads were performed by
444 GATC Biotech according to the manufacturer's recommendations. The sequences were then
445 analyzed with the software program Mothur version 1.26.0 (<http://www.mothur.org>) (Schloss
446 *et al.*, 2009). Preprocessing of unaligned sequences included removing of short sequences <
447 550 pb, all sequences containing ambiguous characters, and sequences with more than eight
448 homopolymers. We further removed sequences that did not align over the same span of
449 nucleotide positions. Identical sequences were grouped and representative sequences were
450 aligned against the SILVA eukaryotic reference database using the Needleman–Wunsch
451 algorithm (Needleman and Wunsch, 1970). Chimeric sequences were detected and removed
452 using the implementation of Chimera Uchime (Edgar *et al.*, 2011). A further screening step
453 (pre-cluster) was applied to reduce sequencing noise by clustering reads differing by only 1
454 base every 100 bases (Huse *et al.*, 2010). The remaining high-quality reads were used to
455 generate a distance matrix and clustering into operational taxonomic units (OTUs) defined at
456 97% cutoff using the average neighbor algorithm. Next, the taxonomic affiliation of the OTUs
457 was determined using the Naïve Bayesian Classifier trained on the RDP's taxonomic outline
458 implemented in Mothur with a confidence threshold of 80% and a modified SILVA
459 eukaryotic database. Note that OTUs classified as Metazoa and Coniferospida were excluded
460 from further analyses to reduce the number of non microbial eukaryotes.

461

462 ***Statistical analyses***

463 OTU based diversity indices Chao1 and Shannon (H), as well as the coverage and
464 rarefaction curves for each sample were calculated with Mothur at a level of 97% sequence
465 similarity. We performed these calculations using full data sets and using reduced data sets in
466 which the number of sequences per sample was made equal through random resampling (565

467 sequences). To estimate community similarity among samples, a distance matrix was
468 performed based on the Bray-Curtis dissimilarities of the OTUs composition of all sampling
469 sites. To down-weight the influence of rare OTUs, OTUs occurring as singletons were
470 removed from the dataset as suggested by Behnke *et al.*, (2011). The normalized abundances
471 of OTUs were square root transformed before the analysis. A second dissimilarity matrix was
472 performed based on Euclidean distance of $\log(N+1)$ environmental variables (Table 1).
473 Mantel test was used on these matrices to determine the relationships between the eukaryotic
474 assemblage structure and the environmental factors (Mantel *et al.*, 1967). To identify potential
475 explanatory variables environmental vector fitting was applied and projected to the NMDS
476 ordination of normalized OTUs abundances. In addition to pair wise comparison, Venn
477 diagrams for graphical descriptions of unshared, as well as shared OTUs between two, three,
478 four or five sample were constructed.

479

480 **Acknowledgments**

481 The study was financed by the “Observatoire de Recherche Méditerranéen en Environnement”
482 (OSU-OREME). Aurélie Volant were supported by a grant from the French Ministry of
483 Education and Research. This work was performed in the framework of the “Groupement de
484 recherche: Métabolisme de l’Arsenic chez les Microorganismes” (GDR2909-CNRS).

485

486 **References**

- 487 Aguilera, A., Zettler, E., Gómez, F., Amaral-Zettler, L., Rodríguez, N. and Amils, R. (2007)
488 Distribution and seasonal variability in the benthic eukaryotic community of Río Tinto
489 (SW, Spain), an acidic, high metal extreme environment. *Syst Appl Microbiol* **30**: 531-546.
- 490 Amaral-Zettler, L. A., Gomez, F., Zettler, E., Keenan, B. G., Amils, R. and Sogin, M. L.
491 (2002) Microbiology: Eukaryotic diversity in Spain's river of fire. *Nature* **417**: 137-137.
- 492 Amaral-Zettler, L. A., Messerli, M. A., Laatsch, A. D., Smith, P. J. and Sogin, M. L. (2003)
493 From genes to genomes: beyond biodiversity in Spain's Río Tinto. *Biol Bull* **204**: 205-209.
- 494 Amaral-Zettler, L. A., Zettler, E. R., Theroux, S. M., Palacios, C., Aguilera, A. and Amils, R.
495 (2011) Microbial community structure across the tree of life in the extreme Río Tinto.
496 *ISME J* **5**: 42-50.
- 497 Amils, R., González-Toril, E., Fernández-Remolar, D., Gómez, F., Aguilera, Á., Rodríguez,
498 N., *et al.* (2007) Extreme environments as Mars terrestrial analogs: The Río Tinto case.
499 *Planet Space Sci* **55**: 370-381.
- 500 Baker, B. J., Lutz, M. A., Dawson, S. C., Bond, P. L. and Banfield, J. F. (2004) Metabolically
501 active eukaryotic communities in extremely acidic mine drainage. *Appl Environ Microbiol*
502 **70**: 6264-6271.

- 503 Baker, B. J., Tyson, G. W., Goosherst, L. and Banfield, J. F. (2009) Insights into the diversity
504 of eukaryotes in acid mine drainage biofilm communities. *Appl Environ Microbiol* **75**:
505 2192-2199.
- 506 Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R. R. and Stoeck, T. (2011) Depicting
507 more accurate pictures of protistan community complexity using pyrosequencing of
508 hypervariable SSU rRNA gene regions. *Environ Microbiol* **13**: 340-349.
- 509 Brake, S. S., Dannelly, H. K. and Connors, K. A. (2001) Controls on the nature and
510 distribution of an alga in coal mine-waste environments and its potential impact on water
511 quality. *Environ Geol* **40**: 458-469.
- 512 Bruneel, O., Volant, A., Gallien, S., Chaumande, B., Casiot, C., Carapito, C., *et al.* (2011)
513 Characterization of the active bacterial community involved in natural attenuation
514 processes in arsenic-rich creek sediments. *Microb Ecol* **61**: 793-810.
- 515 Caron, D. A., Countway, P. D. and Brown, M. V. (2004) The growing contributions of
516 molecular biology and immunology to protistan ecology: molecular signatures as
517 ecological tools. *J Eukaryot Microbiol* **51**: 38-48.
- 518 Casamayor, E. O., Massana, R., Benlloch, S., Ovreas, L., Diez, B., Goddard, V. J., *et al.*
519 (2002) Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient
520 by comparison of genetic fingerprinting methods in a multipond solar saltern. *Environ*
521 *Microbiol* **4**: 338-348.
- 522 Casiot, C., Morin, G., Juillot, F., Bruneel, O., Personné, J.C., Leblanc, M., *et al.* (2003)
523 Bacterial immobilization and oxidation of arsenic in acid mine drainage (Carnoulès creek,
524 France). *Water Res* **37**: 2929-2936.
- 525 Casiot, C., Bruneel, O., Personné, J. C., Leblanc, M. and Elbaz-Poulichet, F. (2004) Arsenic
526 oxidation and bioaccumulation by the acidophilic protozoan, *Euglena mutabilis*, in acid
527 mine drainage (Carnoules, France). *Sci Total Environ* **320**: 259-267.
- 528 Currie, C. R., Wong, B., Stuart, A. E., Schultz, T. R., Rehner, S. A., Mueller, U. G., *et al.*
529 (2003) Ancient tripartite coevolution in the attine ant-microbe symbiosis. *Science* **299**:
530 386-388.
- 531 Das, B. K., Roy, A., Koschorreck, M., Mandal, S. M., Wendt-Potthoff, K. and Bhattacharya,
532 J. (2009) Occurrence and role of algae and fungi in acid mine drainage environment with
533 special reference to metals and sulfate immobilization. *Water Res* **43**: 883-894.
- 534 Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. and Knight, R. (2011) UCHIME
535 improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-2200.
- 536 Egal, M., Casiot, C., Morin, G., Parmentier, M., Bruneel, O., Lebrun, S. and Elbaz-Poulichet,
537 F. (2009) Kinetic control on the formation of tooeleite, schwertmannite and jarosite by
538 *Acidithiobacillus ferrooxidans* strains in an As(III)-rich acid mine water. *Chem Geol* **265**:
539 432-441.
- 540 Finlay, B. J. and Esteban, G. F. (1998) Freshwater protozoa: biodiversity and ecological
541 function. *Biodivers Conserv* **7**: 1163-1186.
- 542 Halter, D., Casiot, C., Heipieper, H. J., Plewniak, F., Marchal, M., Simon, S., *et al.* (2012)
543 Surface properties and intracellular speciation revealed an original adaptive mechanism to
544 arsenic in the acid mine drainage bio-indicator *Euglena mutabilis*. *Appl Microbiol*
545 *Biotechnol* **93**: 1735-1744.

546 Hohmann, C., Morin, G., Ona-Nguema, G., Guigner, J.-M., Brown Jr, G. E. and Kappler, A.
547 (2011) Molecular-level modes of As binding to Fe(III) (oxyhydr)oxides precipitated by the
548 anaerobic nitrate-reducing Fe(II)-oxidizing *Acidovorax* sp. strain BoFeN1. *Geochim*
549 *Cosmochim Acta* **75**: 4699-4712.

550 Huse, S. M., Welch, D. M., Morrison, H. G. and Sogin, M. L. (2010) Ironing out the wrinkles
551 in the rare biosphere through improved OTU clustering. *Environ Microbiol* **12**: 1889-1898.

552 Johnson, D. B. and Hallberg, K. B. (2003) The microbiology of acidic mine waters. *Res*
553 *Microbiol* **154**: 466-473.

554 Johnson, D. B. (2012) Geomicrobiology of extremely acidic subsurface environments. *FEMS*
555 *Microbiol Ecol* **81**: 2-12.

556 López-García, P., Philippe, H., Gail, F. and Moreira, D. (2003) Autochthonous eukaryotic
557 diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic
558 Ridge. *Proc Natl Acad Sci USA* **100**: 697-702.

559 Maillot, F., Morin, G., Juillot, F., Bruneel, O., Casiot, C., Ona-Nguema, *et al.* Structure and
560 reactivity of As(III)- and As(V)-rich schwertmannites and ferric arsenate sulfate from the
561 Carnoulès acid mine drainage, France: comparison with biotic and abiotic model
562 compounds and implications for As remediation. Submitted to *Geochim Cosmochim Acta*

563 Mantel, N. (1967) The detection of disease clustering and a generalized regression approach.
564 *Cancer Res* **27**: 209-220.

565 McGinness, S. and Johnson, D. B. (1992) Grazing of acidophilic bacteria by a flagellated
566 protozoan. *Microb Ecol* **23**: 75-86.

567 Monchy, S., Sancier, G., Jobard, M. and other authors (2011) Exploring and quantifying
568 fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU
569 tag pyrosequencing. *Environ Microbiol* **13**: 1433-1453.

570 Morin, G., Juillot, F., Casiot, C., Bruneel, O., Personné, J. C., Elbaz-Poulichet, F., *et al.*
571 (2003) Bacterial formation of tooeleite and mixed arsenic(III) or arsenic(V)-iron(III) gels
572 in the Carnoules acid mine drainage, France. A XANES, XRD, and SEM study. *Environ*
573 *Sci Technol* **37**: 1705-1712.

574 Motsi, T., Rowson, N. A. and Simmons, M. J. H. (2009) Adsorption of heavy metals from
575 acid mine drainage by natural zeolite. *Int J Miner Process* **92**: 42-48.

576 Needleman, S. B. and Wunsch, C. D. (1970) A general method applicable to the search for
577 similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**: 443-453.

578 Ona-Nguema, G., Morin, G., Juillot, F., Calas, G. and Brown, G. E., Jr. (2005) EXAFS
579 analysis of arsenite adsorption onto two-line ferrihydrite, hematite, goethite, and
580 lepidocrocite. *Environ Sci Technol* **39**: 9147-9155.

581 Ragon, M., Fontaine, M. C., Moreira, D. and Lopez-Garcia, P. (2012) Different
582 biogeographic patterns of prokaryotes and microbial eukaryotes in epilithic biofilms. *Mol*
583 *Ecol* **21**: 3852-3868.

584 Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., *et al.*
585 (2009) Introducing mothur: open-source, platform-independent, community-supported
586 software for describing and comparing microbial communities. *Appl Environ Microbiol*
587 **75**: 7537-7541.

- 588 Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., Huse, S. M., Neal, P. R., *et al.*
589 (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc*
590 *Natl Acad Sci U S A* **103** : 12115-12120.
- 591 Stoeck, T., Hayward, B., Taylor, G. T., Varela, R. and Epstein, S. S. (2006) A multiple PCR-
592 primer approach to access the microeukaryotic diversity in environmental samples. *Protist*
593 **157**: 31-43.
- 594 Volant, A., Desoeuvre, A., Casiot, C., Lauga, B., Delpoux, S., Morin, G., *et al.* (2012)
595 Archaeal diversity: temporal variation in the arsenic-rich creek sediments of Carnoulès
596 Mine, France. *Extremophiles* **16**: 645-657.
- 597 Zafar, S., Aqil, F. and Ahmad, I. (2007) Metal tolerance and biosorption potential of
598 filamentous fungi isolated from metal contaminated agricultural soil. *Bioresour Technol*
599 **98**: 2557-2561.

Table 1. Physicochemical characteristics of the water body (mg L^{-1}) measured directly above the sediment at the sampling stations; r^2 – correlation coefficient and corresponding p-value indicating goodness of fit and significance determined by vector fitting to the NMDS (bold font with *: $p < 0.05$).

Sampling stations	S1	COWG	GAL/ GALm	CONF	Vector fit	
					r^2	p
Location	44°07'02.16"N 4°00'08.92"W	44°07'01.80"N 4°00'06.90"W	44°06'59.90"N 4°00'01.97"W	44°06'24.75"N 3°59'37.67"W		
T (°C)	14.3	13.1	11.2	11.2	0.65	0.262
Eh (mV)	476	542	560	715	0.96	0.017*
pH (\pm SD)	3.3 (\pm 0.05)	3.3 (\pm 0.05)	3.1 (\pm 0.05)	3.2 (\pm 0.05)	0.52	0.487
DO (\pm SD)	0.5 (\pm 0.005)	7.9 (\pm 0.1)	8.5 (\pm 0.1)	9.4 (\pm 0.1)	0.36	0.436
SO ₄ ²⁻ (\pm SD)	2226 (\pm 223)	3300(\pm 330)	3004 (\pm 300)	853 (\pm 85)	0.69	0.341
As(V) (\pm SD)	9 (\pm 0.05)	33(\pm 2)	12 (\pm 1)	0.3 (\pm 0.02)	0.75	0.257
As(III) (\pm SD)	213 (\pm 11)	161 (\pm 8)	40 (\pm 2)	0.02 (\pm 0.001)	0.99	0.025*
Fe(II) (\pm SD)	1289 (\pm 64)	1440 (\pm 72)	968 (\pm 48)	5 (\pm 0.25)	0.90	0.094
Fe(total) (\pm SD)	1904 (\pm 95)	1735 (\pm 87)	1063 (\pm 53)	23 (\pm 1)	0.94	0.041*

SD: standard deviation

Table 2. Diversity indices and statistics calculated for the five sampling stations.

Sampling stations	TOTAL DATASET							NORMALIZED DATASET					
	No. of sequences after quality filtering	No. of sequences excluding Metazoa and Coniferopsida	No. of OTUs	Singletons	Good's coverage	Shannon diversity	Chao1 richness	No. of sequences	No. of OTUs	Singletons	Good's coverage	Shannon diversity	Chao1 richness
S1	3203	792	42	14	98	2.21	65	565	29	6	99	2.10	37
COWG	4319	4183	697	326	92	5.18	1124	565	215	135	76	4.68	466
GAL	2268	583	55	22	96	2.13	88	565	37	4	99	1.96	38
GALm	2887	786	41	19	98	2.31	98	565	25	5	99	2.19	28
CONF	2384	2344	124	53	98	3.31	277	565	61	19	97	3.18	95

Table 3. Distribution of sequences of high-level taxonomic groups in the five samples for the full dataset. Number in brackets represents the relative abundance (%) of taxa. Others Fungi correspond to Blastocladiomycota and unclassified Fungi.

Kingdoms/phyla	Total*	S1	COWG	GAL	GALm	CONF
Fungi; Ascomycota	3451 (39.7)	237 (29.9)	1698 (40.6)	91 (15.6)	123 (15.6)	1302 (55.5)
Fungi; Basidiomycota	1512 (17.4)	427 (53.9)	216 (5.2)	385 (66)	399 (50.8)	85 (3.6)
Alveolata	1007 (11.6)	58 (7.3)	834 (19.9)	10 (1.7)	1 (0.1)	104 (4.4)
Stramenopiles	867 (10)	9 (1.1)	457 (10.9)	27 (4.6)	53 (6.7)	321 (13.7)
Viridiplantae; Streptophyta	669 (7.7)	26 (3.3)	172 (4.1)	41(7)	202 (25.7)	228 (9.7)
Viridiplantae; Chlorophyta	402 (4.6)	30 (3.8)	214 (5.1)	0	0	158 (6.7)
Fungi; others	115 (1.3)	2 (0.3)	101 (2.4)	8 (1.4)	0	4 (0.2)
Fungi; Chytridiomycota	57 (0.7)	0	49 (1.2)	0	8 (1)	0
Choanoflagellida	37 (0.4)	0	37 (0.9)	0	0	0
Rhizaria	37 (0.4)	0	37 (0.9)	0	0	0
Amoebozoa	25 (0.3)	0	22 (0.5)	0	0	3 (0.1)
Centroheliozoa	9 (0.1)	0	9 (0.2)	0	0	0
Apusozoa	10 (0.1)	0	10 (0.2)	0	0	0
Jakobida	1 (<0.1)	0	1 (<0.1)	0	0	0
Unclassified	489 (5.6)	3 (0.4)	326 (7.8)	21 (3.6)	0	139 (5.9)

*Calculated relatively to the total number of sequences retrieved in this study.

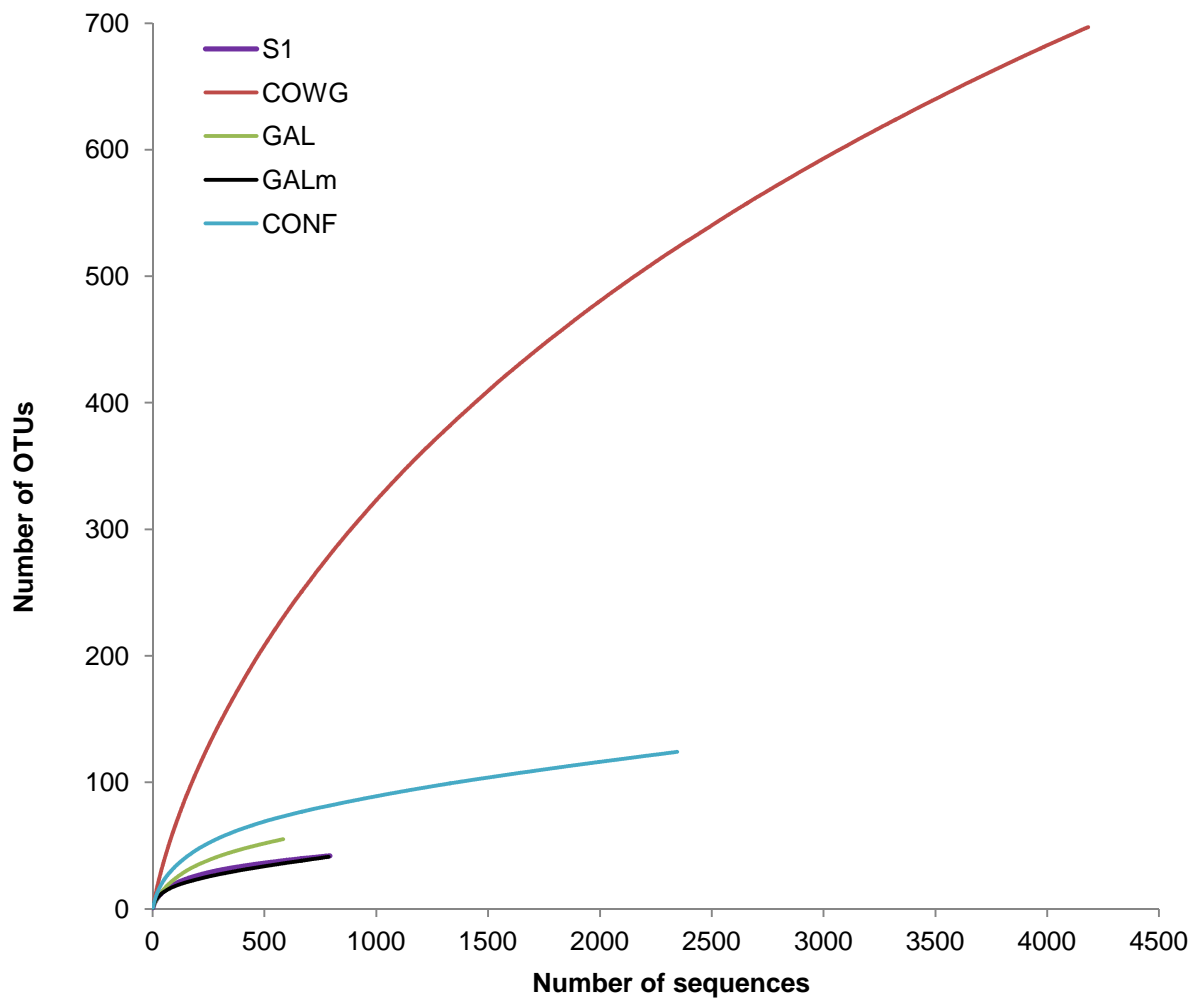


Fig. 1. Rarefaction curves of the eukaryotic 18S rRNA sequences from Carnoulès mine sediments based on OTUs calculated at 97% identity for the full dataset. The total number of sequences analyzed is plotted against the number of OTUs observed in the same library.

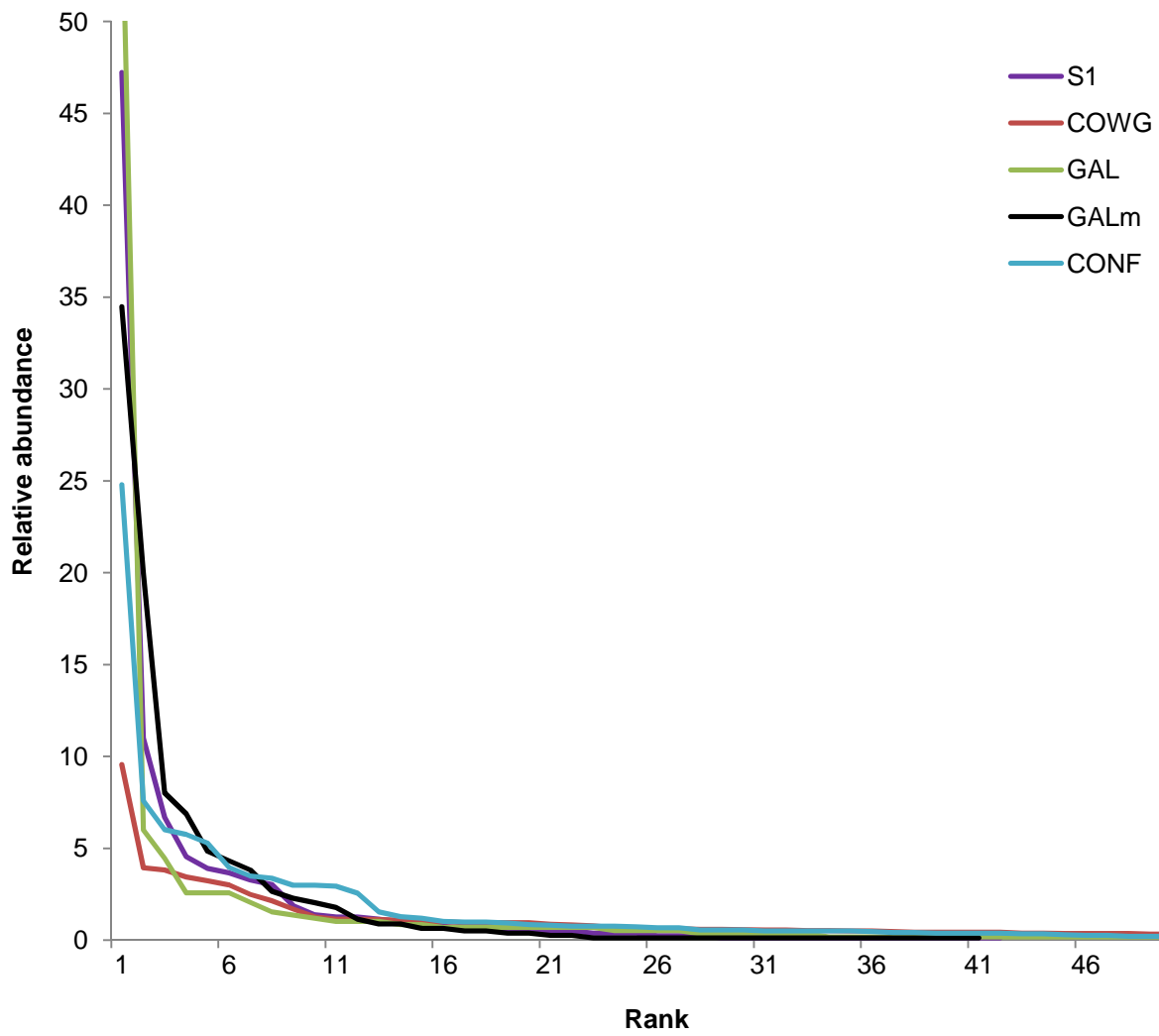


Fig. 2. Rank-abundance curves of the eukaryotic 18S rRNA sequences from Carnoulès mine sediments based on the 50 most abundant OTUs in each sample for the full dataset.

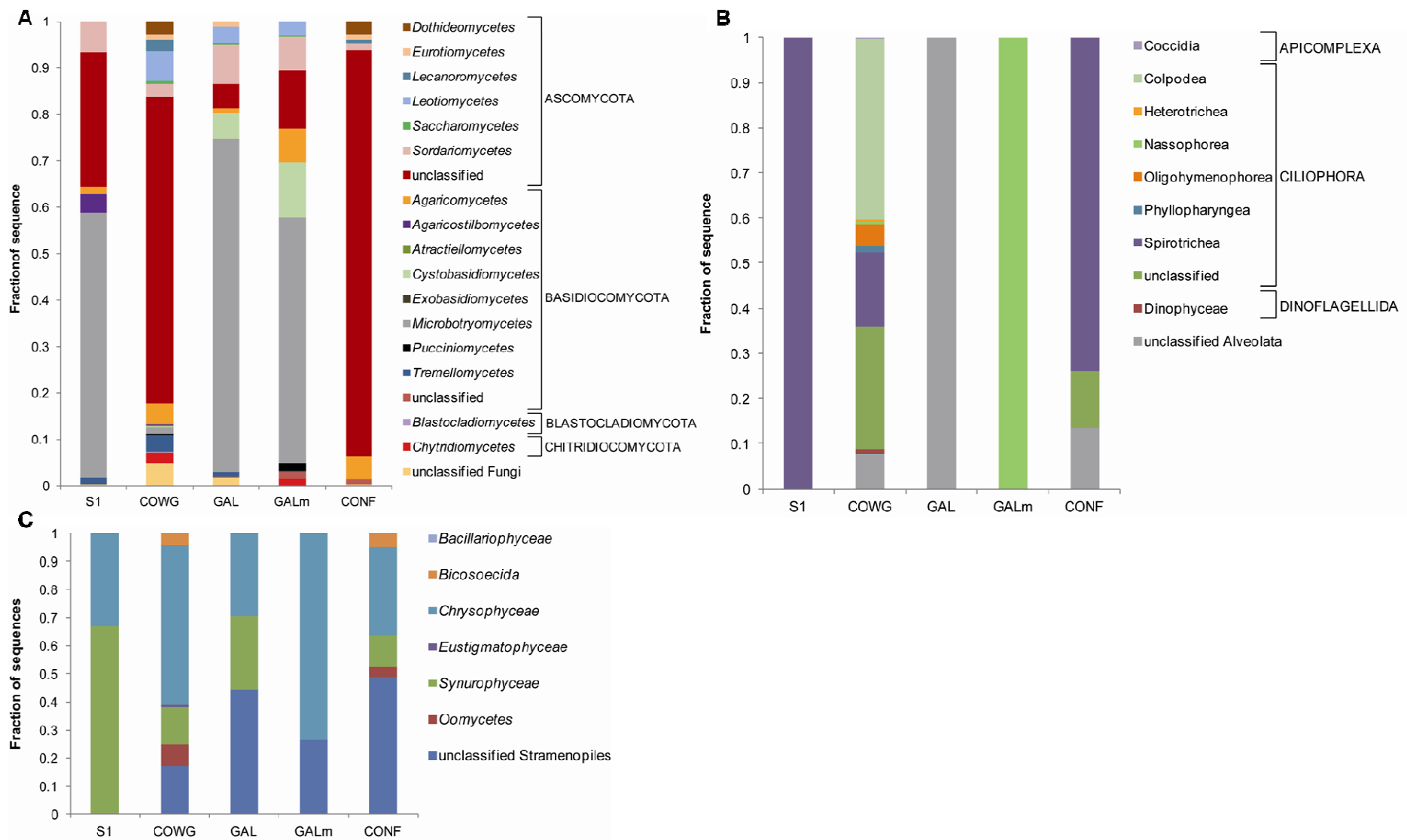
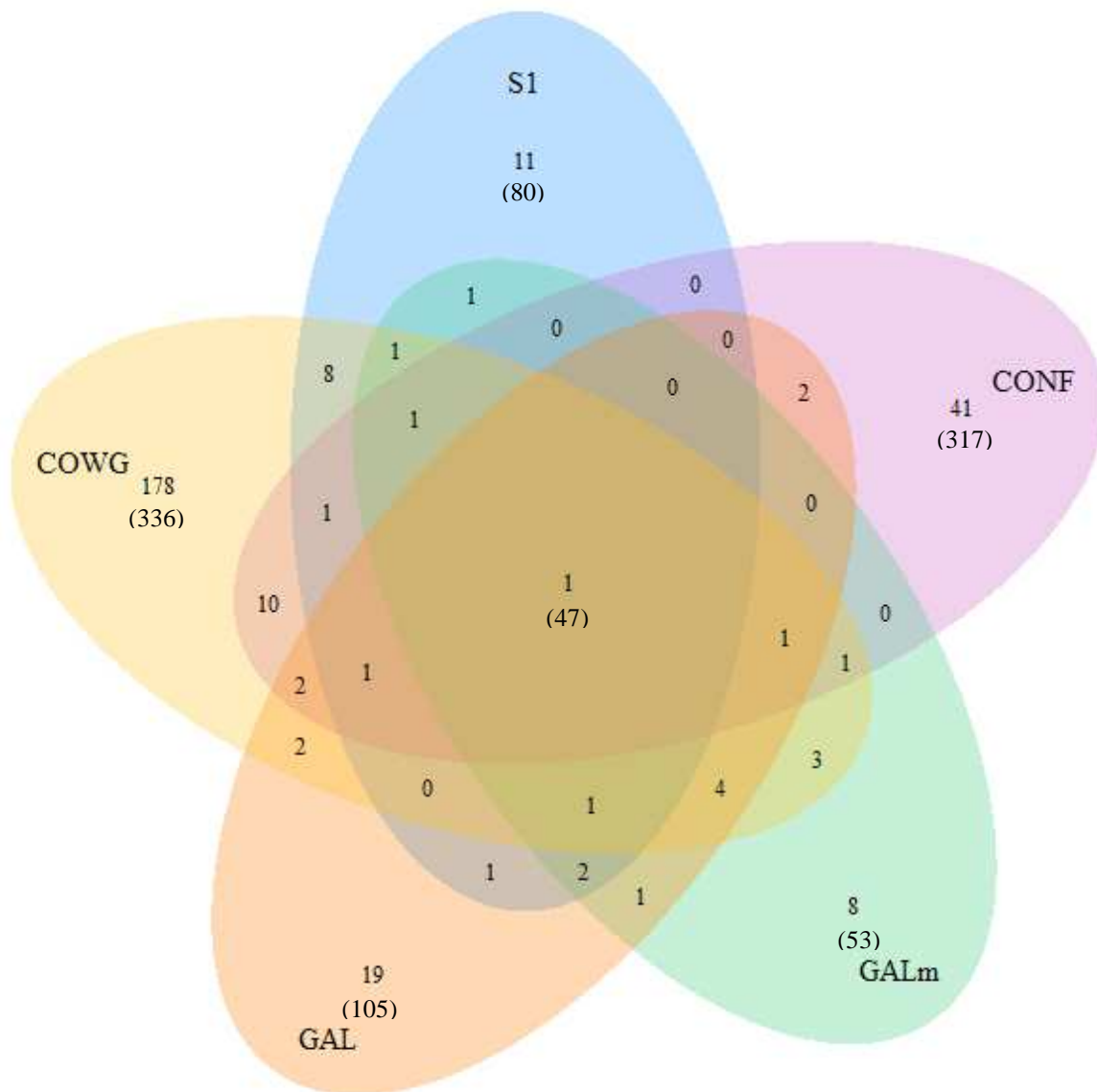


Fig. 3. Taxonomic distribution of the reads retrieved from the five samples and assigned to the Fungi (A), Alveolata (B) and Stramenopiles (C).



Proportion of unique OTUs: 85% (257/301)
 Proportion of unique sequences: 32% (891/2825)

Fig. 4. Venn diagram showing the number of shared and unique OTUs between the five sampling sites for normalized OTUs abundances. The number of sequences associated with OTU is shown in parentheses. OTUs are defined by 97% sequence similarity.

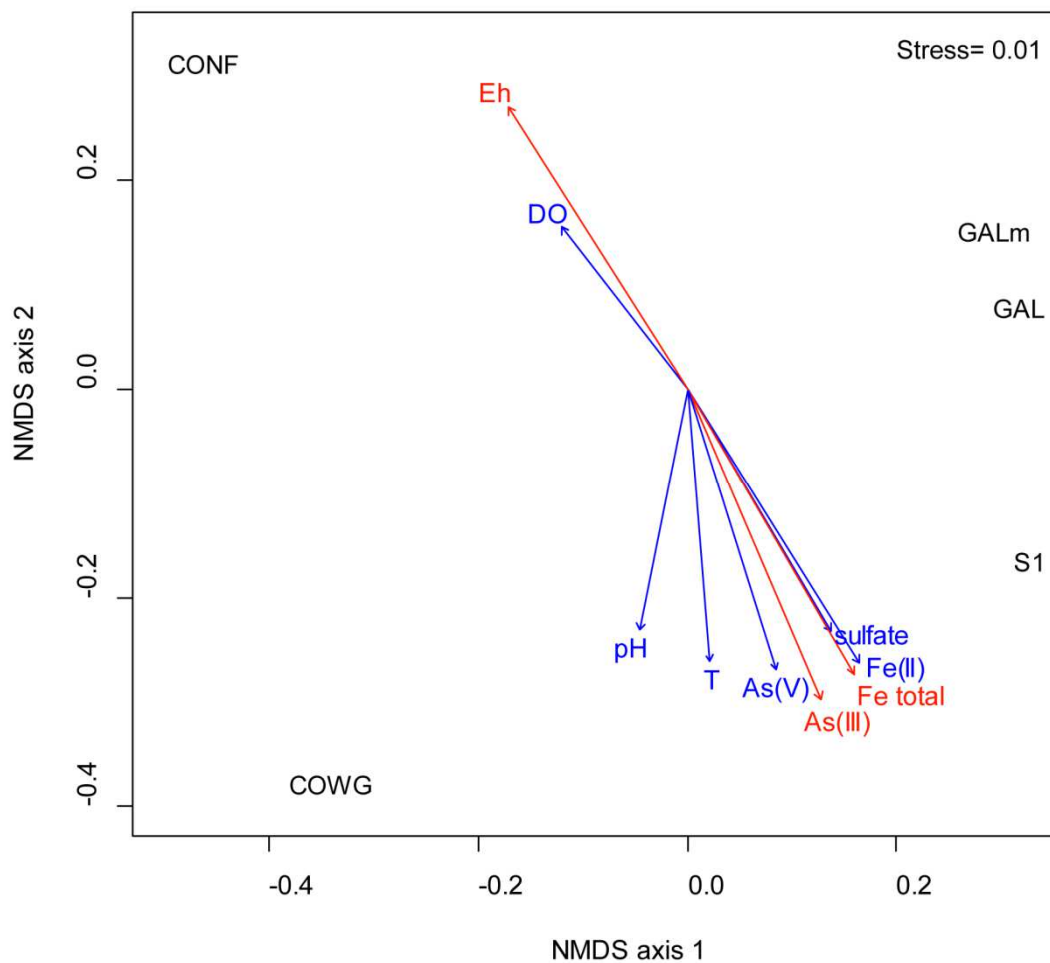


Fig. 5. Nonmetric MultiDimensional Scaling (NMDS) for normalized OTUs abundances at different sampling sites. Arrows are projections of possible explanation variables gained by vector fitting. Significant explanatory variables are shown in red arrows, whereas supplementary variables are shown in blue.

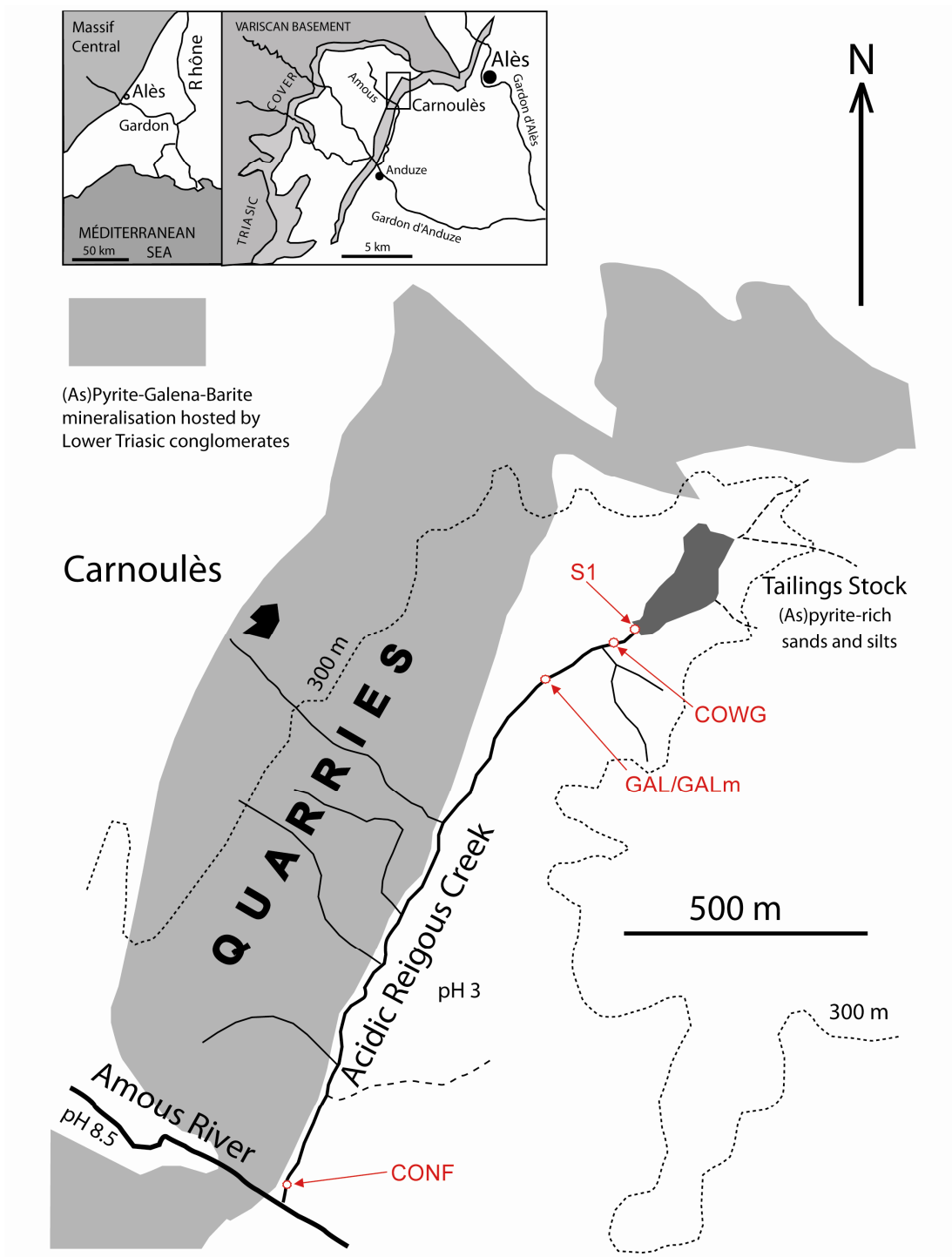


Fig. 6. Map of the Carnoulès mining site and localization of the sampling stations

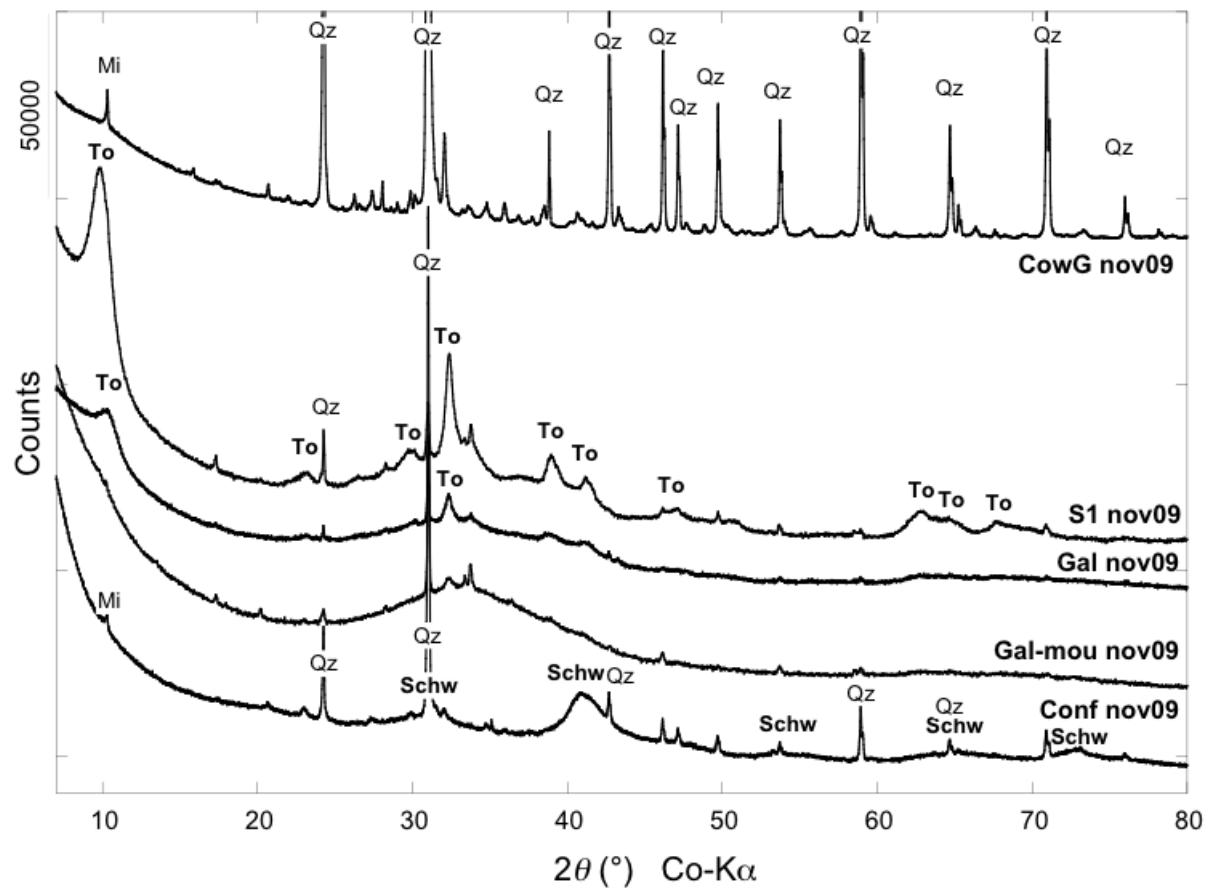


Fig. S1. X-ray Powder Diffraction patterns of selected sediment samples collected along the Reigous River in November 09.

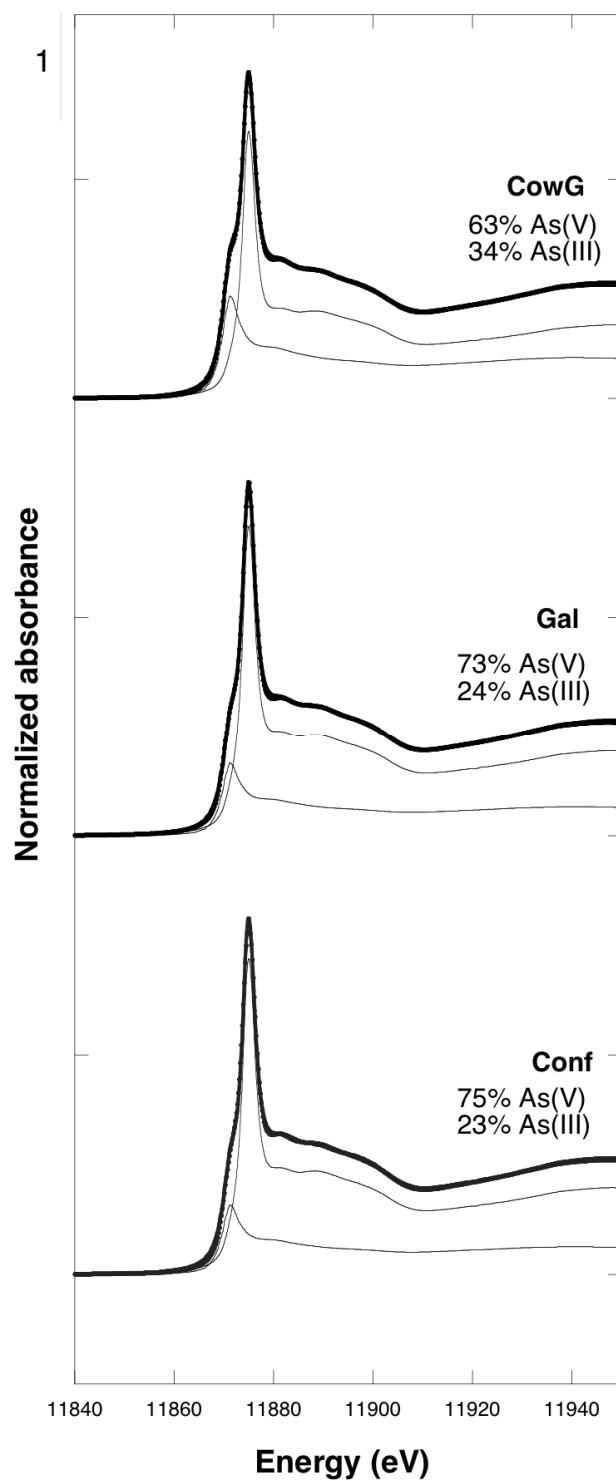


Fig. S2. XANES spectra of selected sediment samples. Arsenic oxidation states are determined using linear combination fit of As(III) and As(V) ferric hydroxides spectra. Precision is $\pm 2\%$.

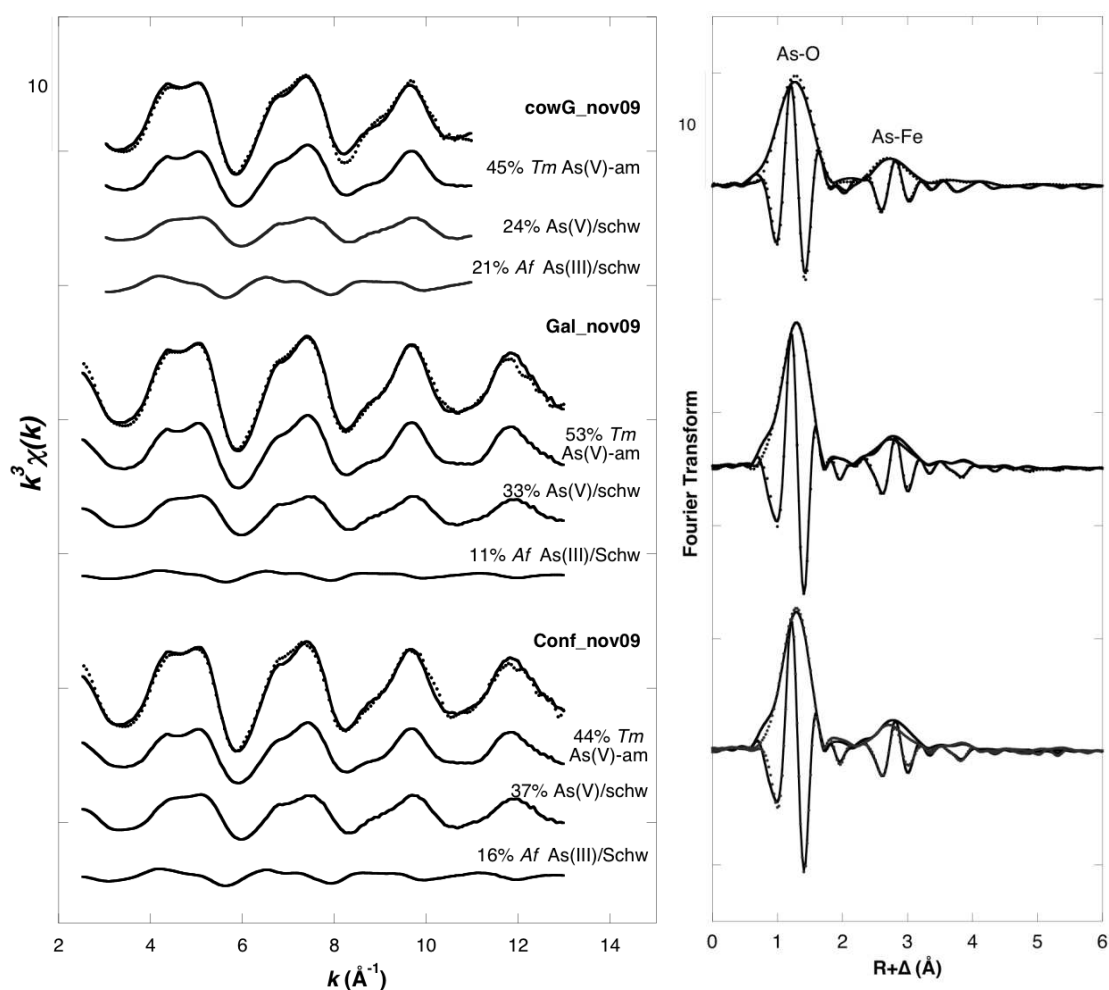


Fig. S3. Extended X-ray Absorption Fine Structure (EXAFS) data at the As K-edge for selected sediment samples. Experimental spectra are interpreted using Linear Composition Fitting (LCF) using 3 model compound spectra: biogenic amorphous ferric arsenate hydroxysulfate from *Thiomonas* sp. strain B2 (Tm As(V)-am); arsenic(V) sorbed onto synthetic schwertmannite (As(V)/schw); arsenic(III) sorbed onto biogenic schwertmannite from *Acidithiobacillus ferrooxidans* strain CC1(Af As(III)/schw). See Morin *et al.* (2003), Egal *et al.* (2009) and Maillot *et al.*, submitted, for details on these biogenic and abiotic mineral model compounds. As(III) proportions are underestimated with respect to XANES fit results because of the lower sensitivity of EXAFS to redox composition.

Chapitre IV – Dynamiques spatiotemporelles de la communauté bactérienne et drivers physicochimiques

A. PROBLÉMATIQUE ET DÉMARCHE SCIENTIFIQUE

1. Contexte général

Après avoir réalisé des études portant principalement sur la description des communautés microbiennes (bactéries, archaea et eucaryotes) au sein du DMA de Carnoulès, nous avons choisi d'aborder dans ce dernier chapitre des questions plus générales d'écologie microbienne abordant notamment la notion de biogéographie. Mon travail de thèse s'est en effet inséré dans le cadre d'un projet FRB, MIGRAMD, déposé en 2009 pour 2 ans par Béatrice Lauga (EEM de Pau) intitulé : « Microbial biogeography of Acid Mine Drainage: a study of genetic diversity and species diversity from an evolutionary perspective ». La biogéographie, c'est-à-dire l'organisation et la répartition des communautés dans l'environnement, constitue un domaine étudié depuis longtemps dans le cas des organismes supérieurs mais beaucoup moins en ce qui concerne les microorganismes (Green *et al.*, 2008). La compréhension des mécanismes régissant l'organisation spatiale de ces communautés microbiennes et leur diversité est donc fondamentale pour l'écologie microbienne au vue de l'importance que joue ces microorganismes dans les écosystèmes, mais est également indispensable et utile pour de nombreux domaines où la protection de l'environnement devient une réelle préoccupation (Haferburg & Kothe, 2007). Les conditions de vie extrêmes qui règnent dans les DMA (*i.e.* pH acides, fortes concentrations en métaux et métalloïdes toxiques et qui diffèrent d'un site à l'autre) en font en effet de bons candidats pour aborder un certain nombre de questions fondamentales de biogéographie et contribuer à documenter et comprendre les profils de diversité microbienne. Les fortes contraintes physicochimiques qui s'exercent sur ces écosystèmes donnent en effet l'opportunité d'étudier l'effet des pressions de sélection sur la diversité microbienne et les communautés simplifiées qui les caractérisent peuvent être comparées plus facilement contrairement aux environnements plus complexes (Baker & Banfield, 2003; Banfield *et al.*, 2005). Les méthodes de séquençages haut-débit sont ici un

outil de choix pour aborder ces problématiques. Ainsi elles ont été utilisées pour étudier de façon approfondie les assemblages microbiens de 12 DMA localisés dans trois régions minières françaises, hispano-portugaises et boliviennes. Parallèlement à cette approche spatiale, la dynamique spatiotemporelle des communautés bactériennes a été étudiée le long du continuum du DMA de Carnoulès en France. C'est sur cette dernière partie que mon travail de thèse s'est essentiellement focalisé.

Le DMA de Carnoulès constitue en effet un site privilégié pour étudier à une échelle locale ces nombreuses questions car il présente un gradient spatial de pollution résultant de processus naturels de remédiation. En effet, près de 95 % de l'arsenic en solution disparaît entre la source du Reigous et sa confluence avec la rivière Amous, 1,5 km en aval. De plus, ce site présente de fortes contraintes physicochimiques (jusqu'à 10 g L⁻¹ d'As dans l'eau interstitielle du stérile minier, concentration parmi les plus importantes au monde) qui donne l'opportunité d'étudier l'effet des pressions de sélection sur la biodiversité microbienne.

2. Objectifs de l'étude

Afin de mieux comprendre les mécanismes d'adaptation et de colonisation des microorganismes dans cet écosystème, un suivi spatiotemporel des communautés bactériennes a été réalisé et corrélé avec les paramètres physicochimiques mesurés *in situ*. Pour cela, deux approches de biologie moléculaire ont été utilisées, une technique d'empreinte moléculaire (*i.e.* la T-RFLP, Terminal-Restriction Fragment Length Polymorphism) permettant l'étude de la structure de ces communautés et une technique de séquençage à haut débit (*i.e.* pyroséquençage 454) permettant une caractérisation plus fine de la diversité bactérienne. Ce niveau de détails n'avait encore jamais été exploré jusqu'à présent pour les bactéries à Carnoulès et avait pour but de mieux comprendre la dynamique spatiotemporelle des communautés bactériennes et d'identifier les paramètres physicochimiques structurant ces communautés.

B. RÉSULTATS

Le pyroséquençage a permis d'obtenir un total de 66016 séquences qui ont conduit à l'identification de 6801 OTUs incluant 4629 singletons. Bien que représentant une forte proportion du nombre total d'OTUs (68 %), ces singletons ne représentaient qu'un faible

pourcentage (7 %) du nombre total de séquences mettant ainsi en évidence la présence de d'un grand nombre de taxons rares. Les analyses taxonomiques ont permis d'identifier un total de 23 phyla bactériens dont les plus abondants sont communément retrouvés dans les DMA (Proteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes, Firmicutes et Nitrospira) Sur l'ensemble des échantillons analysés, les protéobactéries représentaient le phylum majoritaire constituant à lui seul près de 68 % du nombre total de séquences. Dans une moindre mesure, les Actinobacteria, les Acidobacteria et les Firmicutes faisaient également partie des phyla les plus abondants de notre jeu de données avec des abondances relatives d'environ 5 %. L'analyse des classes constituant chacun de ces phyla a révélé une variation en termes d'abondance au sein des échantillons. Une augmentation générale des bêtaprotéobactéries a par exemple été observée d'amont en aval du Reigous. Cette étude a également révélé la présence d'un faible nombre d'OTUs dominant l'ensemble des échantillons. Les trois OTUs majoritairement présents étaient apparentés à des espèces trouvées précédemment dans ce DMA dont *Gallionella ferruginea*, *Acidithiobacillus ferrooxidans* et *Thiobacillus* sp.. *G. ferruginea* constituait la population majoritaire dans presque tous les échantillons et plus particulièrement au niveau de GAL et COWG. Parallèlement, *Acidithiobacillus ferrooxidans* ne représentait qu'une fraction mineure de la communauté bactérienne excepté au niveau de S5 où cette espèce était dominante.

Les deux techniques moléculaires utilisées, la T-RFLP et le pyroséquençage, ont mis en évidence une faible diversité générale caractérisée par un total de 58 TRFs (2 à 20 TRFs par échantillons) pour la T-RFLP et 6801 OTUs représentant 23 phyla pour le pyroséquençage. Ces deux techniques ont par ailleurs révélé des profils de diversité légèrement différents. La T-RFLP a mis en évidence la présence d'une diversité plus importante au niveau de S1 et COWG tandis que les communautés les moins diversifiées se trouvaient au niveau de GAL et S5. Par contre, les données de pyroséquençage ont identifié une plus grande diversité au niveau de S1 et CONF alors que la plus faible diversité se trouvait au niveau de S5.

Alors que les analyses n'ont pas montré de structuration spécifique des communautés basées sur les profils T-RFLP, une séparation spatiale évidente a pu être mise en évidence par les données de pyroséquençage. Une analyse canonique des correspondances a par ailleurs identifié trois paramètres physicochimiques (concentration en arsenic, température et potentiel redox) comme des facteurs potentiellement responsables de cette structuration.

C. DISCUSSION ET CONCLUSION

En accord avec des études précédentes (Bruneel *et al.*, 2006; Bruneel *et al.*, 2011), les deux approches moléculaires utilisées ont mis en évidence une faible diversité bactérienne dans ce milieu extrême. L'analyse des profils de T-RFLP a révélé la présence de 2 à 20 TRFs par stations échantillonnées, représentant un total de 58 TRFs. En parallèle, les données de pyroséquençage ont conduit à l'identification de 6801 OTUs, incluant une majorité de singletons (68 %) constituant la « rare » biosphère de cet écosystème. Ces taxons rares, représenteraient une source potentielle de phylotypes susceptibles de devenir dominants et actifs suite à une modification des conditions environnementales (Galand *et al.*, 2009). Ces microorganismes seraient alors capables de maintenir les fonctions microbiennes des organismes les plus abondants qui auront été affecté par cette perturbation, ce phénomène permettant ainsi de limiter les effets des modifications environnementales (Sogin *et al.*, 2006; Monchy *et al.*, 2011). Malgré leur prévalence, nous n'avons jusqu'à présent que très peu d'informations sur le rôle écologique et fonctionnel de ces organismes, ceci résultant en partie du fait que la majorité des études de diversité réalisées jusqu'ici étaient basées sur la méthode classique de clonage/séquençage ne permettant pas la détection de ces taxons rares (Sogin *et al.*, 2006). Il apparaît donc nécessaire de pousser plus loin les investigations afin de pouvoir déterminer leur rôle potentiel au niveau du DMA de Carnoulès.

D'autre part, l'analyse NMDS des données de pyroséquençage a révélé une structuration spatiale des communautés bactériennes, caractérisée par une composition bactérienne significativement différente entre les cinq stations échantillonnées. En accord avec les données de T-RFLP, des valeurs moyennes de diversité relativement élevées ont été trouvées au niveau de S1 tandis que la plus faible diversité a été rencontrée au niveau des stations GAL et S5. Cette structuration des communautés bactériennes le long du Reigous semble être corrélée avec les concentrations en arsenic, la température et le potentiel redox.

Cette étude a également permis d'atteindre une plus grande profondeur dans l'analyse de la diversité bactérienne présente sur le site de Carnoulès en mettant en évidence un nombre bien plus élevé de taxons que dans les études précédemment réalisées, incluant des taxons non détectés auparavant. Ces nouveaux taxons, ayant pour la plupart une faible abondance au niveau de chaque site (<1 %), incluaient des membres des Bacteroidetes, Chlorobi, Chloroflexi, Elusimicrobia, Chlamydiae, Cyanobacteria, Deinococcus-Thermus, Spirochaetes,

Fibrobacteres, Fusobacteria, Gemmatimonadetes, Plantctomycetes, Synergistetes, Verrumicrobia ainsi que des membres des clades non cultivés OD1-PO11-TM7. La plupart des taxons révélés par cette étude étaient affiliés à des ordres dont les genres sont connus pour être commun au sein des DMA et pour la plupart impliqués dans les cycles du fer, de l'arsenic et du soufre tels que les Gallionellales (Betaproteobacteria), Acidithiobacilliales (Gamaproteobacteria), Desulfobacterales (Deltaproteobacteria), Desulfomonadales (Deltaproteobacteria), Acidimicrobiales (Actinobacteria) ou les Nitrospirales (Nitrospira) (Bruneel *et al.*, 2005; Bruneel *et al.*, 2006; Bertin *et al.*, 2011; Bruneel *et al.*, 2011). Des variations dans l'abondance relative des OTUs les plus abondants apparentés à *Gallionella ferruginea*, *Acidithiobacillus ferrooxidans* et *Thiobacillus* sp., ont été mises en évidence au cours de la période d'échantillonnage au niveau des différentes stations.

En conclusion, les résultats obtenus ont souligné l'intérêt de l'utilisation des nouvelles techniques de séquençage à haut débit permettant de mettre en évidence une diversité majoritairement sous estimée par les techniques classiques de clonage/séquençage Sanger. Cette constatation ouvre la voie à de nombreuses interrogations concernant le rôle écologique des nombreux taxons rares mis en évidence. Il sera donc nécessaire de prendre en compte dans les études futures, cette part importante et jusqu'alors méconnue de la diversité présente dans le DMA de Carnoulès et de tenter d'identifier les fonctions potentiellement réalisées par ces organismes. Cette étude a également mis en évidence une structuration spatiale des communautés bactériennes le long du Reigous en relations avec certains paramètres physicochimiques dont il faudra décrire avec plus de précision l'impact réel sur ces communautés.

D. CONTRIBUTION SCIENTIFIQUE

Ce travail va être valorisé sous la forme d'une publication scientifique qui est en cours de préparation et dont le manuscrit préliminaire est présenté ici (Article D).

Volant, A., Bruneel, O., Desoeuvre, A., Héry, M., Casiot, C., Bru, N., Delpoux, S., Fahy, A., Javerliat, F., Duran, R., Bertin P. N., Elbaz-Poulichet, F. & Lauga, B. Spatiotemporal dynamics of bacterial community composition: physicochemical drivers in an acid mine drainage heavily contaminated by arsenic (Carnoulès, France)
En préparation.

Chapitre IV – Article D

Spatiotemporal dynamic of bacterial community composition: physicochemical drivers in an acid mine drainage heavily contaminated by arsenic (Carnoulès, France)

A. Volant, O. Bruneel, A. Desoeuvre, M. Héry, C. Casiot, N. Bru, S. Delpoux, A. Fahy, F. Javerliat, R. Duran, P. N. Bertin, F. Elbaz-Poulichet, B. Lauga

In preparation

1 **Spatiotemporal dynamics of bacterial community composition:**
2 **physicochemical drivers in acid mine drainage**

3
4
5 A. Volant¹, O. Bruneel¹, A. Desoeuvre¹, M. Héry¹, C. Casiot¹, N. Bru², S. Delpoux¹, A.
6 Fahy³, F. Javerliat³, R. Duran³, P. Bertin⁴, F. Elbaz-Poulichet¹ and B. Lauga³

7
8
9 *In preparation*

10
11 ¹. Laboratoire HydroSciences Montpellier, HSM, UMR 5569 (IRD, CNRS, Universités
12 Montpellier 1 et 2), Université Montpellier 2, Place E. Bataillon, CC MSE, 34095
13 Montpellier, France

14
15 ². Laboratoire de Mathématiques et de leurs Applications, UMR 5142 (CNRS), Université de
16 Pau et des Pays de l'Adour, Bât IPRA, BP 1155, 64013 Pau, France

17
18 ³. Équipe Environnement et Microbiologie, EEM, UMR 5254 (IPREM, CNRS), Université de
19 Pau et des Pays de l'Adour, BP 1155, 64013 Pau, France

20
21 ⁴. Laboratoire de Génétique Moléculaire, Génomique, Microbiologie, GMGM, UMR 7156
22 (Université de Strasbourg, CNRS), Département Microorganismes, Génomes,
23 Environnement, 28 Rue Goethe, 67083 Strasbourg, France

24
25
26
27 *Corresponding author. Aurélie VOLANT. Laboratoire HydroSciences Montpellier,
28 UMR5569, Université Montpellier 2, Place E. Bataillon, CC MSE, 34095 Montpellier,
29 France. Phone: (+33)4 67 14 90 19; Fax: (+33)4 67 14 47 74. Email:
30 aurelie.volant@msem.univ-montp2.fr

31
32 **Key words:** Spatiotemporal dynamic, Bacterial diversity, Acid mine drainage, Arsenic..

33 Abstract

34 The understanding of biotic and abiotic forces that shape microbial community structure is an
35 important question in microbial ecology. The Carnoulès mine (France), characterized by acid
36 waters and very high concentrations of arsenic, iron, and sulfate provides a good opportunity
37 to address the factors that structure communities inside the tailing stock and along the
38 pollution gradient in Reigous Creek, issuing from the mine tailings. Indeed, As concentration
39 subsequently decreases by around 95% between the Reigous spring and its confluence with
40 the Amous River, 1.5 km downstream. Biodiversity and distribution patterns of microbial
41 populations were identified by two molecular approaches, T-RFLP and 454-pyrosequencing
42 of 16S rRNA genes to try to decipher the factors that shape communities structure in the
43 Carnoulès mine. Five stations located along the AMD were sampled at six sampling
44 campaigns over approximately two years. Both methods have highlighted a low diversity,
45 with a total of 58 T-RFs for TRFLP analysis, and 6801 OTUs of which 68% were singletons
46 for the pyrosequencing approach. NMDS indicated that the bacterial community composition
47 significantly differed spatially between the five stations with a high values of diversity in S1
48 and lowest values for GAL and S5. CCA seemed to indicate that the environmental variables
49 significantly correlated with the variation of the bacterial communities were arsenic,
50 temperature and redox potential. The pyrosequencing analysis enable us to capture a broader
51 phylogenetic breadth of taxa than in any previous clone library-based diversity surveys of the
52 Carnoulès AMD, including sequences of several taxa never detected before like members of
53 Bacteroidetes, Chlorobi, Chloroflexi, Elusimicrobia, Chlamydiae, Cyanobacteria,
54 Deinococcus-Thermus, Spirochaetes, Fibrobacteres, Fusobacteria, Gemmatimonadetes,
55 Plantctomycetes, Synergistetes, Verrumicrobia and members of the uncultured OD1-PO11-
56 TM7 clade. However, most of the taxa detected in this study were related to orders whose
57 genera are known to be common in AMD and for a large part to be involved in the Fe, As,
58 and S cycles, namely Gallionellales (Betaproteobacteria), Burkholderiales
59 (Betaproteobacteria), Acidithiobacilliales (Gamaproteobacteria), Desulfobacterales
60 (Deltaproteobacteria), Desulfomonadales (Deltaproteobacteria), Acidimicrobiales
61 (Actinobacteria) or Nitrospirales (Nitrospira).

62 Introduction

63 Acid Mine Drainages (AMD) are one of the most pernicious forms of pollution in the
64 world, and are widely accepted as responsible for costly environmental and socio-economical
65 impacts (Hallberg, 2010). They are generated when the wastes from the mining and processing
66 of sulfide ores come into contact with oxygenated water. Such waters are often characterized
67 by acidic pH and generally contain high concentrations of sulfates, metals and metalloids
68 including arsenic. Although very hostile to life, some microorganisms are able to thrive in this
69 extreme environment. Their roles in the oxidation reactions of sulfide minerals leading to the
70 bioleaching are well known, as well as their roles in the bioremediation of these polluted
71 waters (Edwards *et al.*, 2000; Hallberg, 2010; Johnson, 2012). However, despite
72 microorganisms make critical contributions to AMD functioning through their participation in
73 biogeochemical cycles, we still have only a limited understanding of the factors that control
74 the spatiotemporal structure and diversity of microbial communities. For example, although it
75 is clear that microbial community composition is influenced by environmental variations
76 (Aguilera *et al.*, 2006; Amaral-Zettler *et al.*, 2011), the question of how diversity changes
77 along environmental gradients remains generally unresolved. Elucidating how biodiversity is
78 distributed and the mechanisms underlying those patterns is a central goal of ecology. The
79 AMD of Carnoulès mine in southern France provides an excellent opportunity to investigate
80 these fundamental questions regarding microbial ecology since a spatial gradient of pollution
81 was highlight at this site due to efficient natural remediation processes. Indeed, nearly 95% of
82 the arsenic in solution disappears between the source of the Reigous Creek, which collects the
83 acid mine drainage, and its confluence with the river Amous, 1.5 km downstream.
84 Furthermore, to our knowledge, the Carnoulès AMD is one of the most As-rich AMD reported
85 to date (Morin and Calas, 2006). This site thus provides an outstanding example of adaptation
86 of life to an extreme environment characterized by acidic pH (2-4) and very high level of
87 metallic elements, especially As (up to 10 g L⁻¹ in the tailing stock pore water, and 100-350 mg
88 L⁻¹ in Reigous Creek, which collects the acid mine drainage). Moreover, the strong
89 physicochemical constraints observed also provide the opportunity to investigate the impact of
90 selection pressures on biodiversity.

91 In this study, we used a combination of molecular approaches to investigate the
92 spatiotemporal dynamics of bacterial community in relation to the physicochemical
93 parameters in the acid mine drainage of Carnoulès. By using 16S rRNA gene pyrosequencing
94 and Terminal Restriction Fragment Length polymorphism (T-RFLP) we aimed to (i) compare

95 the structure and the composition of the bacterial communities with contrasting environmental
96 conditions, (ii) assess the spatial and temporal variations in the composition of these
97 communities, and (iii) find out whether the dynamics of these communities are linked to
98 changes in environmental conditions.

99

100 **Materials and methods**

101 *Description of the study site*

102 The Pb-Zn Carnoulès mine, located in southern France, has produced 1.2 Mt of solid
103 wastes containing 0.7% Pb, 10% Fe and 0.2% As, which are stored behind a dam. The aquifer
104 is not fed by the vertical percolation of rainwater through the tailings, but rather originates
105 from former natural springs that were buried under the tailings (Koffi *et al.*, 2003). The water
106 table occurs at 1 to 10 m below the surface of the tailings stock according to season and
107 location. With the exception of the temperature, almost constant with average values around
108 15°C, the other physicochemical parameters of ground water change in relation to the
109 hydrological conditions (Casiot *et al.*, 2003a). The groundwater below the tailings contains
110 extremely high levels of arsenic: up to 10000 mg L⁻¹ found in 2001 (Casiot *et al.*, 2003a). The
111 water emerges at the bottom of the dam forming the source of the Reigous Creek.

112 This AMD is acid (pH≤3) strongly enriched in sulfate (2000-7700 mg L⁻¹), iron (500-
113 1000 mg L⁻¹) and arsenic (50-350 mg L⁻¹), the later being predominantly in reduced forms
114 Fe(II) and As(III) (Casiot *et al.*, 2003b). Natural attenuation of As concentrations occurs as a
115 result of microbiologically mediated As-Fe coprecipitation (Morin *et al.*, 2003, Bruneel *et al.*,
116 2006). From 10 to 47% of the Fe and 20 to 60% of the As are removed from the aqueous
117 phase within the first 30 m of the stream. After the station COWG, Reigous Creek receives
118 some inputs from quarries and mine galleries, especially after rainy events; these strongly
119 influence acidity and metallic content (Egal *et al.*, 2010).

120

121 *Sampling procedure and measurement of physicochemical properties*

122 Six sampling campaigns were carried out in November 2007, February 2008, October
123 2008, March 2009, November 2009 and January 2010 at five sampling stations giving a
124 dataset of 30 samples. Groundwaters were collected in a borehole (S5, between 10 and 12 m
125 deep) located in the center of the tailings. Water samples were also taken at four locations
126 along the Reigous Creek (collecting downstream seepage waters from the surroundings) at the
127 spring (S1), 30 m downstream from the spring (COWG), 150 m downstream (GAL) and 1500

128 m downstream (CONF) just before the confluence between the Reigous and the Amous River.
129 Figure 1 shows the location of sampling stations. Water samples (300 ml) were immediately
130 filtered through sterile 0.22 µm Nucleopore filters that were then transferred to a collection
131 tube (Nunc), frozen in liquid nitrogen, and stored at -80 °C until DNA extraction. This
132 sampling was done in three replicates. Measurements of water conductivity, temperature,
133 redox potential, pH, and dissolved oxygen concentration were carried out as described
134 previously (Bruneel *et al.*, 2011). For chemical analysis, water samples (500 ml) were
135 immediately filtered through 0.22 µm Millipore membranes fitted on Sartorius polycarbonate
136 filter holders. For total Fe and As determination, the filtered water was acidified to pH 1 with
137 HNO₃ (14.5 M) and stored at 4 °C in polyethylene bottles until analysis. A 20 µl aliquot of
138 filtered water sample was added to either a mixture of acetic acid and EDTA (Samanta and
139 Clifford 2005) for As speciation or to a mixture of 0.5 ml acetate buffer (pH 4.5) and 1 ml of
140 1,10-phenanthroline chloride solution (Rodier *et al.*, 1996) for Fe(II) determination. The
141 vials were completed to 10 ml with deionized water. The samples for Fe and As speciation
142 and sulfate determination were stored in the dark and analyzed within 24 h. Chemical analysis
143 were carried out as previously described (Bruneel *et al.*, 2011).

144

145 ***DNA isolation***

146 Triplicate genomic DNA was extracted from filtered water using the UltraClean Soil
147 DNA Isolation Kit according to the manufacturer's recommendations (MoBio Laboratories
148 Inc., Carlsbad, CA, USA). These triplicates were pooled before PCR amplification. All
149 extracted genomic DNA samples were stored at -20 °C until further analysis.

150

151 ***Terminal-restriction fragment length polymorphism***

152 The 16S rRNA genes were amplified by PCR and the bacterial community structure
153 was studied by T-RFLP. The fluorescent labeled primers HEX 357F (5'-hexa-chloro-
154 fluorescein-phosphoramidite-CCT ACG GGA GGC AGC AG-3') (Lane, 1991) and 926R (5'-
155 CCG TCA ATT CMT TTR AGT-3') (Muyzer and Ramsing, 1995), described as universal
156 within the bacterial domain, were used. Triplicate PCR amplifications were performed for
157 each sample. The reaction mixture contained 1 µl of DNA template, 1µl of both primers (10
158 µM) and 12.5 µl of PCR Master Mix Ampli Taq Gold 360 (Applied Biosystems, Foster City,
159 CA, USA). Sterile distilled water was added to reach a final volume of 25 µl. The PCR
160 conditions used were 95 °C for 10 min, 35 cycles at 95 °C for 45 s (denaturation), 55 °C for

161 45 s (annealing) and 72 °C for 45 s (extension), followed by 10 min at 72 °C. The 90 PCR
162 products were purified with the illustra GFX™ PCR DNA and Gel Band Purification Kit (GE
163 Healthcare, Munich, Germany). The amount of PCR product was determined by comparison
164 to known concentrations by the “dots method” (Smartlader, Eurogentec) after migration on
165 agarose gel. Approximately 100 ng of purified amplicons was digested in 10 µl reaction with
166 0.3 U of enzyme HpaII or AluI (New England Biolabs Inc., Ipswich, MA, USA) for 3 h at
167 37°C. Terminal Restriction Fragments (T-RFs) profiles were obtained from the digested
168 amplicons by suspending 1 µl aliquots in 8.75 µl formamide with 0.25 µl of GeneScan ROX
169 500 size standard (Applied Biosystems Foster City, CA, USA). Separation of T-RFs was
170 carried out on an ABI PRISM 3130 xl Genetic Analyser (Applied Biosystems). Data were
171 collected and analyzed by comparison with internal standard by using the GeneMapper
172 software (version 1.4, Applied Biosystem). To enhance stringency for TRF profiles of 16S
173 rRNA genes, TRFs not comprised in the size range of the standard (35-500 bp) and with a
174 height lower than 30 fluorescence units were considered background noise and were manually
175 removed from the TRF profiles in each sample. The T-align software
176 (<http://inismor.ucd.ie/~talign/index.html>) was used to compare replicate profiles and generate
177 consensus profiles containing only TRFs that occurred in replicate reactions. Consensus
178 profiles were then aligned on the basis of TRFs lengths and individual peak areas as
179 previously described by Smith *et al.* (2005) with the confidence interval set at 0.5, resulting in
180 the generation of datasets of aligned TRFs that gave individual relative peak areas as a
181 proportion (%) of the overall profile. TRFs were included in the subsequent analysis if they
182 represented > 1 % of the cumulative peak height for the sample.

183

184 ***Libraries construction, 454-pyrosequencing and sequence quality control***

185 The 16S rRNA genes were also amplified by PCR for multiplexed pyrosequencing
186 using barcoded primers. The primers pairs used, targeting the V3 to V5 variable regions of the
187 16 rRNA gene, were 357F (5'-AxxxCCTACGGGAGGCAGCAG-3') and 926R (5'-
188 BxxxCCGTCAATTCMTTTRAGT-3') where A and B represent the two FLX Titanium
189 adapters (A adapter sequence: 5'-CGTATCGCCTCCCTC GCGCCATCAG-3'; B adapter
190 sequence: 5'-CTATGCGCCTTGCCAGCCCGCTCAG-3') and xxx represent the sample-
191 specific barcode sequence. PCR was performed using 30 to 35 cycles under conditions
192 identical to those described above for T-RFLP. Cycle number was varied for samples to
193 obtain a strong band without nonspecific amplification. The 90 PCR products with a proximal

194 length of 600 bp were excised from a 1 % agarose gel and were purified with QIAquick gel
195 extraction kit (QIAGEN Inc., Valencia, CA, USA). Triplicates were pooled prior
196 pyrosequencing to minimize random PCR bias. Pyrosequencing of the 30 amplicon libraries
197 was performed by a Roche 454 Life Sciences Genome Sequencer FLX Titanium sequencer at
198 the Genotoul genomic platform in Toulouse (France) using four separate 1/8 region of a
199 plate.

200

201 ***Processing pyrosequencing data and taxonomic classification***

202 Bioinformaticians at the NG6 platform performed preliminary quality checks, sorting, and
203 trimming of the 454-reads using the NG6 pipeline (<http://vm-bioinfo.toulouse.inra.fr/ng6/>).
204 Tags were extracted from the 454-reads by the Roche's software Sff file and three kinds of
205 analysis were performed as previously described (Ueno *et al.*, 2010): i) BLAST search for *E.*
206 *coli*, phage and yeast contaminants, ii) read quality analysis and iii) removal of sequences that
207 were too long or too short (sequences with more than or less than two standard deviations
208 from the mean), sequences with an excess of errors (more than 4% of N), low-complexity
209 sequences and duplicated reads, using Pyrocleaner program. The sequences were then
210 analyzed with the software program Mothur version 1.26 (<http://www.mothur.org>) (Schloss *et*
211 *al.*, 2009). Preprocessing of unaligned sequences included removing short sequences < 570
212 pb, all sequences containing ambiguous characters, and sequences with more than eight
213 homopolymers. We further removed sequences that did not align over the same span of
214 nucleotide positions. Identical sequences were grouped and representative sequences were
215 aligned against the SILVA bacterial and archaeal reference database using the Needleman–
216 Wunsch algorithm (Needleman and Wunsch, 1970). Chimeric sequences were detected and
217 removed using the implementation of Chimera Uchime. A further screening step (pre-cluster)
218 was applied to reduce sequencing noise by clustering reads differing by only 1 base every 100
219 bases (Huse *et al.*, 2010). The remaining high-quality reads were used to generate a distance
220 matrix and clustering into operational taxonomic units (OTUs) defined at 97% cutoff using
221 the average neighbor algorithm. Next, the OTUs were phylogenetically classified to the genus
222 level using the naïve Bayesian classifier (80% confidence threshold) trained on the RDP's
223 taxonomic outline and implemented in Mothur.

224 ***Diversity estimation and statistical analysis***

225 *Diversity indices.* Nonparametric Chao1 and Shannon alpha diversity estimates as well as the
226 coverage and rarefaction curves for each sample were calculated with Mothur v.1.26.

227 *Cluster analysis.* To compare community composition based on T-RFLP and 454
228 pyrosequencing data, normalized OTUs abundances were square root transformed and
229 pairwise dissimilarities among samples were calculated using the relative abundance-based
230 Bray-Curtis index (BC). Nonmetric multidimensional scaling (nMDS) analysis was
231 performed on the dissimilarity matrices to visualize patterns of community composition.
232 Using the 454 pyrosequencing data, we performed a random sampling procedure to made
233 equal the number of sequences per sample (585 sequences) and we removed OTUs occurring
234 as singletons (unique sequences only occurring once) to down-weight the influence of rare
235 OTUs as suggested by Behnke *et al.* (2011). One-way analysis of similarity (ANOSIM) and
236 multiple pairwise comparisons were used to test whether significant differences in community
237 composition occurred in space. We interpreted R-values > 0.75 as well-separated bacterial
238 compositions; $R > 0.5$ as overlapping, but clearly different; and $R < 0.25$ as practically not
239 separable.

240
241 *CCA analysis.* A canonical correspondence analysis (CCA) was used to investigate the
242 variations in bacterial community under the constraint of our set of environmental variables.
243 Explanatory variables were $\log(x + 1)$ -transformed where necessary to approximate normal
244 distribution. This model was tested with Monte Carlo permutation tests (499 randomized
245 runs) to determine significance, and each environmental parameter was tested by stepwise
246 analysis to detect the significant predictors. All the statistical analyses were performed with R
247 2.14.2 (R Development Core Team, 2011) including the *vegan* package.

248

249 **Results**

250 ***Spatial and temporal analyses of the environmental data set***

251 The physicochemical characteristics of water samples collected from different stations
252 and sampling dates are summarize in Table 1. The temperature of the water in the piezometer
253 (S5) and at the source of the Reigous (S1) varied slightly (13.3 ± 1.4 and 13.6 ± 1 , respectively)
254 and did not differed significantly. A significant decreased of the temperature was observed in
255 the Reigous Creek which is affected by the air temperature resulting in a larger range of
256 values (Fig. 2A). All samples were characterized by low pH values (≤ 3.7), which displayed an

257 increasing trend from S5 to S1 with average values ranging from 2.5 ± 0.8 to 3.0 ± 0.3 . No
258 significant variations in pH were observed between the downstream stations (Fig. 2B).
259 Dissolved oxygen concentrations in the water at the upstream stations S5 and S1 ranged from
260 0.17 mg L^{-1} to 2.79 mg L^{-1} , with a mean of 0.95 ± 0.9 for S5 and 0.84 ± 0.6 for S1, which
261 denoted generally suboxic conditions at these stations. Dissolved oxygen increased sharply
262 between stations S1 and COWG (mean of 6.12 ± 1.8 with values as high as 7.96 mg L^{-1}) (Table
263 1, Fig. 2C). This rapid oxygenation was continuous all along the creek segment reaching
264 means of 9.23 ± 1.3 for GAL and 10.48 ± 1.3 for CONF. The redox potential (Eh) showed
265 average values of 558 ± 85 at station S5. In the Reigous, it increased from 512 ± 39 to 635 ± 99
266 mV across the study site (Fig. 2D). In contrast, average conductivity values decreased from
267 7588 ± 5203 at S5 to 5121 ± 631 at S1 to reach progressively a minimum ($1612\pm 97\text{ }\mu\text{S cm}^{-1}$) at
268 the station CONF (Fig. 2E). Sulfate concentrations ranged from 2504 to 31339 mg L^{-1} in the
269 groundwater at S5, with a mean of $14080\pm 10630\text{ mg L}^{-1}$, exhibiting the highest values. After
270 a sharp diminution at S1 (average values of $2682\pm 1180\text{ mg L}^{-1}$), concentrations decreased
271 gradually along the Reigous Creek (Fig. 2F). At the station CONF, they ranged between 274
272 and 1013 mg L^{-1} (average values of $753\pm 268\text{ mg L}^{-1}$). Dissolved Fe concentrations ranged
273 from 2154 to 9448 mg L^{-1} in water at S5 with average values of $4474\pm 2855\text{ mg L}^{-1}$.
274 Concentrations decreased from the source (S1, average values of $1317\pm 383\text{ mg L}^{-1}$) to the
275 station CONF where Fe concentrations were always below 82 mg L^{-1} (average values of
276 $47\pm 30\text{ mg L}^{-1}$, Fig. 2G). Contrary to the S5 station, the proportion of Fe(III) was generally
277 negligible in Reigous Creek except at the station CONF (Table 1). In the water at S5, the
278 concentrations of dissolved As ranged from 287 to 764 mg L^{-1} with an average value of
279 $440\pm 184\text{ mg L}^{-1}$ (Fig. 2H). Arsenic concentrations decreased with distance from the source to
280 values below 6.1 mg L^{-1} at the station CONF (average value of $3\pm 2.6\text{ mg L}^{-1}$, Fig. 2H), with
281 predominance of the reduced form As(III) (Table 1). Around 72% of sulfate, 96% of iron and
282 99% of arsenic were removed between the station S1 and CONF. All the environmental
283 variables measured showed significant differences between sites (ANOVA, $p<0.05$), except
284 for pH.

285

286 *Spatial variations in bacterial community structure*

287 Distribution patterns of microbial populations were identified by TRFLP analysis and
288 454-pyrosequencing of 16S rRNA genes. The PCR products labeled with Hex were digested
289 separately with two restriction enzymes. As HpaII produced the largest numbers of TRFs

290 (data not shown), the community profiles generated with this enzyme were used to evaluate
291 the differences in the microbial populations across the AMD. On the NMDS plots of the T-
292 RFLP profiles, the samples did not showed clear distinct clusters based on the sampling site
293 as the distribution clearly overlapped (Fig. 3A). Nevertheless, the ANOSIM showed
294 significantly different microbial communities for the site S5 compared with all the others
295 sampling stations, with some communities overlap (R values ranged from 0.58-0.77;
296 $p < 0.005$). The high dissimilarity observed within sites highlighted the temporal variation of
297 the communities structure, especially for S5 and CONF (Fig. 3A). These variations may have
298 masked the spatial pattern.

299 NMDS analyses on 454 pyrosequencing data indicated that the bacterial community
300 composition significantly differed spatially between the five stations (Fig. 3B). Although
301 considerable variations existed among the bacterial communities from each station at different
302 sampling times, the communities from each station did not overlap. Furthermore, the
303 ANOSIM test corroborated the NMDS plot data, showing significantly different ($R = 0.6009$;
304 $P < 0.0001$) bacterial composition in water according to the spatial location, except for GAL
305 and COWG which did not differed significantly ($R = 0.159$; $P = 0.068$). Moreover, a higher
306 temporal variation was highlighted within CONF as shown by the NMDS plot (Fig. 3B).

307

308 *Diversity and species richness estimators of bacterial communities*

309 T-RFLP profiles generated for the five sites showed a total of 58 T-RFs of different
310 sizes and the number of TRFs detected from each sample ranged from 2 to 20, indicating a
311 low biodiversity in this extreme habitat (Table 2). Average OTU richness assed by
312 determining the number of TRFs detected, and average Shannon diversity indices calculated
313 from relative band intensity data, were higher in S1 and COWG, whereas the lowest value
314 was observed for GAL. CONF exhibited the third highest values (Fig. 4). Although the
315 bacterial community diversity showed variable patterns between stations, they were not
316 significantly different (ANOVA, $p = 0.069$). Several detected TRFs were station specific, all
317 stations had 1 to 3 unique TRFs as indicated by red color in Table 2, whereas a few others
318 were found in most of the samples analyzed. In particular, TRF 99 could be detected in almost
319 all samples and in most cases accounted for a high proportion of the total TRFs (Table 2).
320 These results indicated that these microbial communities seemed to be related.

321 A total of 99441 sequences reads were generated in a single run of 454
322 pyrosequencing from the 30 independent 16S rRNA libraries. Note that two samples taken in

323 February 2008 (S5 and S1) failed at pyrosequencing and were excluded from the analysis.
324 After trimming and processing with Mothur, 66016 reads remained with an average length of
325 540 bp. Clustering of the remaining sequences led to the identification of 6801 OTUs
326 (including 4629 singletons) defined at 97% of identity. Although singleton OTUs represented
327 more than half (68%) of the total number of OTUs, they accounted for a small percentage
328 (7%) of the total DNA sequence. The result of rarefaction analysis as well as Chao1, Shannon
329 and coverage values are provided in Table 3. Good's coverage ranged from 77 to 98% with an
330 average value of 88% indicating that about twelve additional phylotypes would be expected
331 for every 100 additional sequenced reads. However, the values were very different according
332 to the sites with a quite good coverage (>90%) for S5 and Gal and a relatively low coverage
333 (<90%) for S1 and CONF. This level of coverage suggested that additional sequencing effort
334 will be required to avoid underestimation of bacterial diversity in the samples. Species
335 richness (Chao1 index) and diversity (Shannon index) of bacterial communities varied across
336 the Reigous Creek (Fig. 5). The non-parametric estimators Chao1 ranged between 230 and
337 3829 estimated OTUs for all the sites considered (Table 3). The highest average OTU
338 richness were found in CONF and S1 (Chao=1687±1183 and 1664±721, respectively),
339 indicating the presence of many more endemic taxa within these two environments. Between
340 these stations, COWG and GAL exhibited intermediates values with no significant difference
341 in their richness (Chao=889±493 and 778±207, respectively). The lowest values was
342 encountered in the northernmost station S5 (Chao=461±152). Moreover, the microbial
343 diversity, as estimated by the Shannon index, also presented significant variations across the
344 study site with values ranging from 1.50 to 5.93 (Table 3). In accordance with the T-RFLP
345 data, high average values of diversity were found in S1 (H=4.13±0.5) while the lowest value
346 was encountered in GAL (H=2.29±0.5) (Fig.5). These patterns of richness and diversity
347 occurred for both full and resampled data sets (data not shown), although the richness
348 estimator and Shannon index became lower compared with full dataset with large number of
349 sequences.

350

351 *Taxonomic assignment of bacterial pyrosequencing*

352 At a confidence threshold of 80%, 58325 out of 66016 qualified reads (that is, 88.3%)
353 could be assigned to a known phylum (Table 4). Most of the unclassified reads were
354 associated with S1 samples (ranging from 9.1 to 45.4% of the qualified reads). Altogether, 23
355 bacterial phyla were recovered from our samples, with 4-8 different phyla found in S5

356 samples, 7-12 in S1, 8-10 in COWG, 9-11 in GAL and 9-17 in CONF (Fig. 6; Table 4). Most
357 of the bacterial sequences (84.4%) belonged to phyla that are most often encountered in acid
358 mines drainages over the world (Proteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes,
359 Firmicutes and Nitrospira). Proteobacteria was the most abundant phylum in all the analyzed
360 samples accounting for 67.5% of the total sequences retrieved. The relative abundance of
361 Actinobacteria, Acidobacteria and Firmicutes varied among samples, and together with
362 Proteobacteria, these phyla were generally the four most abundant phyla in the samples
363 analyzed. To gain more insight into the dynamics of bacterial community, we also used
364 phylum specific sequence analysis to compare different sample stations about the presence of
365 these four phyla (Fig. 7). The Proteobacteria phylum was represented by bacteria belonging to
366 the classes Betaproteobacteria, Gammaproteobacteria, Alphaproteobacteria,
367 Deltaproteobacteria and Epsilonproteobacteria. The most abundant class in all samples was
368 Betaproteobacteria and Gammaproteobacteria, followed by Alphaproteobacteria (Fig. 7A).
369 Deltaproteobacteria and Epsilonproteobacteria represented minor group not always present in
370 the samples. A general increase of the Betaproteobacteria was observed across the Reigous
371 Creek. The most abundant orders in Proteobacteria were Gallionellales and
372 Acidithiobacillales (Fig. 7B). The Proteobacteria order distribution varied among samples
373 with predominance of Acidithiobacillales in the samples from stations S5 and S1, followed by
374 the dominance of Gallionellales in the majority of the others samples. The Actinobacteria
375 were mainly dominated by Acidimicrobiales in almost all the samples except for CONF,
376 where relatively equal proportion of Acidimicrobiales and Actinomycetales were retrieved
377 (Fig. 7C). The same pattern was observed within the Acidobacteria with a net predominance
378 of Acidobacteria Gp1 in all the samples, except for CONF where a higher diversity was
379 encountered (Fig. 7D). Analysis of the Firmicutes phylum revealed dominance of the
380 *Clostridiales* and *Bacillales* orders varying in proportion within samples (Fig. 7E).

381 A relatively small number of OTUs dominated all sites (Fig. 8). The most abundant
382 OTUs, previously found in this AMD, were phylogenetically related to *Gallionella*
383 *ferruginea*, *Acidithiobacillus ferrooxidans*, *Thiobacillus* sp. and other relatively less abundant
384 phylotypes (Fig. 8) and showed variations in their relative abundances over the sampling
385 period. *Gallionella ferruginea* was present in almost all samples and showed relatively high
386 proportion especially in the stations GAL and COWG. In contrast, *Acidithiobacillus*
387 *ferrooxidans* constituted a very minor fraction of the microbial community except at S5 were

388 this species was dominant. A gradual increased of *A. ferrooxidans* along the stream was also
389 observed (Fig. 8).

390

391 ***Linking community structure to environmental variables***

392 A canonical correspondence analysis (CCA) of the whole data set was performed in
393 order to elucidate the main relationships between bacterial community and physicochemical
394 variables (Fig. 9). Samples plot in different areas of the diagram depending on their
395 environmental characteristics. The combination of environmental variables explained 36.7%
396 of the total variance in the bacterial community composition. CCA axis 1 was most closely
397 correlated with iron, arsenic and conductivity, separating the stations in three clusters (CONF,
398 GAL+COWG and S1+S5). The first canonical axe (CCA axis 1) explained 20.8% of the
399 variance. The environmental variables that were significantly correlated with the canonical
400 axes after performing the Monte-Carlo permutation test were arsenic (F-ratio=1.6, p=0.01),
401 temperature (F-ratio=1.7, p=0.01) and Eh (F-ratio=1.3, p=0.01). CCA axis 2 explained 15.7%
402 of the variance and separated samples according to the redox potential and temperature.

403

404 **Discussion**

405 ***Spatial and temporal variations of the environmental data set and bacterial community***

406 This study combined classical fingerprinting method (T-RFLP) with high-throughput
407 bar-coded pyrosequencing of 16S rRNA gene to investigated the diversity, spatial distribution
408 and seasonal variations of the bacterial community in the AMD of Carnoulès mine (France),
409 heavily contaminated with As.

410 Monitoring of physico-chemical parameters of the Reigous Creek confirmed previous
411 results obtained (Casiot *et al.*, 2003b, Egal *et al.*, 2010) showing intense decrease in the
412 concentrations of dissolved As, Fe and sulfate with 72% of sulfate, 96% of iron and 99% of
413 arsenic removed with increasing distance from the source (Table 1). Both microbiological
414 methods emphasized a low biodiversity in this extreme habitat. T-RFLP profiles generated for
415 the five sites showed a total of 58 T-RFs ranging from 2 to 20 T-RFs per sampling station
416 (Table 2). For Pyrosequencing data, a total of 66016 reads led to the identification of 6801
417 OTUs, including 4629 singletons representing 68% of the total number of OTUs. The
418 recovered phylotypes reflect the major trends in the libraries, but additional sequencing effort
419 would be needed to exhaustively characterize the bacterial community as revealed by the

420 relatively low coverage values and the lack of asymptote in the rarefaction curves (data not
421 show), especially for some samples from the less polluted station CONF.

422 NMDS analyses on 454 pyrosequencing data indicated that the bacterial community
423 composition significantly differed spatially between the five stations (Fig. 3). In accordance
424 with the T-RFLP data, relatively high average values of diversity were found in S1 while the
425 lowest value was encountered in GAL and S5. Canonical correspondence analysis of the
426 whole data set seemed to indicate that the environmental variables significantly correlated
427 with the variation of the bacterial communities were arsenic, temperature and Eh (Fig. 9).

428

429 *Taxonomic assignment of bacterial pyrosequencing*

430 In this survey we have been able to capture a broader phylogenetic breadth of taxa
431 than in any previous clone library-based diversity surveys of the Carnoulès AMD, including
432 sequences of several taxa not detected before. These new taxa included members of the
433 Bacteroidetes, Chlorobi, Chloroflexi, Elusimicrobia, Chlamydiae, Cyanobacteria,
434 Deinococcus-Thermus, Spirochaetes, Fibrobacteres, Fusobacteria, Gemmatimonadetes,
435 Plantctomycetes, Synergistetes, Verrumicrobia and members of the uncultured OD1-PO11-
436 TM7 clade. Interestingly, a major part of the phyla that could be assigned to taxonomic
437 groups not detected *via* clone libraries account for <1% of the sequences, explaining why they
438 have been missed with the clone library approach (Pedrós-Alió, 2006).

439 The existence of a high level of low-abundance populations (68% of singletons),
440 corresponding to the term “rare biosphere”, accounted for most of the observed phylogenetic
441 bacterial diversity. This rare biosphere may have a cosmopolitan distribution because of high
442 dispersal and low loss rates, and possibly represents a source of phylotypes that become
443 abundant when environmental conditions change (Galand *et al.*, 2009). Despite the prevalence
444 of the rare biosphere, we still know a little about its ecological and functional role however,
445 by becoming dominant, these taxa may play an important role in AMD, buffering the effects
446 of dramatic environmental shifts (Sogin *et al.*, 2006, Monchy *et al.*, 2011) and further
447 investigation will be needed to determine their possible role in this ecosystem. However, the
448 deep sequencing raised also questions about the accuracy of OTU richness estimates, as
449 sequencing errors coupled with inadequate choice of clustering algorithms can lead to
450 artificially elevated estimates of community richness (Huse *et al.*, 2010).

451 Most of the taxa detected in this study were related to orders whose genera are known
452 to be common in AMD and for a large part to be involved in the Fe, As, and S cycles namely

453 Gallionellales (Betaproteobacteria), Acidithiobacilliales (Gamaproteobacteria),
454 Desulfobacterales (Deltaproteobacteria), Desulfomonadales (Deltaproteobacteria),
455 Acidimicrobiales (Actinobacteria) or Nitrospirales (Nitrospira). The ecological role of these
456 taxa previously found in this ecosystem has been widely characterized (Bruneel *et al.*, 2005,
457 2006, 2011; Bertin *et al.*, 2011). A relatively small number of OTUs dominated all sites (Fig.
458 8). The most abundant OTUs were phylogenetically related to *Gallionella ferruginea*,
459 *Acidithiobacillus ferrooxidans*, and *Thiobacillus* sp. (Fig. 8) and were previously found in this
460 AMD, highlighting the fact that there are persistent in this ecosystem. *G. ferruginea* is an
461 oxygen-dependent ferrous iron-oxidizing bacterium that grows at circumneutral pH (Hallberg,
462 2010). Relatives of *Gallionella*, have often been detected, sometimes as the dominant group in
463 microbial mine water communities, including Carnoulès (Bruneel *et al.*, 2006, Hallberg *et al.*,
464 2006, Heinzl *et al.*, 2009). *A. ferrooxidans*, which readily dominate enrichment cultures from
465 acidic drainage waters, have been widely accepted to be a major contributor to AMD
466 generation (Hallberg, 2010). *Thiobacillus* sp. has also been frequently reported to occur in
467 AMD (Baker and Banfield, 2003). These three organisms showed variation in their relative
468 abundances over the sampling period. *G. ferruginea* was present in almost all samples and
469 showed relatively high proportion especially in the stations GAL and COWG. In contrast,
470 except at S5 where this species was dominant, *A. ferrooxidans* constituted a minor fraction of
471 the microbial community in the Reigous creek characterized by a gradual increase along the
472 stream (Fig. 8). These results corroborate the general observations of the previous studies.
473 Analysis of the bacterial community in the waters originating from the tailings with extremely
474 high levels of arsenic (until to 10000 mg L⁻¹ in 2001, Casiot *et al.*, 2003b) reveal the
475 predominance of organisms related to sulfate reducing bacteria (Giloteaux *et al.*, 2012),
476 together with *Acidithiobacillus ferrooxidans* and *Thiobacillus* spp. (Bruneel *et al.*, 2005). In
477 the Reigous creek where an important system of remediation is observed, an organism
478 affiliated to the neutrophilic Fe-oxidizing *Gallionella* dominates the communities.

479

480 **Acknowledgments**

481 The study was financed by the FRB program («Fondation pour la recherche sur la
482 Biodiversité»), the “Observatoire de Recherche Méditerranéen en Environnement” (OSU-
483 OREME). Aurélie Volant were supported by a grant from the French Ministry of Education
484 and Research. This work was performed in the framework of the “Groupement de recherche:
485 Métabolisme de l’Arsenic chez les Microorganismes” (GDR2909-CNRS).

486 **References**

- 487 Aguilera, A., Manrubia, S. C., Gómez, F., Rodríguez, N. and Amils, R. (2006). Eukaryotic
488 community distribution and its relationship to water physicochemical parameters in an
489 extreme acidic environment, Río Tinto (Southwestern Spain). *Appl Environ Microbiol* **72**:
490 5325-5330.
- 491 Amaral-Zettler, L. A., Zettler, E. R., Theroux, S. M., Palacios, C., Aguilera, A. and Amils, R.
492 (2011). Microbial community structure across the tree of life in the extreme Río Tinto.
493 *ISME J* **5**: 42-50.
- 494 Baker, B. J. and Banfield, J. F. (2003). Microbial communities in acid mine drainage. *FEMS*
495 *Microbiol Ecol* **44**: 139-152.
- 496 Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R. R. and Stoeck, T. (2011). Depicting
497 more accurate pictures of protistan community complexity using pyrosequencing of
498 hypervariable SSU rRNA gene regions. *Environ Microbiol* **13**: 340-349.
- 499 Bertin, P. N., Heinrich-Salmeron, A., Pelletier, E., Goulhen-Chollet, F., Arsène-Ploetze, F.,
500 Gallien, S. *et al.* (2011). Metabolic diversity among main microorganisms inside an
501 arsenic-rich ecosystem revealed by meta- and proteo-genomics. *ISME J* **5**: 1735-1747.
- 502 Bruneel, O., Duran, R., Koffi, K., Casiot, C., Fourçans, A., Elbaz-Poulichet, F. and Personné,
503 J.-C. (2005). Microbial diversity in a pyrite-rich tailings impoundment (Carnoulès,
504 France). *Geomicrobiol J* **22**: 249-257.
- 505 Bruneel, O., Duran, R., Casiot, C., Elbaz-Poulichet, F. and Personné, J. C. (2006). Diversity
506 of microorganisms in Fe-As-rich acid mine drainage waters of Carnoulès, France. *Appl*
507 *Environ Microbiol* **72**: 551-556.
- 508 Bruneel, O., Volant, A., Gallien, S., Chaumande, B., Casiot, C., Carapito, C. *et al.* (2011).
509 Characterization of the active bacterial community involved in natural attenuation
510 processes in arsenic-rich Creek sediments. *Microb Ecol* **61**: 793-810.
- 511 Casiot, C., Leblanc, M., Bruneel, O., Personné, J. C., Koffi, K. and Elbaz-Poulichet, F.
512 (2003a). Geochemical processes controlling the formation of As-rich waters within a
513 tailings impoundment (Carnoulès, France). *Aquat Geochem* **9**: 273-290.
- 514 Casiot, C., Morin, G., Juillot, F., Bruneel, O., Personné, J.-C., Leblanc, M. *et al.* (2003b).
515 Bacterial immobilization and oxidation of arsenic in acid mine drainage (Carnoulès creek,
516 France). *Water Res* **37**: 2929-2936.
- 517 Edwards, K. J., Bond, P. L., Druschel, G. K., McGuire, M. M., Hamers, R. J. and Banfield, J.
518 F. (2000). Geochemical and biological aspects of sulfide mineral dissolution: lessons from
519 Iron Mountain, California. *Chem Geol* **169**: 383-397.
- 520 Egal, M., Casiot, C., Morin, G., Elbaz-Poulichet, F., Cordier, M.-A. and Bruneel, O. (2010).
521 An updated insight into the natural attenuation of As concentrations in Reigous Creek
522 (southern France). *Appl Geochem* **25**: 1949-1957.
- 523 Galand, P. E., Casamayor, E. O., Kirchman, D. L. and Lovejoy, C. (2009). Ecology of the rare
524 microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci U S A* **106**: 22427-22432.
- 525 Giloteaux L., Duran R., Casiot C., Bruneel O., Elbaz-poulichet F. and Goñi-urriza M. A
526 survey of sulfate reducing bacteria in a heavily arsenic contaminated acid mine drainage
527 (Carnoulès, France). *FEMS Microbiology Ecology*. In presse.

528 Hallberg, K. B., Coupland, K., Kimura, S. and Johnson, D. B. (2006). Macroscopic streamer
529 growths in acidic, metal-rich mine waters in North Wales consist of novel and remarkably
530 simple bacterial communities. *Appl Environ Microbiol* **72**: 2022-2030.

531 Hallberg, K. B. (2010). New perspectives in acid mine drainage microbiology.
532 *Hydrometallurgy* **104**: 448-453.

533 Heinzl, E., Hedrich, S., Janneck, E., Glombitza, F., Seifert, J. and Schlömann, M. (2009).
534 Bacterial diversity in a mine water treatment plant. *Appl Environ Microbiol* **75**: 858-861.

535 Huse, S. M., Welch, D. M., Morrison, H. G. and Sogin, M. L. (2010). Ironing out the wrinkles
536 in the rare biosphere through improved OTU clustering. *Environ Microbiol* **12**: 1889-1898.

537 Johnson, D. B. (2012). Geomicrobiology of extremely acidic subsurface environments. *FEMS*
538 *Microbiol Ecol* **81**: 2-12.

539 Koffi, K., Leblanc, M., Jourde, H., Casiot, C., Pistre, S., Gouze, P. and Elbaz-Poulichet, F.
540 (2003). Reverse oxidation zoning in mine tailings generating arsenic-rich acidic waters
541 (Carnoulès, France). *Mine Water Environ* **22**: 7-14.

542 Lane D, J. (1991). *16S/23S rRNA sequencing*, Stackebrandt Stachenbrady and M Goodfellow
543 edn. New York John Wiley and Sons.

544 Morin, G., Juillot, F., Casiot, C., Bruneel, O., Personné, J. C., Elbaz-Poulichet, F. *et al.*
545 (2003). Bacterial formation of tooeleite and mixed arsenic(III) or arsenic(V)-iron(III) gels
546 in the Carnoulès acid mine drainage, France. A XANES, XRD, and SEM study. *Environ*
547 *Sci Technol* **37**: 1705-1712.

548 Morin, G. and Calas, G. (2006). Arsenic in soils, mine tailings, and former industrial sites.
549 *Elementss* **2**: 97-101.

550 Monchy, S., Sancier, G., Jobard, M., Rasconi, S., Gerphagnon, M., Chabe, M. *et al.* (2011).
551 Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA
552 cloning/sequencing and SSU tag pyrosequencing. *Environ Microbiol* **13**: 1433-1453.

553 Muyzer, G. and Ramsing, N. B. (1995). Molecular methods to study the organization of
554 microbial communities. *Water Science and Technology* **32**: 1-9.

555 Needleman, S. B. and Wunsch, C. D. (1970). A general method applicable to the search for
556 similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**: 443-453.

557 Pedrós-Alió, C. (2006). Marine microbial diversity: can it be determined? *Trends Microbiol*
558 **14**: 257-263.

559 Rodier, J., Broutin J.P., Chambon P., Champsaur H., Rodi L. (1996). L'Analyse des Eaux.
560 *Dunod, Paris*: p 1383.

561 Samanta, G. and Clifford, D. A. (2005). Preservation of inorganic arsenic species in
562 groundwater. *Environ Sci Technol* **39**: 8877-8882.

563 Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B. *et al.*
564 (2009). Introducing mothur: open-source, platform-independent, community-supported
565 software for describing and comparing microbial communities. *Appl Environ Microbiol*
566 **75**: 7537-7541.

567 Smith, C. J., Danilowicz, B. S., Clear, A. K., Costello, F. J., Wilson, B. and Meijer, W. G.
568 (2005). T-Align, a web-based tool for comparison of multiple terminal restriction fragment
569 length polymorphism profiles. *FEMS Microbiol Ecol* **54**: 375-380.

- 570 Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., Huse, S. M., Neal, P. R. *et al.*
571 (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc*
572 *Natl Acad Sci U S A* **103**: 12115-12120.
- 573 R Development Core Team (2012). *R: A language and environment for statistical computing.*
574 R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL
575 <http://www.R-project.org/>.
- 576 Ueno, S., Le Provost, G., Leger, V., Klopp, C., Noirot, C., Frigerio, J.-M. *et al.* (2010).
577 Bioinformatic analysis of ESTs collected by Sanger and pyrosequencing methods for a
578 keystone forest tree species: oak. *BMC Genomics* **11**: 650.

Table 1. Physicochemical characteristics of the water at each sampling station and time.

Sampling stations	Sampling period	pH	T (°C)	DO (mg/l)	Eh* (mV)	Cond (µs/cm)	SO ₄ ²⁻ (mg/l)	Fe (total) (mg/l)	Fe(II) (mg/l)	As(total) (mg/l)	As(III) (mg/l)
S5	11/2007	2.0 (±0.05)	12.2 (±0.1)	0.38 (±0.1)	620 (±1)	2228 (±22)	13464 (±1346)	5433 (±272)	3732 (±187)	546 (±27)	115 (±6)
	02/2008	1.7 (±0.05)	12.7(±0.1)	1.59 (±0.1)	644 (±1)	3670 (±37)	31339 (±3134)	9448 (±472)	3912 (±196)	764 (±38)	97 (±5)
	10/2008	2.0 (±0.05)	15.4 (±0.1)	0.17 (±0.1)	623 (±1)	17020 (±17)	14078 (±1408)	5151 (±258)	3891 (±195)	287 (±14)	164 (±8)
	03/2009	2.4 (±0.05)	14.1 (±0.1)	2.79 (±0.1)	550 (±1)	8780 (±88)	19291 (±1929)	2482 (±124)	2275 (±114)	299 (±15)	257 (±13)
	11/2009	3.4 (±0.05)	13.8 (±0.1)	0.49 (±0.1)	451 (±1)	6860 (±69)	3806 (±381)	2154 (±108)	2000 (±100)	382 (±19)	365 (±18)
	01/2010	3.4 (±0.05)	11.6 (±0.1)	0.29 (±0.1)	462 (±1)	6970 (±70)	2504 (±250)	2175 (±109)	2175 (±109)	360 (±18)	337 (±17)
	Mean±SD	2.5±0.8	13.3±1.4	0.95±0.9	558±85	7588±5203	14080±10630	4474±2855	2998±935	440 ±184	222±114
S1	11/2007	3.3 (±0.05)	12.4 (±0.1)	1.47 (±0.1)	534 (±1)	4716 (±47)	1561 (±156)	1040 (±52)	1040 (±52)	148 (±7)	98 (±5)
	02/2008	3.1 (±0.05)	12.9 (±0.1)	0.37 (±0.1)	510 (±1)	4162 (±42)	2302 (±230)	923 (±46)	923 (±46)	88 (±4)	64 (±3)
	10/2008	3.1 (±0.05)	15.4 (±0.1)	0.38 (±0.1)	476 (±1)	5460 (±55)	3972 (±397)	1665 (±83)	1471 (±74)	281 (±14)	258 (±13)
	03/2009	2.9 (±0.05)	13.9 (±0.1)	0.58 (±0.1)	501 (±1)	4930 (±49)	4331 (±433)	1155 (±58)	1155 (±58)	190 (±10)	166 (±8)
	11/2009	3.3 (±0.05)	14.3 (±0.1)	0.46 (±0.1)	476 (±1)	5818 (±58)	2226 (±223)	1904 (±95)	1289 (±64)	222 (±11)	213 (±11)
	01/2010	2.5 (±0.05)	12.8 (±0.1)	1.79 (±0.1)	577 (±1)	5640 (±56)	1700 (±170)	1214 (±61)	1214 (±61)	118 (±6)	104 (±5)
	Mean±SD	3.0±0.3	13.6±1.1	0.84±0.6	512±39	5121±631	2682±1180	1317±383	1182±192	175±71	150±75
COWG	11/2007	3.4 (±0.05)	10.6 (±0.1)	7.15 (±0.1)	522 (±1)	4022 (±40)	2025 (±203)	675 (±34)	675 (±34)	59 (±3)	40 (±2)
	02/2008	3.2 (±0.05)	9.9 (±0.1)	3.99 (±0.1)	550 (±1)	3031 (±30)	1812 (±181)	440 (±22)	440 (±22)	36 (±2)	26 (±1)
	10/2008	3.1 (±0.05)	14.3 (±0.1)	5.72 (±0.1)	524 (±1)	4900 (±49)	3409 (±341)	1248 (±62)	1220 (±61)	154 (±8)	133 (±7)
	03/2009	2.9 (±0.05)	11.6 (±0.1)	4.01 (±0.1)	499 (±1)	4190 (±42)	3313 (±331)	815 (±41)	815 (±41)	124 (±7)	101 (±5)
	11/2009	3.3 (±0.05)	13.1 (±0.1)	7.86 (±0.1)	542 (±1)	5215 (±52)	3300 (±330)	1736 (±87)	1440 (±72)	194 (±10)	161 (±8)
	01/2010	2.8 (±0.05)	9.5 (±0.1)	7.96 (±0.1)	585 (±1)	4250 (±43)	1874 (±187)	812 (±41)	812 (±41)	80 (±4)	70 (±4)
	Mean±SD	3.1±0.2	11.5±1.9	6.12±1.8	537±29	4268±761	2622±791	954±465	900±366	110±61	88±53

GAL	11/2007	3.6 (±0.05)	7.5 (±0.1)	10.17 (±0.1)	545 (±1)	2662 (±27)	1530 (±153)	363 (±18)	191 (±10)	25 (±1)	18 (±1)
	02/2008	3.2 (±0.05)	8.5 (±0.1)	9.20 (±0.1)	602 (±1)	2220 (±22)	1270 (±127)	263 (±13)	263 (±13)	24 (±1)	17 (±1)
	10/2008	3.0 (±0.05)	12.2 (±0.1)	7.72 (±0.1)	571 (±1)	4540 (±45)	3378 (±338)	962 (±48)	962 (±48)	80 (±4)	57 (±3)
	03/2009	3.2 (±0.05)	11.2 (±0.1)	8.58 (±0.1)	559 (±1)	3040 (±30)	2222 (±222)	336 (±17)	336 (±17)	73 (±4)	55 (±3)
	11/2009	3.1 (±0.05)	11.2 (±0.1)	8.48 (±0.1)	560 (±1)	4524 (±45)	3004 (±300)	1063 (±53)	968 (±48)	52 (±3)	40 (±2)
	01/2010	3.2 (±0.05)	5.0 (±0.1)	11.20 (±0.1)	596 (±1)	2340 (±23)	2226 (±223)	329 (±16)	329 (±16)	31 (±2)	26 (±1)
	Mean±SD	3.2±0.2	9.3±2.8	9.23±1.3	572±22	3221±1055	2272±815	552±359	508±358	48±25	35±18
CONF	11/2007	3.3 (±0.05)	5.0 (±0.1)	11.46 (±0.1)	671(±1)	1566 (±16)	810 (±81)	47 (±2)	< DL	8.10 ⁻¹ (±0.04)	4.10 ⁻¹ (±0.1)
	02/2008	3.0 (±0.05)	7.1 (±0.1)	10.82 (±0.1)	689 (±1)	1514 (±15)	630 (±63)	44 (±2)	< DL	ND	ND
	10/2008	3.7 (±0.05)	11.4 (±0.1)	8.60 (±0.1)	440 (±1)	1692 (±17)	938 (±94)	6 (±0.3)	< DL	3.10 ⁻² (±0.01)	2.10 ⁻² (±0.01)
	03/2009	3.3 (±0.05)	9.2 (±0.1)	10.47 (±0.1)	645 (±1)	1669 (±17)	274 (±27)	80 (±4)	61 (±3)	6 (±0.3)	4 (±0.2)
	11/2009	3.2 (±0.05)	11.2 (±0.1)	9.43 (±0.1)	715 (±1)	1729 (±17)	853 (±85)	23 (±1)	< DL	3.10 ⁻¹ (±0.02)	1.10 ⁻² (±0.01)
	01/2010	3.1 (±0.05)	3.5 (±0.1)	12.08 (±0.1)	650 (±1)	1501 (±15)	1013 (±101)	81 (±4)	< DL	4 (±0.2)	3 (±0.2)
	Mean±SD	3.2±0.3	7.9±3.3	10.48±1.3	635±99	1612±97	753±268	47±30	61±3	3 ±2.6	2±2

*Redox potential (Eh) was measured relatively to the Ag/AgCl/KCl electrode

SD: standard deviation

ND: not determined

DL: detection limit

Table 3. Estimated OTU richness, diversity indices and estimated sample coverage for each 16S rRNA libraries. Results are presented for the full data sets reads (Full dataset), and for reduced data sets randomly resampled to equal sample size (Reduced dataset).

Stations Period	FULL DATASET					REDUCED DATASET				
	No. Reads	Obs. OTUs ^a	Chao1	Shannon ^b	Coverage ^c	No. Reads	Obs. OTUs ^a	Chao1	Shannon ^b	Coverage ^c
S5										
11/2007	2090	216	549 (418; 764)	3.17 (3.08; 3.25)	93	485	88	164 (123; 255)	3.34 (3.20; 3.48)	91
10/2008	2523	87	230 (152; 400)	1.50 (1.43; 1.57)	98	485	35	66 (44; 134)	1.63 (1.48; 1.78)	96
03/2009	2838	178	385 (301; 527)	2.22 (2.14; 2.30)	96	485	46	77 (57; 130)	2.03 (1.89; 2.18)	95
11/2009	1086	161	551 (370; 890)	3.24 (3.13; 3.35)	90	485	91	154 (120; 228)	3.31 (3.16; 3.45)	91
01/2010	2354	195	592 (419; 897)	3.42 (3.35; 3.49)	95	485	75	142 (103; 233)	3.39 (3.27; 3.50)	92
S1										
11/2007	2724	441	1346 (1066; 1751)	3.88 (3.79; 3.97)	89	485	134	230 (187; 308)	3.75 (3.59; 3.91)	85
10/2008	5006	871	2683 (2286; 3191)	4.38 (4.31; 4.44)	87	485	132	350 (250; 533)	3.84 (3.70; 3.97)	82
03/2009	2057	366	1119 (877; 1475)	3.97 (3.88; 4.06)	88	485	127	249 (193; 351)	3.79 (3.65; 3.93)	84
11/2009	2091	397	1024 (839; 1285)	3.58 (3.47; 3.69)	87	485	137	346 (253; 516)	3.57 (3.40; 3.74)	81
01/2010	4622	855	2147 (1855; 2523)	4.84 (4.77; 4.90)	88	485	179	363 (288; 488)	4.45 (4.32; 4.57)	78
COWG										
11/2007	704	201	393 (320; 510)	4.12 (3.99; 4.25)	82	485	120	164 (143; 206)	3.78 (3.64; 3.92)	88
02/2008	2214	312	852 (673; 1120)	3.28 (3.18; 3.38)	90	485	101	231 (166; 362)	3.08 (2.91; 3.25)	87
10/2008	2857	305	1228 (880; 1787)	3.07 (2.99; 3.15)	93	485	95	197 (146; 296)	3.04 (2.89; 3.20)	87
03/2009	2756	510	1702 (1358; 2186)	4.11 (4.02; 4.19)	87	485	151	328 (253; 456)	3.92 (3.77; 4.06)	79
11/2009	1323	202	582 (431; 830)	2.82 (2.69; 2.94)	89	485	118	299 (214; 461)	3.05 (2.85; 3.24)	84
01/2010	1847	145	574 (368; 972)	1.57 (1.47; 1.68)	94	485	66	170 (112; 300)	1.60 (1.41; 1.80)	91
GAL										
11/2007	1638	225	523 (412; 699)	2.26 (2.13; 2.38)	91	485	89	136 (112; 185)	2.23 (2.03; 2.44)	90
02/2008	1683	177	762 (502; 1231)	1.58 (1.46; 1.70)	92	485	62	125 (90; 204)	1.39 (1.20; 1.58)	92
10/2008	2094	238	609 (471; 828)	2.65 (2.55; 2.75)	93	485	115	269 (199; 399)	2.52 (2.31; 2.74)	83

03/2009	2489	358	1083 (853; 1421)	2.49 (2.38; 2.60)	90	485	91	155 (122; 222)	2.75 (2.57; 2.93)	89
11/2009	1682	287	751 (594; 990)	2.65 (2.52; 2.79)	88	485	123	373 (257; 591)	2.76 (2.54; 2.97)	81
01/2010	2254	269	939 (697; 1318)	2.12 (2.01; 2.23)	91	485	85	228 (153; 383)	2.03 (1.82; 2.23)	88
CONF										
11/2007	3259	934	1996 (1765; 2291)	5.93 (5.87; 5.98)	84	485	240	483 (396; 618)	5.06 (4.96; 5.17)	69
02/2008	1770	239	956 (681; 1402)	1.94 (1.82; 2.06)	89	485	64	150 (102; 258)	1.50 (1.31; 1.69)	91
10/2008	1735	551	1166 (1008; 1380)	5.03 (4.93; 5.12)	80	485	187	331 (274; 424)	4.45 (4.31; 4.59)	78
03/2009	527	124	454 (301; 738)	2.27 (2.06; 2.48)	81	485	82	192 (136; 307)	1.84 (1.64; 2.04)	88
11/2009	1996	701	1719 (1476; 2039)	5.45 (5.37; 5.54)	77	485	236	502 (408; 648)	4.87 (4.74; 4.99)	68
01/2010	5725	1344	3829 (3383; 4372)	4.79 (4.72; 4.87)	84	485	207	540 (416; 738)	4.26 (4.09; 4.43)	69

^a OTUs were defined at 97% cutoff

^b Coverage: sum of probabilities of observed classes calculated as $(1 - (n/N))$, where n is the number of singleton sequences and N is the total number of sequences

^c Takes into account the number and evenness of species

Values in brackets are 95% confidence intervals.

Table 4. Relative abundance (in %) of total sequences of bacterial 16S rRNA genes from each sample assigned to different phyla. N7: November 2007; O8: October 2008; M9: March 2009; N9: November 2009; J10: January 2010.

Phylum	Total	S5					S1					COWG					GAL					CONF							
		N7	O8	M9	N9	J10	N7	O8	M9	N9	J10	N7	F8	O8	M9	N9	J10	N7	F8	O8	M9	N9	J10	N7	F8	O8	M9	N9	J10
Proteobacteria	67.5	68.1	85	68.8	68.5	40.7	95.8	20.5	43.4	72.5	46.6	70.9	65.0	79.6	58.4	82.8	93.2	90.4	93.9	89.7	85.2	82.2	92.2	54.3	94.2	75.7	92.4	62.8	74.0
Actinobacteria	5.4	4.3	0.7	22.5	3.1	13.3	8.6	12.9	19.6	4.0	7.1	2.8	0.9	1.5	10.6	1.1	0.3	0.7	0.6	0.3	1.5	1.3	0.5	2.6	0.5	2.5	0.6	4.6	1.8
Acidobacteria	4.9	2.3	0	0.7	2.4	4.5	5.3	18.7	18.1	5.8	2.0	6.8	2.3	5.0	11.2	4.2	0.8	2.0	0.7	1.4	3.1	4.2	1.3	4.1	0.9	4.9	1.1	4.4	3.6
Firmicutes	4.5	15.8	10.9	5.5	14.8	20.3	6.0	0.7	6.0	1.9	2.8	5.1	1.5	4.2	6.6	4.3	2.3	1.0	1.4	1.1	3.2	3.4	1.0	1.0	0.1	2.4	0.2	10.1	3.0
Bacteroidetes	2.1	4.7	0	0.5	2.7	5.7	3.6	0.7	1.3	2.0	0.8	0.9	0.2	0.7	3.0	1.0	0.2	0.9	0.3	0.4	1.6	1.8	0.4	7.3	0.7	5.6	2.1	3.2	4.0
Chlorobi	1.5	3.8	0	0	7.2	14.9	1.1	0.6	1.5	1.4	0.6	2.3	0.4	1.0	1.5	1.2	0.8	0.4	0.5	0.3	0.5	1.0	0.5	2.5	0.4	1.5	0.6	1.4	0.8
Nitrospira	1.2	0.4	3.3	1.3	0	0.1	2.4	0	0.6	0.3	1.1	1.3	0.2	0.4	0.6	0.5	0.3	1.1	0.2	0.6	0.4	0.5	0.3	4.1	1.5	0.5	0.4	3.2	3.1
OD1	0.3	0	0	0	0	0	0.2	<0.1	<0.1	0	1.4	0.1	0.0	0	0	0	0.1	0.1	0	0.2	0	0	0.2	3.8	0	0.1	0	0.1	0.2
Elusimicrobia	0.2	<0.1	0	0	0	0	0.6	0.3	0.3	0.5	0.1	0.3	0.2	0.1	1.1	0.4	0	0	0.1	0	0.6	0.8	0.1	<0.1	0.1	0.4	0	0.0	0.3
TM7	0.1	0	0	0.1	0	0	0.2	<0.1	<0.1	0	0.3	0	0	0	0.1	0	0	0.1	0.1	0	0.1	0.1	<0.1	1.5	0	0	0	0.1	0.2
Chloroflexi	0.1	0	0	0	0	0	0.4	<0.1	0	0	0	0.3	<0.1	0	<0.1	0	0.1	0.1	0	0.2	<0.1	0	0	0.3	0	0.3	0	0.7	0.2
Verrucomicrobia	0.1	0	0	0	0	0	0	<0.1	0	0.1	<0.1	0	<0.1	0	0.1	0	0	0	0	0	0.1	0	<0.1	0.6	0.2	0.1	0.2	0.5	0.1
Gemmatimonadetes	0.1	0	0	0.2	0	0	0	<0.1	<0.1	0	<0.1	0	0	<0.1	<0.1	0	0	0	0	0.2	0	0.1	0	0.3	0	0.2	0.2	0.2	0.2
Spirochaetes	<0.1	0	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0.1	0	0	0	<0.1	<0.1	0	0	0.6	0	0.3	0	0	0.1
Fusobacteria	<0.1	0	0	0	0	0	0	0	0	<0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0.2	0.5
Cyanobacteria	<0.1	<0.1	0	0	0	0	0.3	0	0	<0.1	0	0	0	0	<0.1	0	0	0.2	0	0	0	0	0	0.3	0	0.1	0	0	0.1
OP11	<0.1	0	0	0	0	0	0	<0.1	0	0	0.5	0	0	0	0	0	0	0	0.1	0	0	0	0	0.1	0	0	0	0	0
Chlamydiae	<0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0
Planctomycetes	<0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	0	0.1	<0.1
Deinococcus-Thermus	<0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<0.1	0	0	0	0	0.1	0
Fibrobacteres	<0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<0.1

Synergistetes	<0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<0.1	
BRC1	<0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<0.1	
unclassified	11.7	0.4	0.2	0.4	1.3	0.4	10.0	45.4	9.1	11.2	36.6	9.2	29.1	7.5	6.6	4.5	2.0	3.2	2.2	5.5	3.7	4.7	3.5	16.0	1.4	5.3	2.3	8.4	7.8

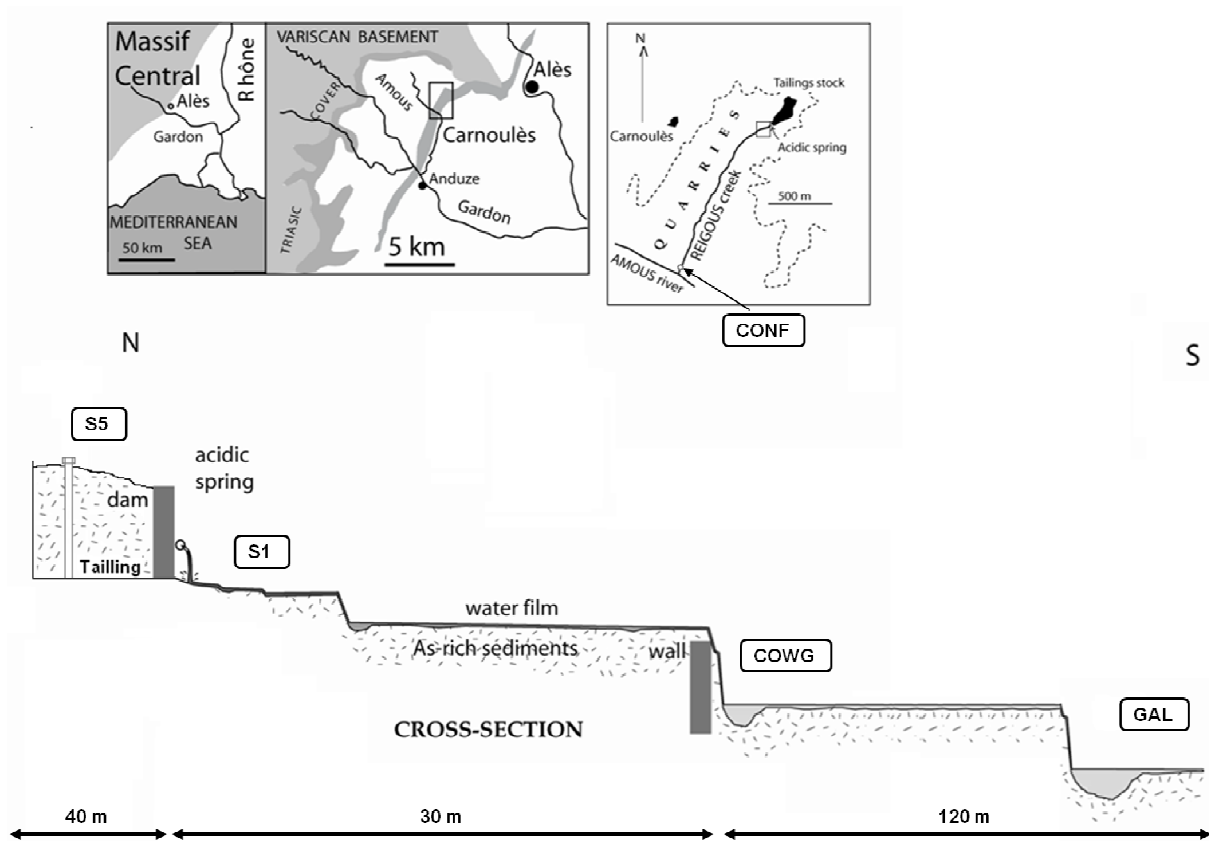


Fig. 1. Map of the Carnoulès mining site and location of sampling stations.

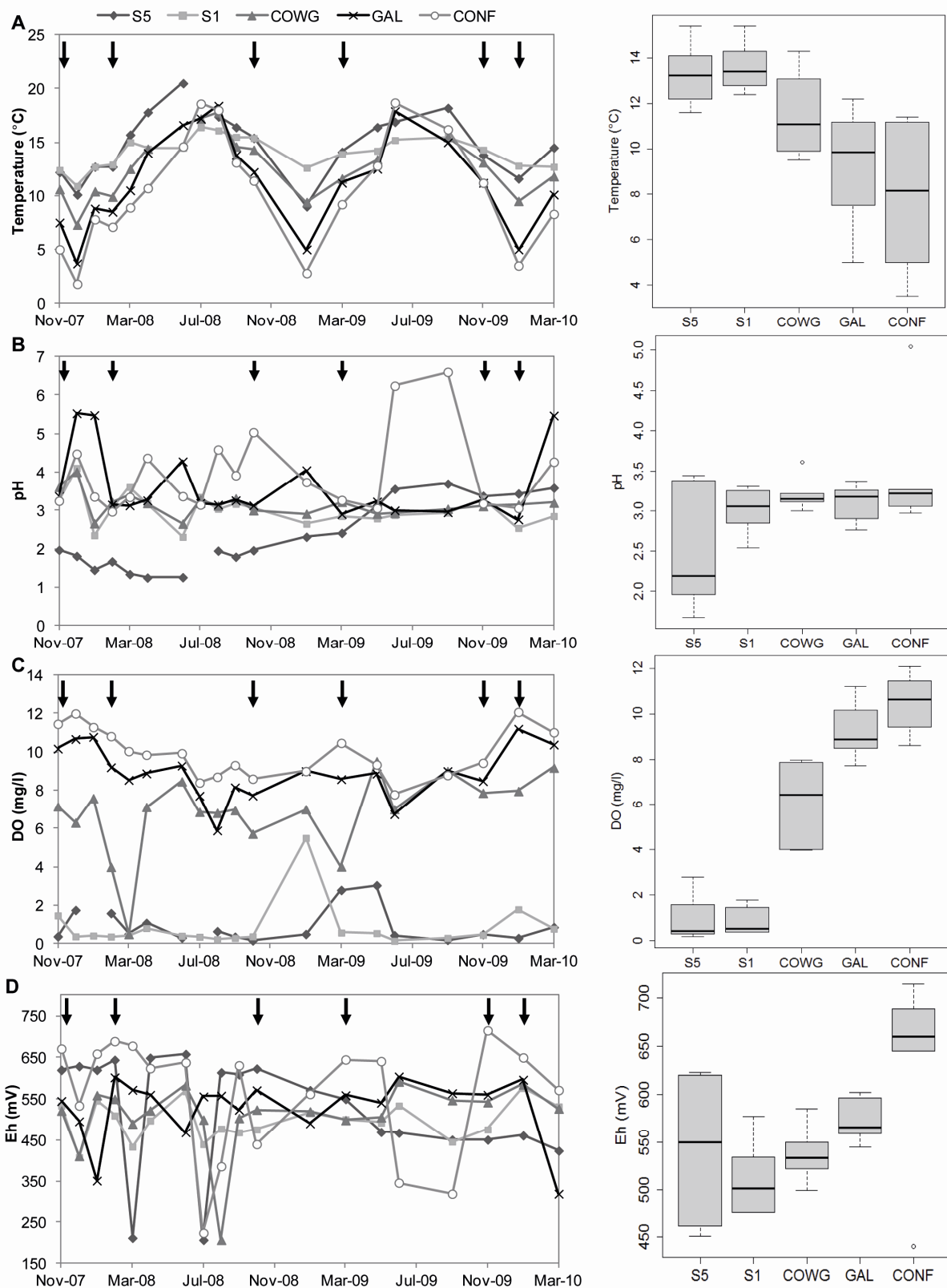


Fig. 2. Variations in the main physicochemical parameters over the course of the study and boxplot of each variable per sampling stations. Arrows indicate samples for T-RFLP and pyrosequencing analysis. Note that some data were absent as shown by a non continuous line in the graph (Part 1).

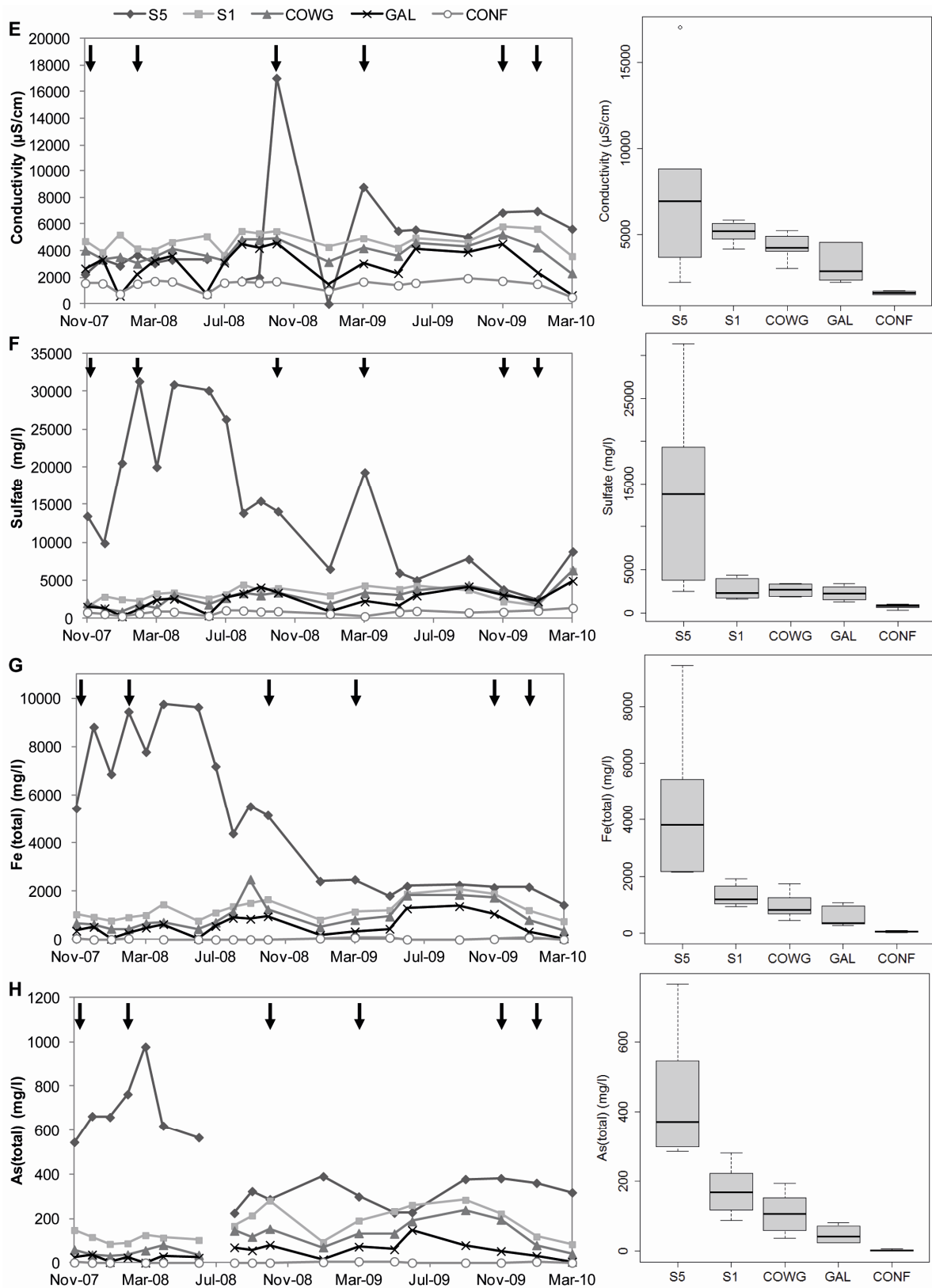


Fig. 2. Variations in the main physicochemical parameters over the course of the study and boxplot of each variable per sampling stations. Arrows indicate samples for T-RFLP and pyrosequencing analysis. Note that some data were absent as shown by a non continuous line in the graph (Part 2).

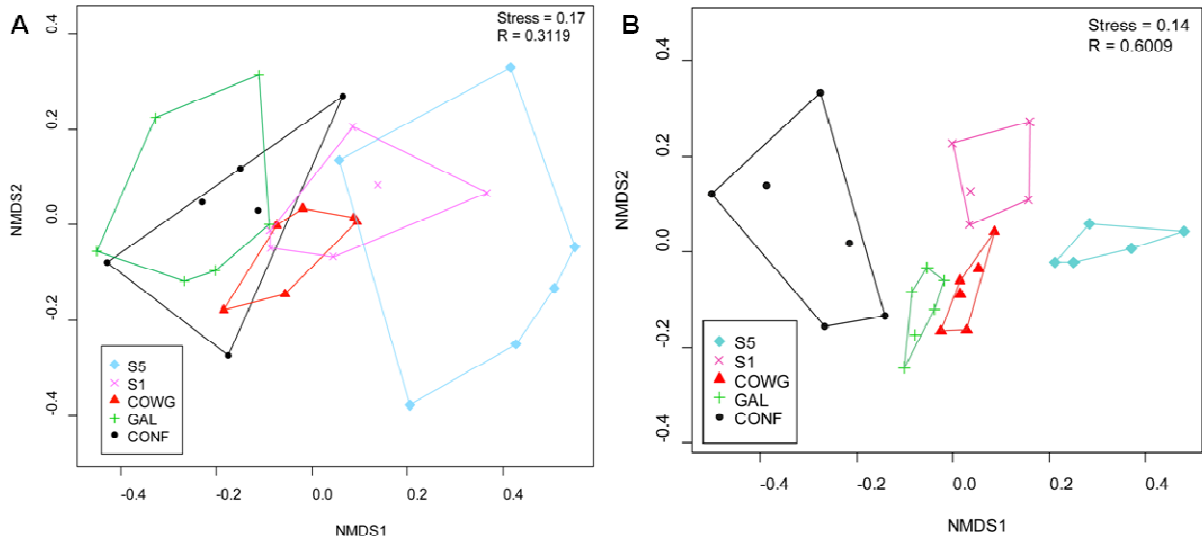


Fig. 3. Nonmetric MultiDimensional Scaling (NMDS) analysis of the bacterial community compositions estimated by (A) T-RFLP and (B) 454 pyrosequencing.

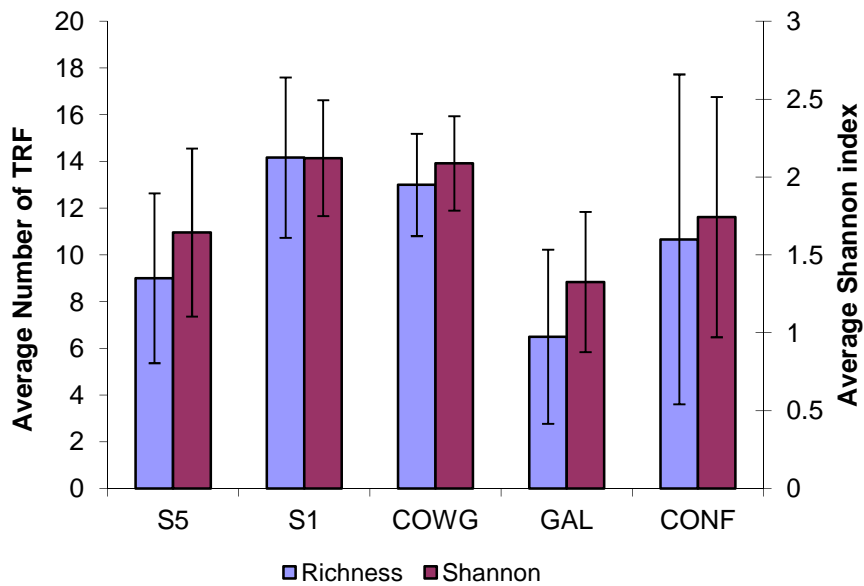


Fig. 4. Average richness and Shannon index per group \pm standard deviation based on T-RFLP profiles.

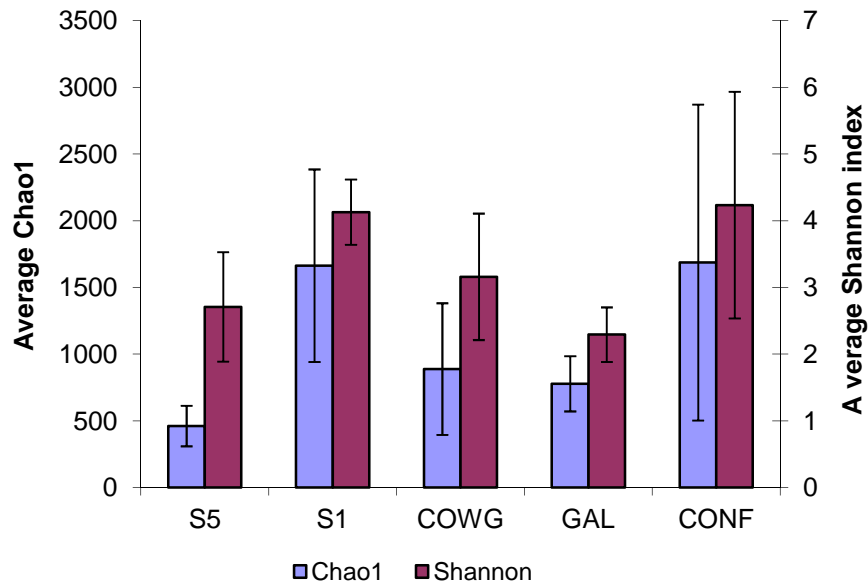


Fig. 5. Average Chao1 and Shannon index per group \pm standard deviation based on pyrosequencing reads.

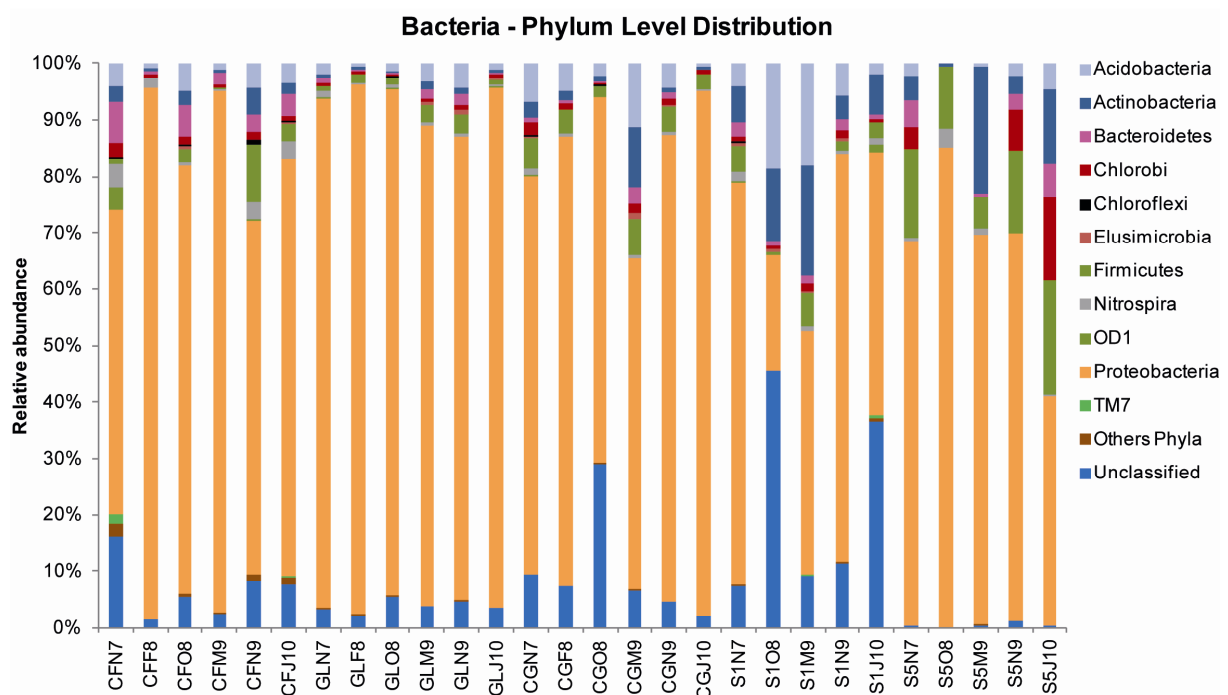
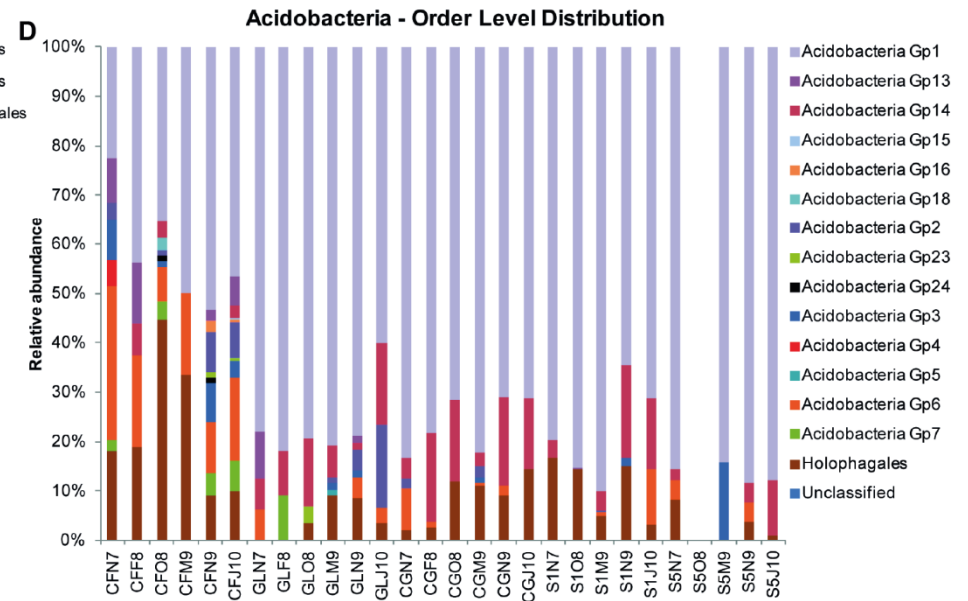
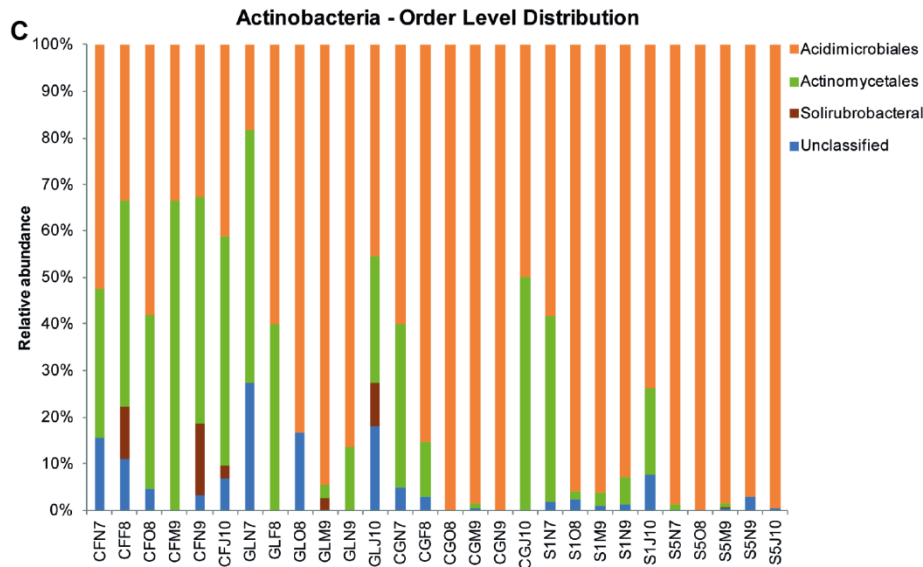
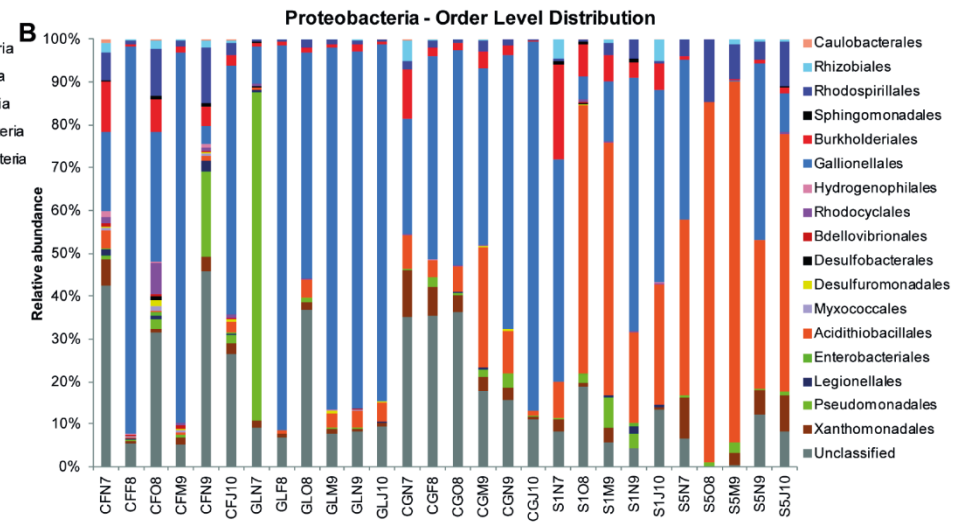
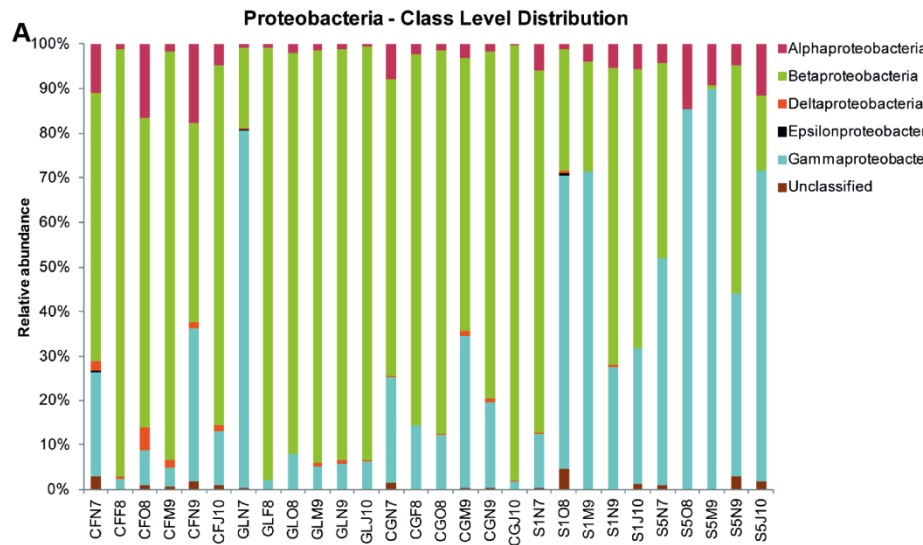


Fig. 6. Taxonomic classification of bacterial reads retrieved from the different water samples at phylum level using RDP classifier with a confidence threshold of 80%. Others phyla refer to minor groups with relative abundance less than 1% related to BRC1, Chlamydiae, Cyanobacteria, Deinococcus-Thermus, Fibrobacteres, Fusobacteria, Gemmatimonadetes, OP11, Plantctomycetes, Spirochaetes, Synergistetes and Verrumicrobia phyla.



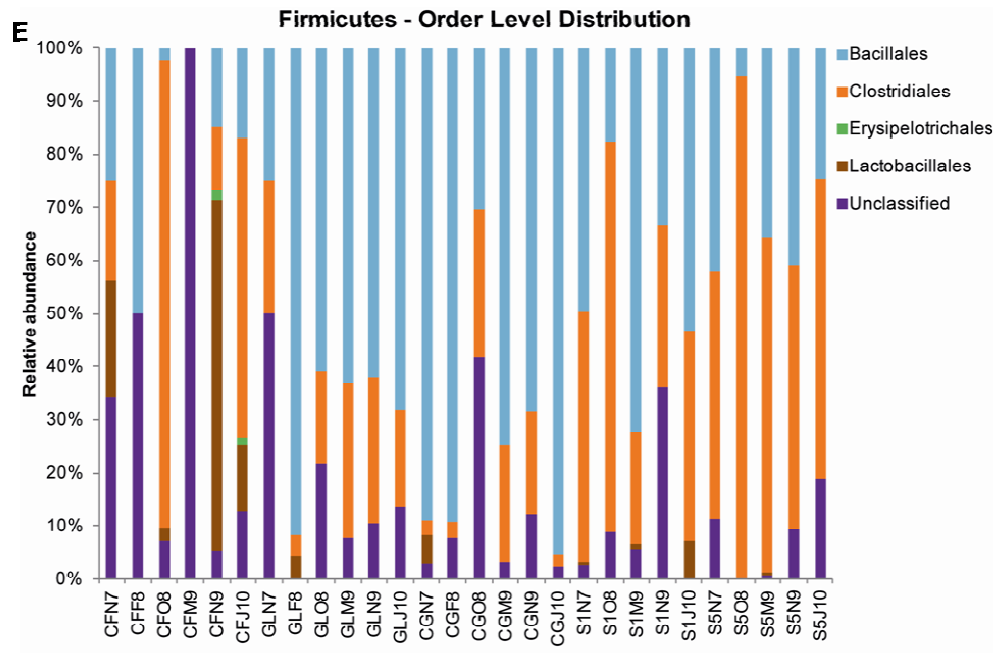


Fig. 7. Composition of different phyla based on classification of 16S rRNA sequences of bacteria from each sample using RDP Classifier: Proteobacteria classes (A), Proteobacteria orders (B), Actinobacteria orders (C), Acidobacteria orders (D) and Firmicutes orders (E).

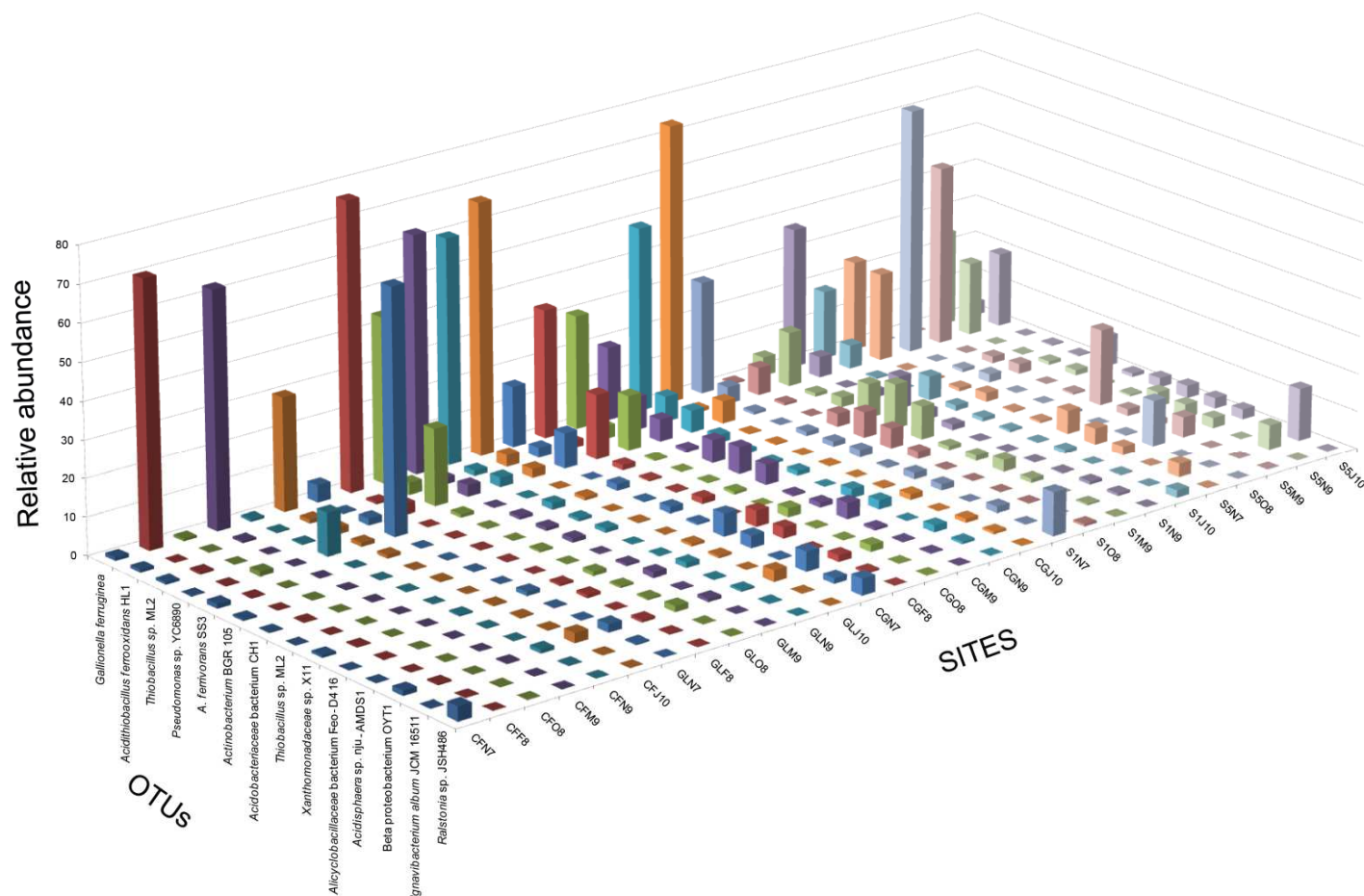


Fig. 8. Histogram of relative abundances of dominant OTUs at Carnoulès study sites (*G. ferruginea*: NC_014394; *A. ferrooxidans* strain HL1: JF815535; *Thiobacillus* sp. ML2-16: DQ145970; *Pseudomonas* sp. YC6890: HM800927; *A. ferrivorans* SS3: NC_015942; *Actinobacterium* BGR 105: GU168008; *Acidobacteriaceae* bacterium CH1: DQ355184; *Xanthomonadaceae* sp. X11: FR874227; *Alicyclobacillaceae* bacterium Feo-D4-16-CH: FN870323; *Acidisphaera* sp. nju-AMDS1: FJ915153; Beta proteobacterium OYT1: AB720116; *Ignavibacterium album* JCM 16511: NC_017464).

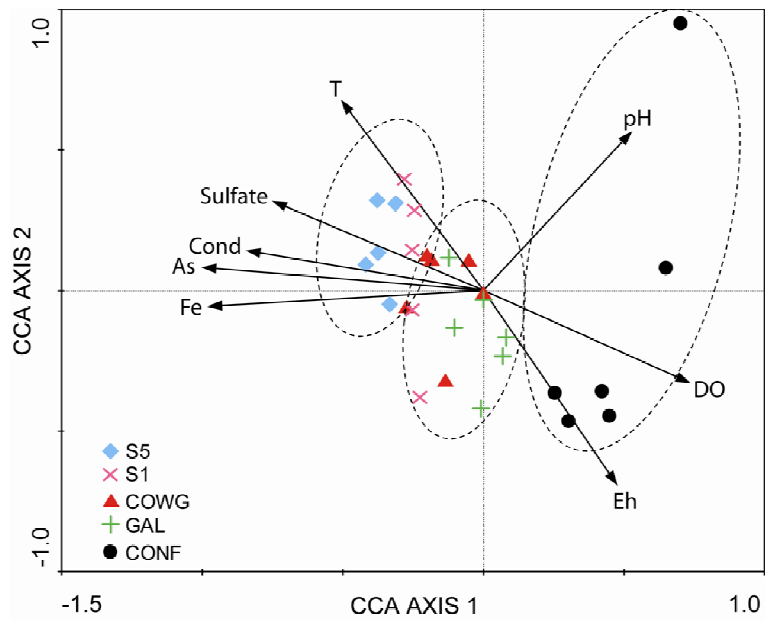


Fig. 9. Canonical Correspondence Analysis (CCA) correlating the bacterial communities structure from each station with arsenic (As), iron (Fe), conductivity (Cond), temperature (T), dissolved oxygen (DO), redox (Eh), pH and sulfate. Three main clusters are highlighted. The bacterial community structures correspond to OTUs abundances from pyrosequencing data.

CONCLUSIONS GÉNÉRALES ET PERSPECTIVES

Conclusions générales

Les drainages miniers acides sont des environnements extrêmes généralement caractérisés par une forte acidité, des concentrations élevées en métaux et métalloïdes et une faible concentration en carbone organique dissous (Kolmert & Johnson, 2001; Hallberg, 2010). Malgré cette toxicité, divers microorganismes sont capables de s'y développer et de nombreuses études ont souligné le rôle crucial joué par ces microorganismes du fait de leurs activités métaboliques impliquées dans les cycles biogéochimiques des éléments (Haferburg & Kothe, 2007; Falkowski *et al.*, 2008) à l'origine notamment des processus de génération et/ou remédiation des drainages miniers acides. La compréhension de ces processus nécessite par conséquent des étapes préalables de caractérisation de la structure et du fonctionnement, y compris les interactions synécologiques, des communautés microbiennes présentes dans ces environnements. Pour ce faire, il est indispensable de prendre en compte l'ensemble de la communauté microbienne comprenant les organismes procaryotes et eucaryotes. Or, jusqu'à présent, les études réalisées sur les DMA se sont presque exclusivement focalisées sur les communautés procaryotes et plus particulièrement sur les bactéries (Aguilera *et al.*, 2006), négligeant ainsi une part importante de la diversité représentée par les Archaea et surtout par les eucaryotes.

La majeure partie du travail de ce doctorat a donc consisté à réaliser un inventaire taxonomique de la diversité microbienne présente au niveau du DMA de Carnoulès en s'attachant à prendre en compte l'ensemble des trois domaines du vivant : Bactérie, Archaea et Eucaryote.

Ainsi, la première partie de ce travail s'est axé sur la description des communautés bactériennes présentes à la fois dans l'eau et dans les sédiments. Ce travail a été couplé à une approche de métaprotéomique qui a permis la détection des communautés bactériennes fonctionnellement actives. La combinaison entre l'approche métaprotéomique et l'approche génomique a ainsi permis de mettre en évidence la présence de bactéries connues pour leurs implications dans le cycle du fer, de l'arsenic et du soufre ainsi que leur activité métabolique *via* l'expression active de protéines. Ce travail a également montré que la structure des communautés n'était pas fondamentalement différente entre l'eau et les sédiments. Enfin, les

résultats de cette étude ont permis de renforcer l'hypothèse selon laquelle les bactéries pourraient être fortement impliquées dans le processus de remédiation observé *in situ* mais également dans la stabilité des sédiments à travers la modification du ratio entre les formes oxydées et réduites du fer, de l'arsenic et des composés soufrés, qui favoriserait ainsi la formation et/ou la dissolution des précipités d'hydroxysulfate de Fe(III)-As(V). Cette étude a fait l'objet d'une publication dans *Microbiology Ecology* (Bruneel *et al.*, 2011).

Dans un second temps, nous nous sommes focalisé sur la communauté d'Archaea et sa dynamique temporelle dans les sédiments. Jusqu'à présent, très peu d'informations étaient disponibles sur cette communauté car seule une étude avait mis en évidence la présence de *Ferroplasma acidiphilum* et de séquences affiliées à des *Thermoplasmatales* non cultivés (Bruneel *et al.*, 2008). Les résultats obtenus ont donc apporté d'importantes informations sur cette communauté contribuant ainsi à l'amélioration de notre compréhension de cet écosystème. En effet, l'inventaire taxonomique réalisé a révélé la présence d'Archaea probablement impliquées dans divers cycles biogéochimiques tels que les cycles du carbone via la méthanogénèse (*Methanomassiliicoccus luminyensis*), de l'azote (*Candidatus Nitrososphaera viennensis* et *Candidatus Nitrosopumilus* sp.) ainsi que du fer et du soufre (*Thermoplasmatales*). Cette étude a ainsi permis de mettre en évidence des organismes susceptibles de participer au processus de remédiation observé *in situ*, conjointement avec les bactéries. Cette communauté pourrait également être impliquée, avec les bactéries, dans la stabilité des sédiments. Néanmoins, la présence de nombreuses séquences affiliées à des Archaea non cultivées ont également souligné l'importance d'améliorer les stratégies de culture à la base de la caractérisation taxonomique, physiologique et métabolique des nouvelles espèces ainsi détectées. Par ailleurs, cette étude a également révélé une variation temporelle significative des communautés d'Archaea caractérisée par un changement brutal de la composition microbienne qui pourrait être lié à des variations importantes des paramètres physicochimiques comme l'apport massif de sables à sulfure sur le site suite à des événements pluvieux intenses et soulignerait ainsi la nécessité de prendre en compte les paramètres environnementaux dans l'appréhension des communautés microbiennes. Cette étude a fait l'objet d'une publication dans *Extremophiles* (Volant *et al.*, 2012).

Dans une troisième partie, nous nous sommes intéressés à la dynamique spatiale de la communauté eucaryote qui n'avait jusqu'à présent jamais été étudiée dans cet écosystème. La

majorité des organismes détectés s'est révélée être proche d'organismes précédemment identifiés dans d'autres DMA (Bacillariophytes, Rhizaires, Alveolates, Amibes, Straménopiles et Champignons) soulignant ainsi la probable adaptation de ces organismes à des environnements extrêmes. Le pyroséquençage a également permis de mettre en évidence de nouveaux taxons de faible abondance non détectés auparavant dans ce type de milieu tels que les Apusozoaires, les Centrohéliozoaires et les Jakobides. Certains de ces organismes pourraient jouer un rôle écologique potentiellement important dans ce milieu. Ainsi, les eucaryotes photosynthétiques appartenant majoritairement aux Chlorophytes (algues verte), aux Streptophytes et à certaines classes autotrophes des Straménopiles (Bacillariophyceae, Chrysophyceae, Eustigmatophyceae et Synurophyceae) seraient les principaux producteurs primaires de cette communauté et pourraient favoriser l'oxydation aérobie du fer et du soufre par la production d'oxygène. Les champignons, majoritairement présents sur l'ensemble du DMA de Carnoulès (représentant 60 % des taxons) pourraient agir en tant que décomposeurs et principaux producteurs de substances inorganiques. De plus, un grand nombre de champignons sont capables d'adsorber des métaux dans leur cellule ou au niveau de formations extracellulaires et pourraient ainsi contribuer à diminuer le niveau de pollution, ce qui permettrait le développement éventuel d'espèces moins tolérantes (Amaral Zettler *et al.*, 2003). La communauté mixotrophe de cet écosystème semble dominée par des Rhizaires de la classe des Cercozoaires et par certains Straménopiles tandis que la communauté des prédateurs serait caractérisée par des protozoaires appartenant à divers groupes (Dinoflagellés, Choanoflagellés, Ciliés, Apicomplexes, Amibes, Apusozoaires et Centrohéliozoaires). Ces organismes sont par ailleurs susceptibles de se nourrir des bactéries acidophiles et pourraient donc jouer un rôle indirect dans les processus de génération et/ou remédiation des DMA par le contrôle de la composition des populations bactériennes et de leur biomasse. Cette étude a également permis de révéler une importante structuration spatiale de ces communautés liée, au moins partiellement, à la physicochimie de l'eau avec l'arsénite, le fer et le potentiel redox comme facteurs significativement responsables de l'assemblage des communautés eucaryotes à Carnoulès. Ces résultats ont contribué à approfondir nos connaissances sur le DMA de Carnoulès en caractérisant une communauté contribuant de façon probablement non négligeable au fonctionnement de cet écosystème. Ces travaux feront l'objet d'une publication prochaine dans *Environmental Microbiology*.

Dans la dernière partie de cette thèse, nous nous sommes employés à décrire plus spécifiquement la dynamique spatiotemporelle des communautés bactériennes présentes au niveau d'habitats fortement contrastés par leur niveau de pollution (*i.e.* stations localisées le long du gradient de pollution du Reigous) afin de déterminer des profils de diversité et de tenter de les corrélérer avec les caractéristiques physicochimiques du milieu. Les analyses réalisées dans cette étude alliant T-RFLP et pyroséquençage ont mis en évidence une faible diversité bactérienne. La majorité des phyla identifiés appartenait à des organismes communément retrouvés dans les DMA (Proteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes, Firmicutes et Nitrospira). Cependant l'utilisation du pyroséquençage a également permis d'atteindre une plus grande profondeur dans l'analyse de la diversité bactérienne présente sur le site de Carnoulès par la détection de nouveaux taxons représentant moins de 1 % du nombre total des séquences obtenues. En accord avec les résultats obtenus pour les eucaryotes, une structuration spatiale des communautés bactériennes a été mise en évidence et semblerait être corrélée avec les concentrations en arsenic, la température et le potentiel redox. L'ensemble des données de cette étude n'a malheureusement pas pu être exploité en détails, comme la dynamique temporelle des communautés, et des études complémentaires sont encore nécessaires afin d'émettre des hypothèses plus précises caractérisant l'impact de ces facteurs environnementaux sur les communautés bactériennes.

Dans leur ensemble, ces résultats nous ont apporté un gain considérable d'informations sur le DMA de Carnoulès et sur les communautés microbiennes qui le composent. Cependant, les travaux réalisés au cours de cette thèse ne représentent qu'une première étape dans la compréhension complète du fonctionnement de cet écosystème. En effet, de nombreuses questions restent en suspens et nécessiteraient d'être approfondies par des études ultérieures.

Perspectives

Plusieurs approches complémentaires aux travaux déjà réalisés pourraient être envisagées afin d'améliorer nos connaissances sur le fonctionnement du DMA de Carnoulès.

Tout d'abord concernant la communauté eucaryote, nous avons mis en évidence un biais relatif à l'utilisation d'une seule paire d'amorces qui ne permettait pas de révéler la diversité globale présente dans cet écosystème comme indiqué par l'absence de détection d'*Euglena mutabilis* dans nos analyses. Une meilleure évaluation de la diversité eucaryote pourrait alors être envisagée par l'utilisation de plusieurs jeux d'amorces comme mentionné par Stoeck *et al.* (2006).

D'autre part, après la réalisation de ces inventaires taxonomiques, il serait nécessaire d'établir un lien plus détaillé entre la composition et la fonction des communautés microbiennes identifiées. En effet, bien que les méthodes utilisées dans cette thèse aient permis de caractériser les communautés microbiennes d'un point de vue phylogénétique, elles apportent généralement peu de renseignements sur le rôle de ces microorganismes dans l'écosystème. Dans ce contexte, il existe plusieurs alternatives permettant de lier la structure phylogénétique à la structure fonctionnelle des communautés microbiennes mais nous ne nous intéresserons qu'aux techniques basées sur la biologie moléculaire. La première approche consiste à étudier les gènes ribosomiaux et en parallèle un ou plusieurs gènes fonctionnels d'intérêt. Les puces à ADN reposent sur ce principe et sont plus exhaustives, permettant le criblage d'un grand nombre de gènes fonctionnels mais ne permettant cependant de suivre que des gènes préalablement connus. Il est également possible que les gènes ciblés ne représentent qu'une fraction minoritaire des gènes fonctionnels présents dans l'échantillon (Cardenas & Tiedje, 2008). Cependant le développement des techniques de séquençage à haut débit ont permis de contourner cette limite par le criblage de l'ensemble des gènes présents dans une communauté. Ainsi en comparant et en réassemblant les séquences obtenues par rapport à des génomes de références, il est possible (i) de caractériser la structure phylogénétique des communautés microbiennes (gènes ribosomiaux), (ii) d'identifier les gènes fonctionnels par homologie aux génomes de référence, (iii) d'identifier de nouvelles fonctions et (iv) de reconstruire des génomes d'organismes non cultivables. L'utilisation complémentaire de la

métatranscriptomique permet d'identifier les ARNm présents dans l'échantillon et donne donc accès au niveau d'expression des gènes fonctionnels (Cardenas & Tiedje, 2008).

Ce type d'approche a été précédemment utilisé par Bertin *et al.* (2011) qui ont réalisé une reconstruction quasi complète de 7 génomes bactériens « majoritaires » présents sur le site de Carnoulès à la station COWG. L'analyse subséquente de ces 7 génomes, associée aux données de métaprotéomique issues de ces mêmes échantillons a permis l'élaboration d'un modèle de fonctionnement de la communauté microbienne. Cependant, cette analyse a été menée sur une seule station du site de Carnoulès or, comme nous l'avons vu précédemment, le Reigous est caractérisé par une succession d'écosystèmes présentant des communautés microbiennes distinctes. Il pourrait donc être intéressant de mettre en place ce type d'analyse sur des stations différentes afin de vérifier si les mêmes fonctions sont présentes sur l'ensemble de l'écosystème.

Enfin, l'ensemble des observations réalisées au cours des ces travaux a souligné l'importance du maintien des méthodes traditionnelles de culture et d'isolement pour l'identification précise et la compréhension du rôle fonctionnel des microorganismes dans le DMA de Carnoulès. En effet, bien que les techniques moléculaires permettent d'accéder aux organismes non cultivables, elles présentent souvent l'inconvénient de produire des inventaires composés pour une large part d'organismes inconnus, empêchant ainsi toute hypothèse sur leur rôle fonctionnel. De plus, les méthodes culturales permettent de contourner les biais inhérents aux approches moléculaires comme la résistance de certaines bactéries à la lyse cellulaire ou bien la difficulté à détecter les microorganismes appartenant à la biosphère rare sans faire un séquençage massif.

Références bibliographiques

A

- Adams, J. B., Palmer, F. and Staley, J. T. (1992).** Rock weathering in deserts: Mobilization and concentration of ferric iron by microorganisms. *Geomicrobiol J* **10**: 99-114.
- Aebersold, R. and Mann, M. (2003).** Mass spectrometry-based proteomics. *Nature* **422**: 198-207.
- Aguilera, A., Manrubia, S. C., Gómez, F., Rodríguez, N. and Amils, R. (2006).** Eukaryotic community distribution and its relationship to water physicochemical parameters in an extreme acidic environment, Río Tinto (Southwestern Spain). *Appl Environ Microbiol* **72**: 5325-5330.
- Aguilera, A., Zettler, E., Gómez, F., Amaral-Zettler, L., Rodríguez, N. and Amils, R. (2007).** Distribution and seasonal variability in the benthic eukaryotic community of Río Tinto (SW, Spain), an acidic, high metal extreme environment. *Syst Appl Microbiol* **30**: 531-546.
- Akcil, A. and Koldas, S. (2006).** Acid mine drainage (AMD): causes, treatment and case studies. *J Clean Prod* **14**: 1139-1145.
- Amann, R. I., Ludwig, W. and Schleifer, K. H. (1995).** Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143-169.
- Amaral-Zettler, L. A., Gomez, F., Zettler, E., Keenan, B. G., Amils, R. and Sogin, M. L. (2002).** Microbiology: Eukaryotic diversity in Spain's river of fire. *Nature* **417**: 137-137.
- Amaral-Zettler, L. A., Messerli, M. A., Laatsch, A. D., Smith, P. J. and Sogin, M. L. (2003).** From genes to genomes: beyond biodiversity in Spain's Río Tinto. *Biol Bull* **204**: 205-209.
- Amils, R., González-Toril, E., Fernández-Remolar, D., Gómez, F., Aguilera, Á., Rodríguez, N. et al. (2007).** Extreme environments as Mars terrestrial analogs: The Río Tinto case. *Planet Space Sci* **55**: 370-381.
- Anderson, C. R. and Cook, G. M. (2004).** Isolation and characterization of arsenate-reducing bacteria from arsenic-contaminated sites in New Zealand. *Curr Microbiol* **48**: 341-347.
- Anderson, L. and Seilhamer, J. (1997).** A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18**: 533-537.
- Avery, S. V., Codd, G. A. and Gadd, G. M. (1992).** Caesium transport in the cyanobacterium *Anabaena variabilis*: Kinetics and evidence for uptake via ammonium transport system(s). *FEMS Microbiol Lett* **95**: 253-258.

B

Bailey, N. J., Oven, M., Holmes, E., Nicholson, J. K. and Zenk, M. H. (2003). Metabolomic analysis of the consequences of cadmium exposure in *Silene cucubalus* cell cultures via ¹H NMR spectroscopy and chemometrics. *Phytochemistry* **62**: 851-858.

Baker, B. J. and Banfield, J. F. (2003). Microbial communities in acid mine drainage. *FEMS Microbiol Ecol* **44**: 139-152.

Baker, B. J., Lutz, M. A., Dawson, S. C., Bond, P. L. and Banfield, J. F. (2004). Metabolically active eukaryotic communities in extremely acidic mine drainage. *Appl Environ Microbiol* **70**: 6264-6271.

Baker, B. J., Tyson, G. W., Goosherst, L. and 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., Verberkmoes, N. C., Hettich, R. L. and Thelen, M. P. (2005). Proteogenomic approaches for the molecular characterization of natural microbial communities. *Omic* **9**: 301-333.

Bastida, F., Moreno, J. L., Nicolás, C., Hernández, T. and García, C. (2009). Soil metaproteomics: a review of an emerging environmental science. Significance, methodology and perspectives. *Eur J Soil Sci* **60**: 845-859.

Battaglia-Brunet, F., Joulain, C., Garrido, F., Dictor, M. C., Morin, D., Coupland, K. et al. (2006). Oxidation of arsenite by *Thiomonas* strains and characterization of *Thiomonas arsenivorans* sp. nov. *Antonie Van Leeuwenhoek* **89**: 99-108.

Bertin, P. N., Heinrich-Salmeron, A., Pelletier, E., Goulhen-Chollet, F., Arsène-Ploetze, F., Gallien, S. et al. (2011). Metabolic diversity among main microorganisms inside an arsenic-rich ecosystem revealed by meta- and proteo-genomics. *ISME J* **5**: 1735-1747.

Bini, E. (2010). Archaeal transformation of metals in the environment. *FEMS Microbiol Ecol* **73**: 1-16.

Birch, L. and Bachofen, R. (1990). Complexing agents from microorganisms. *Experientia* **46**: 8.

Bomar, L., Maltz, M., Colston, S. and Graf, J. (2011). Directed culturing of microorganisms using metatranscriptomics. *MBio* **2**: e00012-00011.

Bond, P. L., Smriga, S. P. and Banfield, J. F. (2000). Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl Environ Microbiol* **66**: 3842-3849.

Bossemeyer, D., Schlösser, A. and Bakker, E. P. (1989). Specific cesium transport via the *Escherichia coli* Kup (TrkD) K⁺ uptake system. *J Bacteriol* **171**: 2219-2221.

Botes, E., Van Heerden, E. and Litthauer, D. (2007). Hyper-resistance to arsenic in bacteria isolated from an antimony mine in South Africa. *S Afr J Sci* **103**: 279-281.

Brake, S. S., Dannelly, H. K. and Connors, K. A. (2001). Controls on the nature and distribution of an alga in coal mine-waste environments and its potential impact on water quality. *Environ Geol* **40**: 458-469.

Brake, S. S., Hasiotis, S. T., Dannelly, H. K. and Connors, K. A. (2002). Eukaryotic stromatolite builders in acid mine drainage: Implications for Precambrian iron formations and oxygenation of the atmosphere? *Geology* **30**: 599-602.

Bridge, T. A. M. and Johnson, D. B. (1998). Reduction of soluble iron and reductive dissolution of ferric iron-containing minerals by moderately thermophilic iron-oxidizing bacteria. *Appl Environ Microbiol* **64**: 2181-2186.

Brochier-Armanet, C., Boussau, B., Gribaldo, S. and Forterre, P. (2008). Mesophilic *Crenarchaeota*: proposal for a third archaeal phylum, the *Thaumarchaeota*. *Nat Rev Microbiol* **6**: 245-252.

Brochier-Armanet, C., Forterre, P. and Gribaldo, S. (2011). Phylogeny and evolution of the *Archaea*: one hundred genomes later. *Curr Opin Microbiol* **14**: 274-281.

Bruneel, O., Personné, J. C., Casiot, C., Leblanc, M., Elbaz-Poulichet, F., Mahler, B. J. et al. (2003). Mediation of arsenic oxidation by *Thiomonas* sp. in acid-mine drainage (Carnoulès, France). *J Appl Microbiol* **95**: 492-499.

Bruneel, O., Duran, R., Koffi, K., Casiot, C., Fourçans, A., Elbaz-Poulichet, F. and Personné, J. C. (2005). Microbial diversity in a pyrite-rich tailings impoundment (Carnoulès, France). *Geomicrobiol J* **22**: 249-257.

Bruneel, O., Duran, R., Casiot, C., Elbaz-Poulichet, F. and 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., Goni-Urriza, M. S., Elbaz-Poulichet, F. et al. (2008). Archaeal diversity in a Fe-As rich acid mine drainage at Carnoulès (France). *Extremophiles* **12**: 563-571.

Bruneel, O., Volant, A., Gallien, S., Chaumande, B., Casiot, C., Carapito, C. et al. (2011). Characterization of the active bacterial community involved in natural attenuation processes in arsenic-rich creek sediments. *Microb Ecol* **61**: 793-810.

C

Campos Paulino, L., Faria Bergamo, R., de Mello, M., Garcia, O., Manfio, G. and Ottoni, L. (2001). Molecular characterization of *Acidithiobacillus ferrooxidans* and *A. thiooxidans* strains isolated from mine wastes in Brazil. *Antonie van Leeuwenhoek* **80**: 65-75.

Canfield, D. E. and Des Marais, D. J. (1991). Aerobic sulfate reduction in microbial mats. *Science* **251**: 1471-1473.

Cardenas, E. and Tiedje, J. M. (2008). New tools for discovering and characterizing microbial diversity. *Curr Opin Biotechnol* **19**: 544-549.

Caron, D. A., Countway, P. D. and Brown, M. V. (2004). The growing contributions of molecular biology and immunology to protistan ecology: molecular signatures as ecological tools. *J Eukaryot Microbiol* **51**: 38-48.

Cascante, M. and Marin, S. (2008). Metabolomics and fluxomics approaches. *Essays Biochem* **45**: 67-81.

Casiot, C., Leblanc, M., Bruneel, O., Personné, J. C., Koffi, K. and Elbaz-Poulichet, F. (2003a). Geochemical processes controlling the formation of As-rich waters within a tailings impoundment (Carnoulès, France). *Aquat Geochem* **9**: 273-290.

Casiot, C., Morin, G., Juillot, F., Bruneel, O., Personné, J. C., Leblanc, M. et al. (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. and Elbaz-Poulichet, F. (2004). Arsenic oxidation and bioaccumulation by the acidophilic protozoan, *Euglena mutabilis*, in acid mine drainage (Carnoulès, France). *Sci Total Environ* **320**: 259-267.

Casiot, C., Lebrun, S., Morin, G., Bruneel, O., Personné, J. C. and Elbaz-Poulichet, F. (2005). Sorption and redox processes controlling arsenic fate and transport in a stream impacted by acid mine drainage. *Sci Total Environ* **347**: 122-130.

Casiot, C., Egal, M., Elbaz-Poulichet, F., Bruneel, O., Bancon-Montigny, C., Cordier, M.-A. et al. (2009). Hydrological and geochemical control of metals and arsenic in a Mediterranean river contaminated by acid mine drainage (the Amous River, France); preliminary assessment of impacts on fish (*Leuciscus cephalus*). *Appl Geochem* **24**: 787-799.

Casneuf, T., Van de Peer, Y. and Huber, W. (2007). In situ analysis of cross-hybridisation on microarrays and the inference of expression correlation. *BMC Bioinformatics* **8**: 461.

Cebron, A. and Garnier, J. (2005). *Nitrobacter* and *Nitrospira* genera as representatives of nitrite-oxidizing bacteria: detection, quantification and growth along the lower Seine River (France). *Water Res* **39**: 4979-4992.

Chaban, B., Ng, S. Y. and Jarrell, K. F. (2006). Archaeal habitats — from the extreme to the ordinary. *Can J Microbiol* **52**: 73-116.

Chait, B. T. (2006). Chemistry. Mass spectrometry: bottom-up or top-down? *Science* **314**: 65-66.

Chen, W., Parette, R., Zou, J., Cannon, F. S. and Dempsey, B. A. (2007). Arsenic removal by iron-modified activated carbon. *Water Res* **41**: 1851-1858.

Cloonan, N., Forrest, A. R., Kollé, G., Gardiner, B. B., Faulkner, G. J., Brown, M. K. et al. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods* **5**: 613-619.

Cohan, F. M. and Koeppe, A. F. (2008). The origins of ecological diversity in prokaryotes. *Curr Biol* **18**: R1024-1034.

Conrad, R. (1996). Soil microorganisms as controllers of atmospheric trace gases (H₂, CO, CH₄, OCS, N₂O, and NO). *Microbiol Rev* **60**: 609-640.

Cravatt, B. F., Simon, G. M. and Yates, J. R., 3rd (2007). The biological impact of mass-spectrometry-based proteomics. *Nature* **450**: 991-1000.

D

Das, B. K., Roy, A., Koschorreck, M., Mandal, S. M., Wendt-Potthoff, K. and Bhattacharya, J. (2009). Occurrence and role of algae and fungi in acid mine drainage environment with special reference to metals and sulfate immobilization. *Water Res* **43**: 883-894.

Delavat, F., Phalip, V., Forster, A., Lett, M. C. and Lievremont, D. (2012). Deciphering the role of *Paenibacillus* strain Q8 in the organic matter recycling in the acid mine drainage of Carnoulès. *Microb Cell Fact* **11**: 16.

DeLong, E. F. (2009). The microbial ocean from genomes to biomes. *Nature* **459**: 200-206.

Dettmer, K., Aronov, P. A. and Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrom Rev* **26**: 51-78.

Devars, S., Hernandez, R. and Moreno-Sanchez, R. (1998). Enhanced heavy metal tolerance in two strains of photosynthetic *Euglena gracilis* by preexposure to mercury or cadmium. *Arch Environ Contam Toxicol* **34**: 128-135.

Dilling, W. and Cypionka, H. (1990). Aerobic respiration in sulfate-reducing bacteria. *FEMS Microbiol Lett* **71**: 123-127.

Dohm, J. C., Lottaz, C., Borodina, T. and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res* **36**: e105.

Domon, B. and Aebersold, R. (2006). Mass spectrometry and protein analysis. *Science* **312**: 212-217.

Dugat-Bony, E., Peyretailade, E., Parisot, N., Biderre-Petit, C., Jaziri, F., Hill, D. et al. (2012). Detecting unknown sequences with DNA microarrays: explorative probe design strategies. *Environ Microbiol* **14**: 356-371.

Duquesne, K., Lebrun, S., Casiot, C., Bruneel, O., Personné, J. C., Leblanc, M. et al. (2003). Immobilization of arsenite and ferric iron by *Acidithiobacillus ferrooxidans* and its relevance to acid mine drainage. *Appl Environ Microbiol* **69**: 6165-6173.

Dworkin, M. (2012). Sergei Winogradsky: a founder of modern microbiology and the first microbial ecologist. *FEMS Microbiol Rev* **36**: 364-379.

E

Edwards, K. J., Gihring, T. M. and Banfield, J. F. (1999). Seasonal variations in microbial populations and environmental conditions in an extreme acid mine drainage environment. *Appl Environ Microbiol* **65**: 3627-3632.

Edwards, K. J., Bond, P. L., Druschel, G. K., McGuire, M. M., Hamers, R. J. and Banfield, J. F. (2000a). Geochemical and biological aspects of sulfide mineral dissolution: lessons from Iron Mountain, California. *Chem Geol* **169**: 383-397.

Edwards, K. J., Bond, P. L., Gihring, T. M. and Banfield, J. F. (2000b). An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science* **287**: 1796-1799.

Edwards, K. J., Hu, B., Hamers, R. J. and Banfield, J. F. (2001). A new look at microbial leaching patterns on sulfide minerals. *FEMS Microbiol Ecol* **34**: 197-206.

Egal, M., Casiot, C., Morin, G., Elbaz-Poulichet, F., Cordier, M.-A. and Bruneel, O. (2010). An updated insight into the natural attenuation of As concentrations in Reigous Creek (southern France). *Appl Geochem* **25**: 1949-1957.

Ellis, P. J., Conrads, T., Hille, R. and 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.

Erlich, Y., Mitra, P. P., delaBastide, M., McCombie, W. R. and Hannon, G. J. (2008). Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nat Methods* **5**: 679-682.

Evangelou, V. P. (1995). Pyrite oxidation and its control. Boca Raton: CRC press.

F

Falkowski, P. G., Fenchel, T. and Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034-1039.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. (1989). Electrospay ionization for mass spectrometry of large biomolecules. *Science* **246**: 64-71.

Ferris, F. G., Hallberg, R. O., Lyvén, B. and Pedersen, K. (2000). Retention of strontium, cesium, lead and uranium by bacterial iron oxides from a subterranean environment. *Appl Geochem* **15**: 1035-1042.

Fiehn, O. (2001). Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp Funct Genomics* **2**: 155-168.

Fiehn, O. (2002). Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol* **48**: 155-171.

Fowler, T. A., Holmes, P. R. and Crundwell, F. K. (1999). Mechanism of pyrite dissolution in the presence of *Thiobacillus ferrooxidans*. *Appl Environ Microbiol* **65**: 2987-2993.

Franche, C., Lindström, K. and Elmerich, C. (2009). Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant Soil* **321**: 35-59.

Frias-Lopez, J., Shi, Y., Tyson, G. W., Coleman, M. L., Schuster, S. C., Chisholm, S. W. and Delong, E. F. (2008). Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci U S A* **105**: 3805-3810.

Friedrich, C. G., Rother, D., Bardischewsky, F., Quentmeier, A. and Fischer, J. (2001). Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism? *Appl Environ Microbiol* **67**: 2873-2882.

Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A. and Fischer, J. (2005). Prokaryotic sulfur oxidation. *Curr Opin Microbiol* **8**: 253-259.

Fuchs, T., Huber, H., Teiner, K., Burggraf, S. and Stetter, K. O. (1995). *Metallosphaera prunae*, sp. nov., a novel metal-mobilizing, thermoacidophilic *Archaeum*, isolated from a uranium mine in Germany. *Syst Appl Microbiol* **18**: 560-566.

Gadanho, M., Libkind, D. and Sampaio, J. P. (2006). Yeast diversity in the extreme acidic environments of the Iberian Pyrite Belt. *Microb Ecol* **52**: 552-563.

G

Gadd, G. M. (2004). Microbial influence on metal mobility and application for bioremediation. *Geoderma* **122**: 109-119.

Gadd, G. M. (2010). Metals, minerals and microbes: geomicrobiology and bioremediation. *Microbiology* **156**: 609-643.

Galand, P. E., Casamayor, E. O., Kirchman, D. L. and Lovejoy, C. (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci U S A* **106**: 22427-22432.

García, C., Moreno, D. A., Ballester, A., Blazquez, M. L. and Gonzalez, F. (2001). Bioremediation of an industrial acid mine water by metal-tolerant sulphate-reducing bacteria. *Miner Eng* **14**: 997-1008.

Gebel, T. (1997). Arsenic and antimony: comparative approach on mechanistic toxicology. *Chem Biol Interact* **107**: 131-144.

Gehrke, T., Telegdi, J., Thierry, D. and Sand, W. (1998). Importance of extracellular polymeric substances from *Thiobacillus ferrooxidans* for bioleaching. *Appl Environ Microbiol* **64**: 2743-2747.

Gihring, T. M., Bond, P. L., Peters, S. C. and Banfield, J. F. (2003). Arsenic resistance in the archaeon "*Ferroplasma acidarmanus*": new insights into the structure and evolution of the ars genes. *Extremophiles* **7**: 123-130.

Gilbert, J. A., Field, D., Huang, Y., Edwards, R., Li, W., Gilna, P. and Joint, I. (2008). Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. *PLoS One* **3**: e3042.

Gill, S. R., Pop, M., Deboy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S. et al. (2006). Metagenomic analysis of the human distal gut microbiome. *Science* **312**: 1355-1359.

Golyshina, O. V., Pivovarova, T. A., Karavaiko, G. I., Kondrat'eva, T. F., Moore, E. R., Abraham, W. R. et al. (2000). *Ferroplasma acidiphilum* gen. nov., sp. nov., an acidophilic, autotrophic, ferrous-iron-oxidizing, cell-wall-lacking, mesophilic member of the *Ferroplasmaceae* fam. nov., comprising a distinct lineage of the Archaea. *Int J Syst Evol Microbiol* **50**: 997-1006.

Golyshina, O. V. and Timmis, K. N. (2005). *Ferroplasma* and relatives, recently discovered cell wall-lacking archaea making a living in extremely acid, heavy metal-rich environments. *Environ Microbiol* **7**: 1277-1288.

Graham, R., Graham, C. and McMullan, G. (2007). Microbial proteomics: a mass spectrometry primer for biologists. *Microb Cell Fact* **6**: 26.

Green, H. H. (1918). Description of a bacterium which oxidizes arsenite to arsenate, and one which reduces arsenate to arsenite, isolated from a cattle-dipping tank. *S Afr J Sci* **14**: 465–467.

Green, J. L., Bohannon, B. J. and Whitaker, R. J. (2008). Microbial biogeography: from taxonomy to traits. *Science* **320**: 1039-1043.

Guo, X., You, X.-Y., Liu, L.-J., Zhang, J.-Y., Liu, S.-J. and Jiang, C.-Y. (2009). *Alicyclobacillus aeris* sp. nov., a novel ferrous- and sulfur-oxidizing bacterium isolated from a copper mine. *Int J Syst Evol Microbiol* **59**: 2415-2420.

Gyaneshwar, P., Naresh Kumar, G., Parekh, L. J. and Poole, P. S. (2002). Role of soil microorganisms in improving P nutrition of plants. *Plant Soil* **245**: 83-93.

H

Haack, E. and Warren, L. A. (2003). Biofilm hydrous manganese oxyhydroxides and metal dynamics in acid rock drainage. *Environ Sci Technol* **37**: 4138-4147.

Haferburg, G. and Kothe, E. (2007). Microbes and metals: interactions in the environment. *J Basic Microbiol* **47**: 453-467.

Hall, C., Brachat, S. and Dietrich, F. S. (2005). Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. *Eukaryot Cell* **4**: 1102-1115.

Hallbeck, L., Stahl, F. and Pedersen, K. (1993). Phylogeny and phenotypic characterization of the stalk-forming and iron-oxidizing bacterium *Gallionella ferruginea*. *J Gen Microbiol* **139**: 1531-1535.

RÉFÉRENCES BIBLIOGRAPHIQUES

Hallberg, K., González-Toril, E. and Johnson, D. (2010). *Acidithiobacillus ferrivorans*, sp. nov.; facultatively anaerobic, psychrotolerant iron-, and sulfur-oxidizing acidophiles isolated from metal mine-impacted environments. *Extremophiles* **14**: 9-19.

Hallberg, K. B. and Johnson, D. B. (2001). Biodiversity of acidophilic prokaryotes. *Adv Appl Microbiol Volume* **49**: 37-84.

Hallberg, K. B. and 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.

Hallberg, K. B., Coupland, K., Kimura, S. and Johnson, D. B. (2006). Macroscopic streamer growths in acidic, metal-rich mine waters in North Wales consist of novel and remarkably simple bacterial communities. *Appl Environ Microbiol* **72**: 2022-2030.

Hallberg, K. B. (2010). New perspectives in acid mine drainage microbiology. *Hydrometallurgy* **104**: 448-453.

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

Halter, D., Goulhen-Chollet, F., Gallien, S., Casiot, C., Hamelin, J., Gilard, F. et al. (2012a). In situ proteo-metabolomics reveals metabolite secretion by the acid mine drainage bio-indicator, *Euglena mutabilis*. *ISME J* **6**: 1391-1402.

Halter, D., Casiot, C., Heipieper, H. J., Plewniak, F., Marchal, M., Simon, S. et al. (2012b). Surface properties and intracellular speciation revealed an original adaptive mechanism to arsenic in the acid mine drainage bio-indicator *Euglena mutabilis*. *Appl Microbiol Biotechnol* **93**: 1735-1744.

Haqqani, A. S., Kelly, J. F. and Stanimirovic, D. B. (2008). Quantitative protein profiling by mass spectrometry using label-free proteomics. *Methods Mol Biol* **439**: 241-256.

He, Z., Xie, X., Xiao, S., Liu, J. and Qiu, G. (2007). Microbial diversity of mine water at Zhong Tiaoshan copper mine, China. *J Basic Microbiol* **47**: 485-495.

Heinrich-Salmeron, A., Cordi, A., Brochier-Armanet, C., Halter, D., Pagnout, C., Abbaszadeh-fard, E. et al. (2011). Unsuspected diversity of arsenite-oxidizing bacteria revealed by a widespread distribution of the *aoxB* gene in prokaryotes. *Appl Environ Microbiol* **77**: 4685-4692.

Heinzel, E., Hedrich, S., Janneck, E., Glombitza, F., Seifert, J. and Schlömann, M. (2009). Bacterial diversity in a mine water treatment plant. *Appl Environ Microbiol* **75**: 858-861.

Hillenkamp, F., Karas, M., Beavis, R. C. and Chait, B. T. (1991). Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal Chem* **63**: 1193A-1203A.

Hiraishi, A., Matsuzawa, Y., Kanbe, T. and Wakao, N. (2000). *Acidisphaera rubrifaciens* gen. nov., sp. nov., an aerobic bacteriochlorophyll-containing bacterium isolated from acidic environments. *Int J Syst Evol Microbiol* **50**: 1539-1546.

Huber, R., Sacher, M., Vollmann, A., Huber, H. and Rose, D. (2000). Respiration of arsenate and selenate by hyperthermophilic archaea. *Syst Appl Microbiol* **23**: 305-314.

Hudson-Edwards, K. A., Schell, C. and Macklin, M. G. (1999). Mineralogy and geochemistry of alluvium contaminated by metal mining in the Río Tinto area, southwest Spain. *Appl Geochem* **14**: 1015-1030.

Huse, S. M., Huber, J. A., Morrison, H. G., Sogin, M. L. and Welch, D. M. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* **8**: R143.

J

Jareonmit, P., Mehta, M., Sadowsky, M. and Sajjaphan, K. (2012). Phylogenetic and phenotypic analyses of arsenic-reducing bacteria isolated from an old tin mine area in Thailand. *World J Microbiol Biotechnol* **28**: 2287-2292.

Johnson, D. B. and Rang, L. (1993). Effects of acidophilic protozoa on populations of metal-mobilizing bacteria during the leaching of pyritic coal. *J Gen Microbiol* **139**: 1417-1423.

Johnson, D. B. (1995). Acidophilic microbial communities: candidates for bioremediation of acidic mine effluents. *Int Biodeterior Biodegradation* **35**: 41-58.

Johnson, D. B. (1998). Biodiversity and ecology of acidophilic microorganisms. *FEMS Microbiol Ecol* **27**: 307-317.

Johnson, D. B., Rolfe, S., Hallberg, K. B. and Iversen, E. (2001). Isolation and phylogenetic characterization of acidophilic microorganisms indigenous to acidic drainage waters at an abandoned Norwegian copper mine. *Environ Microbiol* **3**: 630-637.

Johnson, D. B. and Bridge, T. A. M. (2002). Reduction of ferric iron by acidophilic heterotrophic bacteria: evidence for constitutive and inducible enzyme systems in *Acidiphilium* spp. *J Appl Microbiol* **92**: 315-321.

Johnson, D. B., Dziurla, M. A., Kolmert, A. and Hallberg, K. B. (2002). The microbiology of acid mine drainage: genesis and biotreatment. *S Afr J Sci* **98**: 249-255.

Johnson, D. B. and Hallberg, K. B. (2003). The microbiology of acidic mine waters. *Res Microbiol* **154**: 466-473.

Johnson, D. B. and Hallberg, K. B. (2005). Acid mine drainage remediation options: a review. *Sci Total Environ* **338**: 3-14.

Johnson, D. B., Bacelar-Nicolau, P., Okibe, N., Thomas, A. and Hallberg, K. B. (2009). *Ferrimicrobium acidiphilum* gen. nov., sp. nov. and *Ferrithrix thermotolerans* gen. nov., sp.

RÉFÉRENCES BIBLIOGRAPHIQUES

nov.: heterotrophic, iron-oxidizing, extremely acidophilic actinobacteria. *Int J Syst Evol Microbiol* **59**: 1082-1089.

Johnson, D. B. (2012). Geomicrobiology of extremely acidic subsurface environments. *FEMS Microbiol Ecol* **81**: 2-12.

Jong, T. and Parry, D. L. (2003). Removal of sulfate and heavy metals by sulfate reducing bacteria in short-term bench scale upflow anaerobic packed bed reactor runs. *Water Res* **37**: 3379-3389.

Joyce, A. R. and Palsson, B. O. (2006). The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Biol* **7**: 198-210.

Juniper, S. K. and Tebo, B. M. (1995). Microbe-metal interactions and mineral deposition at hydrothermal vents. In *The microbiology of deep-sea hydrothermal vents*. (Karl Eds). Boca Raton: pp. 219-253.

K

Katsoyiannis, I. A. and Zouboulis, A. I. (2004). Application of biological processes for the removal of arsenic from groundwaters. *Water Res* **38**: 17-26.

Kell, D. B. (2004). Metabolomics and systems biology: making sense of the soup. *Curr Opin Microbiol* **7**: 296-307.

Keller, M. and Hettich, R. (2009). Environmental proteomics: a paradigm shift in characterizing microbial activities at the molecular level. *Microbiol Mol Biol Rev* **73**: 62-70.

Kinzler, K., Gehrke, T., Telegdi, J. and Sand, W. (2003). Biobleaching—a result of interfacial processes caused by extracellular polymeric substances (EPS). *Hydrometallurgy* **71**: 83-88.

Kishimoto, N., Kosako, Y. and Tano, T. (1991). *Acidobacterium capsulatum* gen. nov., sp. nov.: An acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. *Curr Microbiol* **22**: 1-7.

Kolmert, Å. and Johnson, D. B. (2001). Remediation of acidic waste waters using immobilised, acidophilic sulfate-reducing bacteria. *J Chem Technol Biotechnol* **76**: 836-843.

Korlach, J., Bjornson, K. P., Chaudhuri, B. P., Cicero, R. L., Flusberg, B. A., Gray, J. J. et al. (2010). Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol* **472**: 431-455.

Korte, N. E. and Fernando, Q. (1991). A review of arsenic (III) in groundwater. *Crit Rev Environ Control* **21**: 1-39.

Kulp, T. R., Hoefft, S. E., Asao, M., Madigan, M. T., Hollibaugh, J. T., Fisher, J. C. et al. (2008). Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science* **321**: 967-970.

Kupferberg, E. D. (2003). A field of great promise: soil bacteriology in America, 1900-1925. *Endeavour* **27**: 16-21.

Küsel, K., Dorsch, T., Acker, G. and Stackebrandt, E. (1999). Microbial reduction of Fe(III) in acidic sediments: isolation of *Acidiphilium cryptum* JF-5 capable of coupling the reduction of Fe(III) to the oxidation of glucose. *Appl Environ Microbiol* **65**: 3633-3640.

Küsel, K., Roth, U. and Drake, H. L. (2002). Microbial reduction of Fe(III) in the presence of oxygen under low pH conditions. *Environ Microbiol* **4**: 414-421.

L

Larsson, L., Olsson, G., Hoist, O. and Karlsson, H. T. (1993). Oxidation of pyrite by *Acidianus brierleyi*: Importance of close contact between the pyrite and the microorganisms. *Biotechnol Lett* **15**: 99-104.

Lear, G., Niyogi, D., Harding, J., Dong, Y. and Lewis, G. (2009). Biofilm bacterial community structure in streams affected by acid mine drainage. *Appl Environ Microbiol* **75**: 3455-3460.

Leblanc, M., Achard, B., Ben Othman, D., Luck, J. M., Bertrand-Sarfati, J. and Personné, J. C. (1996). Accumulation of arsenic from acidic mine waters by ferruginous bacterial accretions (stromatolites). *Appl Geochem* **11**: 541-554.

Lebrun, E., Brugna, M., Baymann, F., Muller, D., Lièvremon, D., Lett, M.-C. and Nitschke, W. (2003). Arsenite oxidase, an ancient bioenergetic enzyme. *Mol Biol Evol* **20**: 686-693.

Ledin, M. and Pedersen, K. (1996). The environmental impact of mine wastes — Roles of microorganisms and their significance in treatment of mine wastes. *Earth-Science Reviews* **41**: 67-108.

Ledoux, E. (1996). Données et concepts utiles pour la gestion optimale de l'environnement hydrogéologique, aspect méthodologique. In *Les Techniques*, numéro spécial "Mines et Carrières". pp. 11-13.

Lee, G., Bigham, J. M. and Faure, G. (2002). Removal of trace metals by coprecipitation with Fe, Al and Mn from natural waters contaminated with acid mine drainage in the Ducktown Mining District, Tennessee. *Appl Geochem* **17**: 569-581.

Lee, Y.-J., Romanek, C. S. and Wiegel, J. (2007). *Clostridium aciditolerans* sp. nov., an acid-tolerant spore-forming anaerobic bacterium from constructed wetland sediment. *Int J Syst Evol Microbiol* **57**: 311-315.

Lièvremon, D., Bertin, P. N. and Lett, M. C. (2009). Arsenic in contaminated waters: Biogeochemical cycle, microbial metabolism and biotreatment processes. *Biochimie* **91**: 1229-1237.

Lilley, K. S., Razzaq, A. and Dupree, P. (2002). Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. *Curr Opin Chem Biol* **6**: 46-50.

López-Archilla, A. I., Marin, I. and Amils, R. (2001). Microbial community composition and ecology of an acidic aquatic environment: The Tinto river, Spain. *Microb Ecol* **41**: 20-35.

López-García, P. and Moreira, D. (2008). Tracking microbial biodiversity through molecular and genomic ecology. *Res Microbiol* **159**: 67-73.

M

Macur, R. E., Wheeler, J. T., McDermott, T. R. and Inskeep, W. P. (2001). Microbial populations associated with the reduction and enhanced mobilization of arsenic in mine tailings. *Environ Sci Technol* **35**: 3676-3682.

Macy, J. M., Santini, J. M., Pauling, B. V., O'Neill, A. H. and Sly, L. I. (2000). Two new arsenate/sulfate-reducing bacteria: mechanisms of arsenate reduction. *Arch Microbiol* **173**: 49-57.

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet* **24**: 133-141.

Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M. and Gilad, Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* **18**: 1509-1517.

Maron, P. A., Ranjard, L., Mougél, C. and Lemanceau, P. (2007). Metaproteomics: a new approach for studying functional microbial ecology. *Microb Ecol* **53**: 486-493.

Mayr, M., Chung, Y. L., Mayr, U., Yin, X., Ly, L., Troy, H. et al. (2005). Proteomic and metabolomic analyses of atherosclerotic vessels from apolipoprotein E-deficient mice reveal alterations in inflammation, oxidative stress, and energy metabolism. *Arterioscler Thromb Vasc Biol* **25**: 2135-2142.

Mayr, M. (2008). Metabolomics: ready for the prime time? *Circ Cardiovasc Genet* **1**: 58-65.

Mc Guinness, S. and Johnson, D. B. (1992). Grazing of acidophilic bacteria by a flagellated protozoan. *Microb Ecol* **23**: 75-86.

Metzker, M. L. (2010). Sequencing technologies — the next generation. *Nat Rev Genet* **11**: 31-46.

Monchy, S., Sancier, G., Jobard, M., Rasconi, S., Gerphagnon, M., Chabe, M. et al. (2011). Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. *Environ Microbiol* **13**: 1433-1453.

Morin, G., Juillot, F., Casiot, C., Bruneel, O., Personné, J. C., Elbaz-Poulichet, F. et al. (2003). Bacterial formation of tooeleite and mixed arsenic(III) or arsenic(V)-iron(III) gels in

RÉFÉRENCES BIBLIOGRAPHIQUES

the Carnoulès acid mine drainage, France. A XANES, XRD, and SEM study. *Environ Sci Technol* **37**: 1705-1712.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**: 621-628.

Moses, C. O., Kirk Nordstrom, D., Herman, J. S. and Mills, A. L. (1987). Aqueous pyrite oxidation by dissolved oxygen and by ferric iron. *Geochim Cosmochim Acta* **51**: 1561-1571.

Muller, D., Médigue, C., Koechler, S., Barbe, V., Barakat, M., Talla, E. et al. (2007). A tale of two oxidation states: bacterial colonization of arsenic-rich environments. *PLoS Genet* **3**: e53.

Muyzer, G. and Stams, A. J. (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nat Rev Microbiol* **6**: 441-454.

N

Natarajan, K. A. (2008). Microbial aspects of acid mine drainage and its bioremediation. *Trans Nonferrous Met Soc China* **18**: 1352-1360.

Neculita, C. M., Zagury, G. J. and Bussiere, B. (2007). Passive treatment of acid mine drainage in bioreactors using sulfate-reducing bacteria: critical review and research needs. *J Environ Qual* **36**: 1-16.

Nicholson, J. K., Connelly, J., Lindon, J. C. and Holmes, E. (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discov* **1**: 153-161.

Nordstrom, D. K. and Southam, G. (1997). Geomicrobiology of sulfide mineral oxidation. *Rev Mineral Geochem* **35**: 361-390.

Nordstrom, D. K., Alpers, C. N., Ptacek, C. J. and Blowes, D. W. (1999). Negative pH and extremely acidic mine waters from Iron Mountain, California. *Environ Sci Technol* **34**: 254-258.

Normand, P. and Lalonde, M. (1986). The genetics of actinorhizal *Frankia*: A review. *Plant Soil* **90**: 429-453.

Norris, P. R., Marsh, R. M. and Lindström, E. B. (1986). Growth of mesophilic and thermophilic acidophilic bacteria on sulfur and tetrathionate. *Biotechnol Appl Biochem* **8**: 318-329.

Norris, P. R. (1990). Acidophilic bacteria and their activity in mineral sulfide oxidation. In *Microbial Mineral Recovery* (Ehrlich & Brierley Eds). New York: pp. 3-27.

O

O'Brien, S. J. and Goedert, J. J. (1996). HIV causes AIDS: Koch's postulates fulfilled. *Curr Opin Immunol* **8**: 613-618.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **250**: 4007-4021.

Ochman, H., Lawrence, J. G. and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299-304.

Ona-Nguema, G., Morin, G., Juillot, F., Calas, G. and Brown, G. E. (2005). EXAFS analysis of arsenite adsorption onto two-line ferrihydrite, hematite, goethite, and lepidocrocite. *Environ Sci Technol* **39**: 9147-9155.

Oremland, R. S. and Stolz, J. F. (2003). The ecology of arsenic. *Science* **300**: 939-944.

Oremland, R. S. and Stolz, J. F. (2005). Arsenic, microbes and contaminated aquifers. *Trends Microbiol* **13**: 45-49.

Ozsolak, F., Platt, A. R., Jones, D. R., Reifengerger, J. G., Sass, L. E., McInerney, P. et al. (2009). Direct RNA sequencing. *Nature* **461**: 814-818.

P

Parkes, R. J., Cragg, B. A., Bale, S. J., Getliff, J. M., Goodman, K., Rochelle, P. A. et al. (1994). Deep bacterial biosphere in Pacific Ocean sediments. *Nature* **371**: 410-413.

Pedersen, S., Bloch, P. L., Reeh, S. and Neidhardt, F. C. (1978). Patterns of protein synthesis in *E. coli*: a catalog of the amount of 140 individual proteins at different growth rates. *Cell* **14**: 179-190.

Plumb, J. J., Haddad, C. M., Gibson, J. A. and Franzmann, P. D. (2007). *Acidianus sulfidivorans* sp. nov., an extremely acidophilic, thermophilic archaeon isolated from a solfatara on Lihir Island, Papua New Guinea, and emendation of the genus description. *Int J Syst Evol Microbiol* **57**: 1418-1423.

Poretsky, R. S., Hewson, I., Sun, S., Allen, A. E., Zehr, J. P. and Moran, M. A. (2009). Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environ Microbiol* **11**: 1358-1375.

Porter, J. R. (1976). Antony van Leeuwenhoek: tercentenary of his discovery of bacteria. *Bacteriol Rev* **40**: 260-269.

Pradet-Balade, B., Boulme, F., Beug, H., Mullner, E. W. and Garcia-Sanz, J. A. (2001). Translation control: bridging the gap between genomics and proteomics? *Trends Biochem Sci* **26**: 225-229.

Prescott, L. M., Harley, J. P. and Klein, D. A. (2010). Microbiologie, 3ème édition française. (DeBoeck Eds). Bruxelles: pp.643-660.

Pronk, J. T., de Bruyn, J. C., Bos, P. and Kuenen, J. G. (1992). Anaerobic growth of *Thiobacillus ferrooxidans*. *Appl Environ Microbiol* **58**: 2227-2230.

Q

Qin, J., Lehr, C. R., Yuan, C., Le, X. C., McDermott, T. R. and Rosen, B. P. (2009). Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga. *Proc Natl Acad Sci U S A* **106**: 5213-5217.

Qiu, R., Zhao, B., Liu, J., Huang, X., Li, Q., Brewer, E. et al. (2009). Sulfate reduction and copper precipitation by a *Citrobacter* sp. isolated from a mining area. *J Hazard Mater* **164**: 1310-1315.

R

Raes, J. and Bork, P. (2008). Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol* **6**: 693-699.

Ram, R. J., VerBerkmoes, N. C., Thelen, M. P., Tyson, G. W., Baker, B. J., Blake, R. C. et al. (2005). Community Proteomics of a Natural Microbial Biofilm. *Science* **308**: 1915-1920.

Raymond, J., Siefert, J. L., Staples, C. R. and Blankenship, R. E. (2004). The natural history of nitrogen fixation. *Mol Biol Evol* **21**: 541-554.

Ritchie, A. I. M. (1994). The waste-rock environment. In *Short Course Handbook on Environmental Geochemistry of Sulfide Mine Waste*. (Jambor & Blowes Eds). Ottawa: pp.133-161.

Rodriguez-Valera, F., Ruiz-Berraquero, F. and Ramos-Cormenzana, A. (1981). Characteristics of the heterotrophic bacterial populations in hypersaline environments of different salt concentrations. *Microb Ecol* **7**: 235-243.

Rodriguez-Valera, F. (2004). Environmental genomics, the big picture? *FEMS Microbiol Lett* **231**: 153-158.

Röling, W. F. M., Ferrer, M. and Golyshin, P. N. (2010). Systems approaches to microbial communities and their functioning. *Curr Opin Biotechnol* **21**: 532-538.

Rowe, O. F., Sanchez-Espana, J., Hallberg, K. B. and Johnson, D. B. (2007). Microbial communities and geochemical dynamics in an extremely acidic, metal-rich stream at an abandoned sulfide mine (Huelva, Spain) underpinned by two functional primary production systems. *Environ Microbiol* **9**: 1761-1771.

Rowe, O. F. and Johnson, D. B. (2008). Comparison of ferric iron generation by different species of acidophilic bacteria immobilized in packed-bed reactors. *Syst Appl Microbiol* **31**: 68-77.

S

Sánchez-Andrea, I., Rodríguez, N., Amils, R. and Sanz, J. L. (2011). Microbial diversity in anaerobic sediments at Río Tinto, a naturally acidic environment with a high heavy metal content. *Appl Environ Microbiol* **77**: 6085-6093.

Sánchez España, J., Pamo, E. L., Pastor, E. S., Andrés, J. R. and Rubí, J. A. M. (2005). The natural attenuation of two acidic effluents in Tharsis and La Zarza-Perrunal mines (Iberian Pyrite Belt, Huelva, Spain). *Environ Geol* **49**: 253-266.

Sand, W., Gehrke, T., Jozsa, P. G. and Schippers, A. (2001). (Bio)chemistry of bacterial leaching-direct vs. indirect bioleaching. *Hydrometallurgy* **59**: 159-175.

Santini, J. M., Sly, L. I., Schnagl, R. D. and 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.

Sanz, J. L., Rodríguez, N., Díaz, E. E. and Amils, R. (2011). Methanogenesis in the sediments of Río Tinto, an extreme acidic river. *Environ Microbiol* **13**: 2336-2341.

Schadt, E. E., Turner, S. and Kasarskis, A. (2010). A window into third-generation sequencing. *Hum Mol Genet* **19**: R227-R240.

Schippers, A. and Sand, W. (1999). Bacterial leaching of metal sulfides proceeds by two indirect mechanisms via thiosulfate or via polysulfides and sulfur. *Appl Environ Microbiol* **65**: 319-321.

Schippers, A., Breuker, A., Blazejak, A., Bosecker, K., Kock, D. and Wright, T. L. (2010). The biogeochemistry and microbiology of sulfidic mine waste and bioleaching dumps and heaps, and novel Fe(II)-oxidizing bacteria. *Hydrometallurgy* **104**: 342-350.

Schneider, T. and Riedel, K. (2010). Environmental proteomics: analysis of structure and function of microbial communities. *Proteomics* **10**: 785-798.

Sehlin, H. M. and Lindström, E. B. (1992). Oxidation and reduction of arsenic by *Sulfolobus acidocaldarius* strain BC. *FEMS Microbiol Lett* **93**: 87-92.

Shen, Y. and Smith, R. D. (2002). Proteomics based on high-efficiency capillary separations. *Electrophoresis* **23**: 3106-3124.

Shendure, J. and Ji, H. (2008). Next-generation DNA sequencing. *Nat Biotechnol* **26**: 1135-1145.

Siggins, A., Gunnigle, E. and Abram, F. (2012). Exploring mixed microbial community functioning: recent advances in metaproteomics. *FEMS Microbiol Ecol* **80**: 265-280.

Silberbach, M., Huser, A., Kalinowski, J., Puhler, A., Walter, B., Kramer, R. and Burkovski, A. (2005). DNA microarray analysis of the nitrogen starvation response of *Corynebacterium glutamicum*. *J Biotechnol* **119**: 357-367.

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

Silverman, M. P., Ehrlich, H. L. and Wayne, W. U. (1964). Microbial formation and degradation of minerals. *Adv Appl Microbiol* **6**: 153-206.

Singer, P. C. and Stumm, W. (1970). Acidic mine drainage: the rate-determining step. *Science* **167**: 1121-1123.

Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., Huse, S. M., Neal, P. R. et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci U S A* **103**: 12115-12120.

Souza-Egipsy, V., Altamirano, M., Amils, R. and Aguilera, A. (2011). Photosynthetic performance of phototrophic biofilms in extreme acidic environments. *Environ Microbiol* **13**: 2351-2358.

Stoeck, T., Hayward, B., Taylor, G. T., Varela, R. and Epstein, S. S. (2006). A multiple PCR-primer approach to access the microeukaryotic diversity in environmental samples. *Protist* **157**: 31-43.

Stolz, J. F. and Oremland, R. S. (1999). Bacterial respiration of arsenic and selenium. *FEMS Microbiol Rev* **23**: 615-627.

Stolz, J. F. and Basu, P. (2002). Evolution of nitrate reductase: molecular and structural variations on a common function. *ChemBiochem* **3**: 198-206.

Stolz, J. F., Basu, P., Santini, J. M. and Oremland, R. S. (2006). Arsenic and selenium in microbial metabolism. *Annu Rev Microbiol* **60**: 107-130.

Stumm, W. and Morgan, J. (1981). Aquatic chemistry: An introduction emphasizing chemical equilibria in natural waters. New York: John Wiley & Sons. pp.1-780

Sultan, M., Schulz, M. H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M. et al. (2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**: 956-960.

T

Tamaki, S. and Frankenberger, W. T. J. (1992). Environmental biochemistry of arsenic. *Rev Environ Contam Toxicol* **124**: 79-110.

Thevenieau, F., Fardeau, M. L., Ollivier, B., Joulian, C. and Baena, S. (2007). *Desulfomicrobium thermophilum* sp. nov., a novel thermophilic sulphate-reducing bacterium isolated from a terrestrial hot spring in Colombia. *Extremophiles* **11**: 295-303.

Treusch, A. H., Leininger, S., Kletzin, A., Schuster, S. C., Klenk, H. P. and Schleper, C. (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ Microbiol* **7**: 1985-1995.

Tributsch, H. (2001). Direct versus indirect bioleaching. *Hydrometallurgy* **59**: 177-185.

Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M. et al. (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. and 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.

V

Velculescu, V. E., Vogelstein, B. and Kinzler, K. W. (2000). Analysing uncharted transcriptomes with SAGE. *Trends Genet* **16**: 423-425.

Vincent, W. F. and James, M. R. (1996). Biodiversity in extreme aquatic environments: Lakes, ponds and streams of the Ross Sea sector, Antarctica. *Biodivers Conserv* **5**: 1451-1471.

W

Wang, Z., Gerstein, M. and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**: 57-63.

Weckwerth, W. (2003). Metabolomics in systems biology. *Annu Rev Plant Biol* **54**: 669-689.

White, C. and Gadd, G. M. (1996). Mixed sulphate-reducing bacterial cultures for bioprecipitation of toxic metals: factorial and response-surface analysis of the effects of dilution rate, sulphate and substrate concentration. *Microbiology* **142**: 2197-2205.

Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* **95**: 6578-6583.

Widdel, F. (1988). Microbiology and ecology of sulfate- and sulfur-reducing bacteria. In *The Biology of Anaerobic Microorganisms*. (Zehnder Eds). New York: pp. 469-586.

Wilmes, P. and Bond, P. L. (2006). Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends Microbiol* **14**: 92-97.

Wolff, S., Otto, A., Albrecht, D., Zeng, J. S., Buttner, K., Gluckmann, M. et al. (2006). Gel-free and gel-based proteomics in *Bacillus subtilis*: a comparative study. *Mol Cell Proteomics* **5**: 1183-1192.

Wu, S., Lourette, N. M., Tolic, N., Zhao, R., Robinson, E. W., Tolmachev, A. V. et al. (2009). An integrated top-down and bottom-up strategy for broadly characterizing protein isoforms and modifications. *J Proteome Res* **8**: 1347-1357.

X

Xie, J., Fu, Y., Jiang, D., Li, G., Huang, J., Li, B. et al. (2008). Intergeneric transfer of ribosomal genes between two fungi. *BMC Evol Biol* **8**: 87.

Xu, Y., Axe, L., Boonfueng, T., Tyson, T. A., Trivedi, P. and Pandya, K. (2007). Ni(II) complexation to amorphous hydrous ferric oxide: an X-ray absorption spectroscopy study. *J Colloid Interface Sci* **314**: 10-17.

Y

Yamanaka, K. and Okada, S. (1994). Induction of lung-specific DNA damage by metabolically methylated arsenics via the production of free radicals. *Environ Health Perspect* **102 Suppl 3**: 37-40.

Z

Zhang, G. and Fang, Z. (2005). The contribution of direct and indirect actions in bioleaching of pentlandite. *Hydrometallurgy* **80**: 59-66.

Zhang, W., Li, F. and Nie, L. (2010). Integrating multiple 'omics' analysis for microbial biology: application and methodologies. *Microbiology* **156**: 287-301.

Zirnstein, I., Arnold, T., Krawczyk-Barsch, E., Jenk, U., Bernhard, G. and Roske, I. (2012). Eukaryotic life in biofilms formed in a uranium mine. *Microbiologyopen* **1**: 83-94.

Liste des figures

Figure 1 : Arbre phylogénétique du vivant simplifié (López-García & Moreira, 2008)	6
Figure 2 : Cycle du carbone dans l'environnement (Prescott <i>et al.</i> , 2010).....	8
Figure 3 : Cycle de l'azote. Les flux se produisant principalement en conditions aérobies correspondent aux flèches vides. Les processus anaérobies correspondent aux flèches pleines. Les processus se déroulant dans les deux conditions aérobies et anaérobies sont représentés par des flèches hachurées (Prescott <i>et al.</i> , 2010).....	10
Figure 4 : Cycle du soufre. Les flèches bleues représentent des processus cataboliques, les flèches hachurées représentent des processus anaboliques (Prescott <i>et al.</i> , 2010)	11
Figure 5 : Modèle de l'oxydation de la pyrite (adaptation de (Stumm & Morgan, 1981)	16
Figure 6 : Diversité phylogénétique basée sur les séquences des gènes codant pour l'ARNr 16S des bactéries des drainages miniers acides (caractère normal) et des sites de biolixiviation (caractère gras). Les lignées anaérobies facultatives sont indiquées par * et les aérobies obligatoires par ϕ . IM = Iron Mountain (Baker & Banfield, 2003)	26
Figure 7 : Exemple d'eucaryotes isolés de DMA. (A) Flagellé du genre <i>Euglena</i> ; (B) amibes du genre <i>Vahlkampfia</i> ; (C) héliozoaire du genre <i>Actinophrys</i>	30
Figure 8 : Arbre phylogénétique basé sur les ADNr 18S des organismes eucaryotes du DMA du Río Tinto. Les espèces en gras représentent des clones environnementaux. RT : séquences des espèces identifiées sur le site du Río Tinto. Les taxons soulignés représentent les espèces identifiés par observations microscopiques (Amaral-Zettler <i>et al.</i> , 2002).....	31
Figure 9 : Approches systémiques pour la caractérisation d'écosystèmes microbiens (Siggins <i>et al.</i> , 2012).....	41
Figure 10 : Résumé des principales étapes de séquençage par la technologie 454 (adaptée de Mardis, 2008)	43
Figure 11 : Principales étapes de séquençage par la technologie illumina (adaptée de Mardis, 2008).....	45
Figure 12 : Etapes relatives au séquençage par la technologie SOLiD. (a) Séquençage par ligation. (b) Principe des deux bases codantes (adaptée de Metzker, 2010).....	48

LISTE DES FIGURES

- Figure 13 :** Localisation géographique de la mine de Carnoulès 63
- Figure 14 :** Topologie du DMA de Carnoulès. En rouge sont indiqués les sites d'échantillonnages. S5 : au centre du stérile minier où a été placé un piézomètre. S1 : à la source du Reigous. COWG : à 30 m de la source. GAL : à 150 m de la source. CONF : à 1,5 km de la source, juste avant la confluence entre le Reigous et l'Amous 64

Liste des tableaux

Tableau 1 : Principaux minéraux sulfurés	13
Tableau 2 : Caractéristiques physicochimiques de quelques sites miniers répartis dans le monde (adapté de Johnson & Hallberg, 2003; Hallberg, 2010).....	24
Tableau 3 : Bactéries détectées au sein de drainages miniers acides.....	27

Liste des abréviations

ADN	Acide désoxyribonucléique
AOA	Ammonia-Oxidizing Archaea
AOB	Ammonia-Oxidizing Bacteria
ARMs	Arsenate-Resistant Microbes
ARN	Acide ribonucléique
BET	Bromure d'éthidium
BLAST	Basic Local Alignment Search Tool
BSR	Bactérie Sulfato-Réductrice
CAOs	Chemoautotrophic Arsenite-Oxidizers
CCA	Canonical Correspondance Analysis
DAR	Drainage Acide de Roche
DARPs	Dissimilatory Arsenate-Respiring Prokaryotes
DMA	Drainage Minier Acide
dNTP	Désoxyribonucléotide triphosphate
EDTA	Ethylène diamine tétraacétate
ETM	Eléments Traces Métalliques
HOAs	Heterotrophic Arsenite Oxidizers
Kb	Kilobases
kDa	Kilodalton
Mb	Megabases
NCBI	National Center for Biotechnology Information
NGS	Next-Generation Sequencing
OMS	Organisation Mondiale de la Santé
OTU	Operational Taxonomic Unit
pb	Paire de bases
PCR	Polymerase Chain Reaction
ppm	Partie par million
RDP	Ribosomal Database Project
RFLP	Restriction Fragment Length Polymorphism
SRB	Sulfate-Reducing Bacteria
T-RF	Terminal-Restriction Fragment
T-RFLP	Terminal-Restriction Fragment Length Polymorphism

Liste des symboles

Al	Aluminium
Ag	Argent
As	Arsenic
As(III)	Arsénite
As(V)	Arséniate
Bi	Bismuth
C	Carbone
Ca	Calcium
Cd	Cadmium
Co	Cobalt
Ni	Nickel
Cr	Chrome
Cu	Cuivre
Fe	Fer
Fe(II)	Fer ferreux
Fe(III)	Fer ferrique
Hg	Mercure
Mg	Magnésium
Mo	Molybdène
N	Azote
Na	Sodium
Pb	Plomb
S	Soufre
Sb	Antimoine
Sb(III)	Antimonite
Sb(V)	Antimoniate
Se	Sélénium
Se(IV)	Sélénite
Se(VI)	Séléniate
Sn	Etain
Te	Tellure
Tl	Thallium
V	Vanadium
Zn	Zinc